GEOMICROBIOLOGY OF HYDROGEN

IN YELLOWSTONE HOT SPRINGS

by

Melody Rose Lindsay

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Microbiology

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2019

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor, Dr. Eric Boyd for everything; his time, effort, advising, guidance, and patience. I thank him for guiding me through my graduate work and for teaching me to be the best scientist I can be.

I would also like to thank my other committee members, past and present; including Dr. John Peters, Dr. Dan Colman, Dr. Mark Skidmore, and Dr. Mark Young. I also extend my sincere gratitude to all the collaborators I've had the privilege to work with: Dr. John Spear, Dr. Everett Shock, and Dr. Tori Hoehler for great research in Yellowstone, and the Ladies of the Lake GSLI crew Dr. Bonnie Baxter and Jaimi Butler.

I also thank my labmates/friends, past and current for always being awesome: Max, Matt, Jayme, Saroj, Dan, Erik A., Eric D., Libby, and Devon. I would also like to thank the MBI staff, especially Ileana and Lisa, for all the help over the years. All the thanks to my mom, Dr. Holly Lindsay and my brother, Christopher for all the support and love during my life. Many thanks to my friends both far and near who supported me, sent cat/dog pictures, and video chatted, especially Karis, Kirsten, Kyle, James, and Liz. I would also like to thank my best friend and honorary field assistant David S. Pumpkins who always comes running with happy meows after every long day.

Finally, I am also very thankful to the NASA Earth and Space Science Fellowship for supporting my graduate research, and to other sources of funding, including the TBI Turner Student Award, the Beverly Ferguson Graduate Student Award from the MBI Department, and the Doyle W. Stephens Award from Friends of Great Salt Lake.

TABLE OF CONTENTS

1. GENERAL INTRODUCTION....1

Hydrogen in Microbial Metabolism....3
Diversification of Hydrogenases....6
Hydrogen Metabolism in Hydrothermal Environments....9
Overarching Goals....12
References....15

2. GEOLOGICAL SOURCE AND BIOLOGICAL FATE OF HYDROGEN IN YELLOWSTONE HOT SPRINGS....22

Contribution of Authors and Co-Authors....22
Manuscript Information Page....24
Abstract....25
Introduction....26
Materials and Methods....28
Determination of hydrogenase protein encoding gene homolog abundance in chemosynthetic community metagenomes....28
Physical and chemical measurements....30
DNA extraction....31
PCR, quantitative PCR, and sequencing....31
Dilution to extinction cultivation assays....32
Metagenomic sequencing and analysis....33
Identification and classification of [NiFe]- and [FeFe]-hydrogenase homologs in the SJ3 metagenome....35
RNA extraction, cDNA synthesis, and quantification of gene expression....36
Results and Discussion....38
Geological sources of H_2 in Yellowstone hot springs....38
Biological fate of hydrogen in Yellowstone hot springs....40
Conclusions....50
Acknowledgments....51
Figures....52
Supplemental Tables....58
Supplemental Figures....62
Additional Supplemental Datasets....68
References....69

3. SUBSURFACE PROCESSES INFLUENCE OXIDANT AVAILABILITY AND CHEMOAUTOTROPHIC HYDROGEN METABOLISM IN YELLOWSTONE HOT SPRINGS....76

Contribution of Authors and Co-Authors....76
Manuscript Information Page....78

TABLE OF CONTENTS CONTINUED

Abstract 79
Introduction 80
Materials and Methods 83
Description of sample sites and sample collection 83
Physical and chemical measurements 84
X-Ray diffraction (XRD) 87
RNA extraction and cDNA synthesis 88
PCR, qPCR, and sequencing 88
Dissolved inorganic carbon assimilation assays 92
Hydrogen transformation assays 93
Dilution to extinction cultivation assays 94
Results 97
Geochemistry 97
Mineralogy 100
DIC and dissolved CO_2 assimilation 100
H_2 transformation 101
Abundance of SSU rRNA gene transcripts 102
Sediment community taxonomic diversity and composition 103
Archaea 103
Bacteria 104
Eukarya 105
Dilution to extinction cultivation assays 106
Discussion 107
Conclusions 118
Acknowledgments 119
Table 120
Figures 121
Supplemental Tables 130
Supplemental Figures 139
References 142

4. CONCLUSIONS AND FUTURE DIRECTIONS 150

Conclusions 150
Future Directions 154
References 156

CUMULATIVE REFERENCES 157

APPENDICES 174

TABLE OF CONTENTS CONTINUED

APPENDIX A: MIXING OF END-MEMBER FLUIDS SUPPORTS HYPERDIVERSE CHEMOSYNTHETIC HYDROTHERMAL COMMUNITIES 175

APPENDIX B: COPIES OF PERMISSIONS TO REPRINT 225

LIST OF TABLES

Table Page

2.1. Supplemental Table 1. Table of primers designed and used to quantify transcript abundances for target gene homologs 58

2.2. Supplemental Table 2. Geochemical measurements and 16S rRNA gene abundances (normalized to $gdws^{-1}$) in waters and sediments from Smokejumper springs.....59

2.3. Supplemental Table 3. Approximate number of autotrophic cells capable of coupling H_2 oxidation to the reduction of specified oxidants (electron acceptors) as measured by dilution series cultivation experiments.....60

2.4. Supplemental Table 4. Metagenome sequence statistics for the SJ3 community 61

3.1. Approximate number of autotrophic cells capable of coupling H_2 oxidation to the reduction of specified oxidants (electron acceptor) as measured by dilution series cultivation experiments.....120

3.2. Supplemental Table 1. Geochemical parameters measured in waters from sampling locations.....130

3.3. Supplemental Table 2. Geochemical parameters measured in waters from sampling locations, all normalized to exclude any concentrations due to evaporative concentration 131

3.4. Supplemental Table 3. Concentrations of dissolved methane, carbon dioxide, and hydrogen as measured along the outflow channels in RSE and RSW.....132

3.5. Supplemental Table 4. Minerals detected by XRD in sediments sampled along the outflow channels of RSE and RSW.....133

3.6. Supplemental Table 5. Rates of microbial activities 134

3.7. Supplemental Table 6. Predicted sequence coverage and inverse Simpson indices of SSU rRNA gene sequences from sediment communities at RSE and RSW sampling sites 135

LIST OF TABLES CONTINUED

Table Page

3.8. Supplemental Table 7. Identities of representative 16S rRNA gene sequences in the enrichment transfers of the most dilute, dilution-to-extinction cultivation experiments conducted on RSE sediments 136

3.9. Supplemental Table 8. Identities of representative 16S rRNA gene sequences in the enrichment transfers of the most dilute, dilution-to-extinction cultivation experiments conducted on RSW sediments 137

3.10. Supplemental Table 9. Percent relative abundances of 16S (Archaea and Bacteria) and 18S (Eukarya) rRNA gene transcript OTUs (as represented by their closest cultivated relatives) found in sites at RSE and RSW 138

LIST OF FIGURES

Figure Page

2.1. Map of Yellowstone National Park (YNP) 52

2.2. Schematic summarizing a conceptual model of interactions between subsurface crustal minerals and water leading to production of gas in hydrothermal fluids, developed based on geochemical and isotopic proxies 53

2.3. Prevalence of hydrogenase protein encoding gene homologs among 50 chemosynthetic community metagenomes from YNP hot springs plotted as a function of spring pH 54

2.4. Concentrations of dissolved gases, 16S rRNA genes, and cultivatable hydrogenotrophic and autotrophic cells in SJ hot springs 55

2.5. Rank abundance plot (log transformed) of genome bins in SJ3 and the distribution of hydrogenase isoforms in these bins 56

2.6. Abundance of gene transcripts associated with dominant SJ3 metagenomic bins 57

2.7. Supplemental Figure 1. H_2 (mol%) in gas samples from 130 springs YNP plotted as a function of the elevation (m) in those springs 62

2.8. Supplemental Figure 2. Helium, methane, and hydrogen in springs across YNP 63

2.9. Supplemental Figure 3. Sulfate/chloride concentrations (mg L^{-1}), water isotope composition (‰), and spring classifications for three Smokejumper hot springs 64

2.10. Supplemental Figure 4. Taxonomic composition of 16S rRNA gene OTUs in sediments from SJ hot springs 65

2.11. Supplemental Figure 5. Emergent self-organizing map (ESOM) of assembled genome contigs 66

2.12. Supplemental Figure 6. Multiple sequencing alignment of H-cluster binding motifs for the large (catalytic) subunit of [FeFe]-hydrogenase (HydA) homologs from SJ3 bins 67

LIST OF FIGURES CONTINUED

Figure Page

3.1. Location of sampling sites along the outflow channels at Roadside East and Roadside West hot springs. 121

3.2. Plots of geochemical parameters, concentrations of ions, and concentrations of dissolved gases measured down the outflow channels of RSE and RSW 122

3.3. Dissolved inorganic carbon uptake in sediment-associated microbial communities sampled along the outflow channels of select RSE and RSW sampling locations....124

3.4. Rates of net H_2 transformation in sediment-associated microbial communities in selected sites along the RSE and RSW hot spring outflow channels 125

3.5. Abundances of 16S (Archaea, Bacteria) and 18S (Eukarya) rRNA gene transcripts associated with sediments sampled along the outflow channels of RSE and RSW hot springs as determined using qPCR on cDNA 126

3.6. Relative abundances of 16S (Archaea and Bacteria) and 18S (Eukarya) rRNA gene transcript OTUs as a function of distance along the outflow channels of RSE and RSW 127

3.7. Schematic summarizing the influence of subsurface and surface processes on oxidant availability and use by hydrogenotrophic and autotrophic populations 128

3.8. Supplemental Figure 1. SO_4^{2-} concentrations (mg L^{-1}) in hot spring waters plotted relative to Cl^- concentrations (mg L^{-1}) in the same waters 139

3.9. Supplemental Figure 2. The isotopic composition of hydrogen (δD) plotted relative to the isotopic composition of oxygen ($\delta^{18}O$) in the same waters 140

LIST OF FIGURES CONTINUED

Figure Page

3.10. Supplemental Figure 3. YNP hot spring samples (442 total) plotted as a function of spring water pH and temperature, or total sulfide concentration (mg L^{-1}) and temperature, indicating presence or absence of phototrophy.....141

ABSTRACT

Hydrogen (H_2) connects the geosphere and biosphere in rock-hosted ecosystems and has likely done so since early in Earth's history. High temperature hydrothermal environments, such as hot springs, can be enriched in H_2 and were likely widespread on early Earth. As such, linking the geological processes that supply H_2 to contemporary hot springs and the distribution of extant thermophilic organisms that can utilize H_2 as a component of their energy metabolism can provide insights into the environment types that supported early H_2 dependent life. Using a series of geochemical proxies, I developed a model to describe variable H_2 concentrations in Yellowstone National Park (YNP) hot springs. The model invokes interaction between water and crustal minerals that generates H_2 that can partition into the vapor phase during decompressional boiling of ascending hydrothermal waters. Fractures and faults in bedrock, combined with topographic features such as high elevation, allow for vapor to migrate and concentrate in certain areas of YNP leading to elevated concentrations of H_2. Metagenomes from chemosynthetic communities in YNP springs sourced with vapor-phase gas are enriched in genes coding for enzymes predicted to be involved in H_2-oxidation. A spring in an area of YNP (Smokejumper, SJ3) sourced with vapor-phase gas, that has the highest concentration of H_2 measured in YNP, and that is enriched in hydrogenase encoding genes was chosen to further examine the biological fate of H_2. SJ3 harbors a hyperdiverse community that is supported by mixing of oxidized meteoric fluids and volcanic gases. Transcripts coding for genes involved in H_2 uptake and CO_2 fixation were detected. The processes that control the availability of oxidants and their effect on the activity and abundance of H_2 dependent organisms was also investigated in two paired hot springs. H_2-oxidizing chemoautotrophs utilized different oxidants in the two springs and this underpinned differences in H_2 oxidation activity and their identity. Together, these observations indicate that the subsurface geological processes of decompressional boiling and phase separation influence the distribution, identity, and activity of hydrogenotrophs through their combined effects on the availability of H_2 and oxidants.

CHAPTER ONE

GENERAL INTRODUCTION

Hydrogen (H_2) links the biosphere and the geosphere and likely has done so since life first emerged on Earth. The rate of H_2 release from the subsurface of early Earth is predicted to have exceeded the rate of H_2 escape from the atmosphere, resulting in a predicted mixing ratio of ~30% H_2 (the balance between slow hydrogen escape and volcanic outgassing) in the primitive atmosphere of early Earth (Tian et al 2005). H_2 is likely to have been available on early Earth and could have contributed to energy-yielding reactions that supported early life (Wächtershäuser 1990). Indeed, H_2 has been hypothesized to be one of the oldest and most widely utilized reductants in microbial metabolism, and early anaerobic organisms could have reduced volcanic carbon dioxide (CO_2) with electrons from H_2 similar to reactions that methanogens and acetogens use in contemporary environments (Martin 2012).

The enzymes that catalyze the reversible oxidation of H_2 are termed hydrogenases, and hydrogenases are encoded in the genomes of taxa representing all three domains of life (Martin and Müller 1998, Peters et al 2015). This includes many deeply rooted lineages in the archaeal and bacterial domains (Peters et al 2015, Vignais et al 2001). Enzymes from one class of hydrogenase, ([NiFe]-hydrogenase), are present in both Bacteria and Archaea, suggesting that [NiFe]-hydrogenase likely emerged prior to the divergence of these domains (Boyd et al 2014, Weiss et al 2016). Further, this suggests that H_2, through the activity of [NiFe]-hydrogenase, could have supported the metabolism of the Last Universal Common Ancestor (LUCA).

Another suggested property of the LUCA is thermophily and this is based primarily on taxa from hydrothermal environments forming deeply (basal) branches on phylogenetic reconstructions of life that are short, indicating that contemporary thermophiles are less derived from their ancestors when compared to mesophilic lineages (Pace 1991, Weiss et al 2016). Consistent with the possibility that life originated as a thermophile, geological evidence indicates that hydrothermal environments were present if not widespread on early Earth (Djokic et al 2016). Given the prevalence of lithogenic H_2 in hydrothermal systems today (Adam and Perner 2018, Spear et al 2005), evidence for the presence of hydrothermal systems on early Earth (Djokic et al 2016), and the hypothesized H_2 dependence and thermophilic character of LUCA (Weiss et al 2016), it is perhaps reasonable to suspect H_2 could have supported early evolving organisms inhabiting hydrothermal environments through the activity of ancestral [NiFe]-hydrogenases (Elderfield and Schultz 1996, Neal and Stanger 1983). Thus, contemporary hydrothermal systems are likely to be useful analogs to study the processes that support modern day descendants of early evolving microbial life.

Over time, the escape of H_2 from the atmosphere (Hunten 1973, Tian et al 2005) and the rise of oxygenic photosynthesis (Anbar 2008, Lyons et al 2014) led to the oxidation of Earth's surface environment (Catling et al 2001). Nonetheless, on modern Earth, H_2 is still abundant locally including in hydrothermal environments such as the hot springs in Yellowstone National Park (YNP), Wyoming, U.S.A. (Bergfeld et al 2014, Lowenstern et al 2015, Spear et al 2005, Windman et al 2007). As such, reactions involving H_2 have been hypothesized to be an important source of energy supporting primary production in YNP hot spring communities, although others are present as well

(Shock et al 2010, Spear et al 2005). Additionally, analyses of chemosynthetic microbial communities have indicated that several taxa hypothesized to be involved in biological H_2 cycling are abundantly distributed across YNP hot springs (Spear et al 2005). However, other studies challenge the importance of H_2 in supporting thermophilic hot spring communities (D'Imperio et al 2008, Inskeep et al 2005). Together, these observations warrant further study to better understand the processes that supply H_2 to hydrothermal environments and how these processes present and influence the distribution of H_2-dependent organisms.

Hydrogen in Microbial Metabolism

The oxidation of a molecule of H_2 yields two protons and two electrons (Reaction 1), both of which could have been involved in primitive energy conservation mechanisms employed by early life (Hoehler 2005).

$$H_2 \leftrightarrow 2\ H^+ + 2\ e^- \qquad \text{Reaction 1}$$

Indeed, H_2 metabolism has been suggested to be among the oldest reactions (such as used in microbial metabolism (Martin 2012). The directionality of the reversible oxidation of H_2 reaction is dependent on concentrations and redox potentials of substrates, products, and electron carriers (e.g., ferredoxin (Fd) or NAD(P)H) and catalytic biases associated with the various types of enzymes that catalyze this redox transformation (Vignais and Colbeau 2004). Importantly, the intracellular and extracellular concentrations of H_2 also impact the thermodynamic favorability of any H_2 associated metabolic strategy (Hoehler et al 2002), as cells rely on intracellular H_2 partial pressures which in turn are strongly

affected by extracellular H_2 levels due to the permeability of lipid membranes to this nonpolar gas (Atkins 1990, Marrink and Berendsen 1996).

The enzymes that catalyze the reversible oxidation of H_2, called hydrogenases, are present in all three domains of life and are associated with a wide variety of metabolic strategies (Vignais et al 2001, Vignais and Colbeau 2004). They are encoded in the genomes of many taxa that are deeply rooted in phylogenetic reconstructions, particularly among Bacteria and Archaea (including methanogens and other taxa such as Aquificae and Thermotogae) (Peters et al 2015, Vignais et al 2001). This observation is often used to suggest the microbial use of H_2 could have been an early biological innovation, and potentially a property of the LUCA (Weiss et al 2016).

Overall, H_2 and hydrogenases are associated with a wide variety of microbial metabolisms, including those of anaerobes and aerobes, autotrophs and heterotrophs, and anoxygenic and oxygenic phototrophs (Hoehler et al 2002, Hoehler 2005). There are three main types of hydrogenases that are termed [FeFe]-hydrogenases, [NiFe]-hydrogenases, and [Fe]-only (Hmd) hydrogenases which are differentiated phylogenetically and by the metallic contents of their active site metal clusters (Greening et al 2016, Peters et al 2015, Thauer 1998, Vignais et al 2001). The Hmd hydrogenases are widespread in methanogens and contain an iron-containing cofactor but they lack iron-sulfur (Fe-S) clusters that participate in electron transfer, and do not formally catalyze the reversible oxidation of H_2 (Vignais and Colbeau 2004). In contrast, both [FeFe]- and [NiFe]- hydrogenases contain a dinuclear metal center (either [FeFe] or [NiFe]) that is the site of the H_2 reaction, as well as accessory Fe-S clusters (Nicolet et al 1999, Peters et al 1998, Peters et al 2015, Shafaat et al 2013).

[FeFe]-hydrogenases have been identified among numerous fermentative and anaerobic Bacteria, as well as several unicellular eukaryotes (Meyer 2007, Peters et al 2015). [FeFe]-hydrogenase likely evolved to couple the production of H_2 with oxidation of electron carriers (e.g., Fd_{red} and NAD(P)H) that accumulate during anaerobic metabolism, although they have also been shown to oxidize H_2 (Adams et al 1989, Therien et al 2017). In contrast, [NiFe]-hydrogenases are widely distributed in both archaeal and bacterial taxa, but have yet to be identified among eukaryotes (Boyd et al 2014, Peters et al 2015). Additionally, many subtypes (groups) of [NiFe]-hydrogenase catalyze H_2 oxidation for energy metabolism, although they can also be bidirectional (reversible) or involved in the production of H_2 (Greening et al 2016, Peters et al 2015, Vignais et al 2001). [NiFe]-hydrogenases involved in the oxidation of H_2 are often coupled metabolically via electron carriers such as Fd and NAD(P)H, to the reduction of oxidants such as sulfate, metals, nitrate, and fumarate [for a recent review, see (Greening et al 2016)]. They can also be involved in H_2 oxidation coupled with the reduction of O_2, as well as reactive oxygen species (Boga and Brune 2003, Trembley and Lovley 2012).

[NiFe]- and [FeFe]-hydrogenases share many common biochemical and mechanistic features, including the active site ligands carbon monoxide and cyanide (Peters et al 2015). Remarkably, despite these similarities, [FeFe]- and [NiFe]-hydrogenases are not evolutionarily related and do not share sequence homology (Vignais et al 2001). Thus, [NiFe]- and [FeFe]-hydrogenases convergently evolved to catalyze the same chemistry: the reversible oxidation of H_2 (Reaction 1). However, unlike [NiFe]-hydrogenases that are widespread among both archaeal and bacterial genomes and were likely a property of the LUCA (Boyd et al 2014), [FeFe]-hydrogenases are not

present in Archaea and it is thus likely that it is likely that [FeFe]-hydrogenase evolved after the divergence of Archaea and Bacteria (Mulder et al 2010).

Diversification of Hydrogenases

Methanogens encode numerous [NiFe]-hydrogenase homologs (Thauer et al 2010). Phylogenetic reconstructions of representatives of each of these enzyme groups reveals a basal branch position relative to homologs from other taxa, indicating that these evolved early in the history of hydrogenases. This observation, combined with the presence of multiple [NiFe]-hydrogenase enzymes in methanogens, suggests that certain different isoforms (phylogenetically related and H_2-oxidizing group 3a/3c and group 4 [NiFe]-hydrogenases) could have arisen through a series of gene duplication events from an ancestoral enzyme complex (Boyd et al 2014). As described below, [NiFe]-hydrogenase in methanogens are either physiologically biased towards H_2 oxidation (Group 3a and 3c) or are bidirectional (Group 4 Eha/Ehb). Collectively, these observations led to the suggestion that the earliest organisms to oxidize H_2 were methanogens and that these could have likely functioned to reduce Fd (ferredoxin; an electron acceptor) in early cells (Boyd et al 2014).

Several types of [NiFe]-hydrogenases are encoded by modern hydrogenotrophic methanogens, including group 3a and group 3c enzymes (Thauer et al 2010). The group 3a [NiFe]-hydrogenases [8-hydroxy-5-deazaflavin (F_{420}) reducing] couple the oxidation of H_2 with the reduction of F_{420} that then serves as an electron carrier during methanogenesis (de Poorter et al 2005, Vignais and Billoud 2007). The group 3c [NiFe]-hydrogenases (Mvh) couple the oxidation of H_2 to the concurrent exergonic reduction of

heterodisulfide and Fd (Kaster et al 2011, Thauer et al 2010) in what has been termed electron bifurcation (Buckel and Thauer 2013, Peters et al 2015). Finally, some methanogen species also encode representatives of the group 4 membrane bound [NiFe]-hydrogenases, termed Eha/Ehb. These reversible enzymes couple ion translocation to the oxidation of H_2 thereby allowing cells to overcome the unfavorable thermodynamics associated with reduction of Fd (Lie et al 2012, Porat et al 2006).

Following an early emergence in the earliest forms of life on Earth, [NiFe]-hydrogenases then diversified, likely due to selective pressure to improve energy yields associated with H_2 oxidation through metabolic coupling to oxidants with increasingly more positive electrochemical potentials (Boyd et al 2014, Schut et al 2016). One example of the diversification of [NiFe]-hydrogenases is seen in response to the increase in O_2 following the emergence of oxygenic photosynthesis (3.2 to 2.8 bya), which led to a gradual rise in O_2 concentrations after ~2.4 bya (Anbar 2008, Moore et al 2017). Production of O_2 allowed for an influx of sulfate (SO_4^{2-}) to oceans via enhanced O_2-dependent weathering of continental sulfide minerals (Canfield and Farquhar 2009) and allowed for the O_2-dependent oxidation of ammonia to generate nitrate (NO_3^-) (Garvin et al 2009, Godfrey and Falkowski 2009). Thus, it has been hypothesized that the emergence of hydrogenases that can metabolically couple H_2 oxidation with the reduction of O_2, NO_3^-, or SO_4^{2-} (group 1 [NiFe]-hydrogenase enzymes) could have diversified to more efficiently harvest the energy from reactions involving these oxidants (Boyd et al 2014). These observations suggest that group 1 [NiFe]-hydrogenase enzymes emerged after the emergence of oxygenic photosynthesis (Boyd et al 2014).

[NiFe]-hydrogenases that form phylogenetic group 2 [reviewed recently in (Greening et al 2016)] are involved in other metabolic processes, including H_2 sensing (Lenz and Friedrich 1998) and harvesting H_2 produced as a byproduct of nitrogen fixation (Bothe et al 1977, Masukawa et al 2002). However, since nitrogen fixation via the nitrogenase enzyme is not thought to be a property of the LUCA (Boyd et al 2011, Boyd and Peters 2013), the particular [NiFe]-hydrogenase enzymes which are closely linked with nitrogenases are not considered to be ancestral. Moreover, it is perhaps not logical to imagine that enzymes that function to sense H_2 would evolve prior to enzymes that can actually oxidize/produce H_2. Thus, group 2 enzymes are considered to be relatively more recently evolved (Boyd et al 2014).

[NiFe]-hydrogenases have also been suggested to represent the progenitors of modern respiratory complexes, and likely proceeded through a series of intermediate respiratory complexes that are still relic in anaerobic metabolism (Boyd et al 2014, Hedderich 2004, Schut et al 2013, Schut et al 2016). Numerous subunits that form the NADH quinone oxidoreductase complex (Nuo, or Complex I), which is an integral component of modern day respiratory chains including those involved in O_2 reduction, are evolutionarily related to those that comprise ion-translocating group 4 [NiFe]-hydrogenases (Hedderich 2004, Schut et al 2016). The earliest evolving member of this oxidoreductase family is a proton reducing (note these enzymes are reversible) membrane-bound hydrogenase (Mbh)-like ancestor and this likely diversified through a series of intermediates including a sulfur reducing respiratory complex (membrane-bound oxidoreductase; MBX) and a cofactor F_{420} reducing complex ($F_{420}H_2$:phenazine oxidoreductase; FPO), and ultimately Complex I (Schut et al 2016). Thus, similar to the

diversification of the group 1 [NiFe]-hydrogenases, the diversification of group 4 [NiFe]-hydrogenase and related respiratory complexes could have been driven by the evolutionary drive to maximize metabolic energy yield (Schut et al 2016).

Hydrogen Metabolism in Hydrothermal Environments

The maximum growth temperature of archaeal and bacterial taxa representing lineages that branch basal on phylogenetic reconstructions indicates that early evolving taxa and the ancestor of these taxa (LUCA) may have been a thermophile (Lineweaver and Schwartzman 2004, Pace 1991). Geological data suggests that hydrothermal systems may have been widespread on early Earth, with reports of laminated structures (resembling stromatolites) in a hydrothermal setting dated to 3.48 bya (Djokic et al 2016). This suggests that life could have inhabited hydrothermal environments early in Earth history and life itself may have even originated in such environments. Indeed, many of the abiotic synthesis reactions that form the basic building blocks of life, including amino acids, are thermodynamically favorable (energy yielding) under hydrothermal conditions (Canovas III 2016, Shock and Canovas 2010).

Prior to the advent of oxygenic photosynthesis ~ 3.2 to 2.8 Ga, life was likely dependent on chemical forms of energy (Anbar 2008, Moore et al 2017, Thauer et al 1977). Potential sources of energy for early life would have included reactions involving H_2, CO_2 and other inorganic substrates (including S^0, H_2S, and Fe^{2+} among others) available on early Earth (Kral et al 1998, Wächtershäuser 1990). These substrates are also typically available in modern hydrothermal environments (Ball et al 2006, Ball et al 2010, McCollom and Shock 1997, Nordstrom et al 2009, Shock et al 2010).

Hydrothermal environments on Earth today could host microbial life analogous to life on early Earth, as many contemporary hydrothermal environments provide many of the geochemical substrates which could support ancient metabolisms (i.e. H_2, CO_2) (Shock et al 2010).

Hydrothermal systems in YNP are among the most widely studied on Earth. This includes an extensive and published history of coupled geological, geochemical, and microbiological investigations (e.g. recent studies such as Amenabar et al 2018, Boyd et al 2007, Boyd et al 2010, Boyd et al 2012, Colman et al 2018, Hamilton et al 2012, Inskeep et al 2005, Inskeep et al 2013, Kozubal et al 2008, Kozubal et al 2012, Lowenstern et al 2015, Shock et al 2010, Spear et al 2005, among many others) conducted on hot springs that span a wide range of pH (1.5 - 10), temperature (up to 93˚C), and geochemical and mineralogical compositions (Allen and Day 1935). The >14,000 hydrothermal features in YNP are incredibly diverse in terms of temperature, pH, and geochemical composition (Nordstrom et al 2009), including wide variation in the availability of reductants and oxidants capable of supporting chemosynthetic metabolisms (Shock et al 2010).

One of the sources of reductant in hydrothermal systems (including those found in YNP) is H_2, although others are present (such as H_2S) (Shock et al 2010, Spear et al 2005, D'Imperio et al 2008) which can be generated by a variety of processes. This includes interaction between water and ferrous iron minerals (Kelley et al 2005, Stevens and McKinley 2000), the radiolysis of water (Lin et al 2005), reactions between mineral surface silicate radicals and water (Kita et al 1982, Telling et al 2015), and the activity of several biological enzymes (Boyd et al 2010, Peters et al 2015). Of particular importance

to hydrothermal systems, such as those found in YNP, are interactions between water and iron-bearing minerals present in igneous rocks (Reaction 2).

$$2H_2O + (FeO)_2(SiO_2) = 2H_2 + 2(FeO_{3/2}) + SiO_2 \qquad \text{Reaction 2}$$

Given that YNP hosts abundant iron-bearing rhyolites and also basalts (to a lesser extent) (Christiansen 2001), it has been postulated that H_2 could be produced in YNP through reaction 2 when hot water interacts with subsurface bedrock, similar to other basalts (Stevens and McKinley 2000). Indeed, previous studies have detected H_2 at elevated concentrations compared with atmospheric H_2 (Bergfeld et al 2014, Spear et al 2005, Windman et al 2007). Due to the prevalence of H_2 in hot springs, it was hypothesized that the energy derived from the oxidation of H_2 could support primary production in YNP springs (Spear et al 2005). However, evidence indicating that H_2 is oxidized and used to support primary production was not firmly established in this study (Spear et al 2005). Moreover, subsequent studies suggested that other reductants, such as hydrogen sulfide, might be more important in supporting microbial primary producers and preferentially utilized over H_2 in hot spring environments (D'Imperio et al 2008). Thus, the extent that H_2 supports primary production in YNP hot springs, and hot springs in general, is not well understood. Moreover, the link between geological processes that drive variation in the geochemical composition of springs, including the availability of H_2 and electron acceptors capable of being used to oxidize H_2, and the distribution of organisms capable of utilizing H_2 to fuel their energy metabolism is not known.

Overarching Goals

This works seeks to develop new understanding between the geological processes that supply H_2 to contemporary hot springs and the distribution of organisms that can utilize H_2 as a component of their energy metabolism. High temperature hot spring environments exclude photosynthesis and thus allow for the study of microbial life that is supported by chemical sources of energy (Boyd et al 2012, Brock 1967), similar to conditions that have been hypothesized for life on early Earth (Kral et al 1998, Martin 2012, Weiss et al 2016). Thus, in this work I focus on high temperature springs that host communities dependent on chemical energy. Throughout this work, I use a framework put forth by (Boyd et al 2012) which defined the habitat range of phototrophs as limited primarily by temperature, pH, and sulfide concentrations to drive the selection of hot spring locations and communities for further study. Within hot spring chemosynthetic communities, I pursued several integrated research questions to address the overarching hypothesis that subsurface geological processes influence the distribution of H_2-dependent organisms and the metabolic strategies that they use to support primary production.

Firstly, I aimed to elucidate where and how H_2 is produced in hot springs to determine why H_2 varies across hot spring environments and how this in turn influences the distribution and abundance of hydrogenotrophic organisms. In Chapter 2, titled "Geological Source and Biological Fate of Hydrogen in Yellowstone Hot Springs", I leverage the variability of H_2 across YNP thermal features to develop a model to describe processes that source H_2 to hot springs and that drive differences in H_2 availability. Using a series of geochemical proxies, I provide evidence that interaction between hot water

and bedrock minerals drives production of H_2 that is partitioned into a vapor phase during decompressional boiling of ascending hydrothermal waters. Differential input of vapor to hot springs, which is controlled by geological features such as faulting and fracturing along the caldera boundary in YNP as well as high elevation that promotes more complete phase separation, leads to variable input of H_2 to springs. To examine how this influences the distribution of organisms capable of integrating H_2 into their energy metabolism, I subjected 50 metagenomes from chemosynthetic communities inhabiting springs across YNP to comparative analyses to determine the relative abundance of genes coding for hydrogenases as a function of geochemical proxies (i.e. pH) for the extent of vapor phase input. I focused on a spring in the high elevation area of the Mirror Plateau that is located on the edge of the caldera: Smokejumper spring number 3 (SJ3) to more comprehensively examine the fate of H_2 in hot springs. This spring hosts the highest concentration of H_2 measured in YNP and is enriched in hydrogenase encoding genes. An integrated suite of geochemical, microbiological, and molecular assays were used to determine the extent that H_2 supports primary production in this spring.

Given that energy yields associated with H_2 oxidation are highly dependent on the oxidant that it is metabolically coupled with (Shock et al 2010), I next focused on determining how variation in the distribution of oxidants selects for different hydrogenotrophic and autotrophic populations and influences rates of their activities (Chapter 3). In a study titled "Subsurface Processes Influence Oxidant Availability and Chemoautotrophic Hydrogen Metabolism in Yellowstone Hot Springs", I examine microbially mediated H_2 cycling and characterize the composition of active microbial communities of two ecosystems subjected to steep gradients in available energy and

oxidants. Through these studies I aimed to determine how subsurface or surface processes influence the availability of oxidants in phase separated hot springs and whether such variation in oxidant availability impacts biological H_2 cycling in these ecosystems. In this work, I determined rates of H_2 cycling, rates of CO_2 fixation, abundances of hydrogenotrophic cells capable of metabolically coupling H_2 oxidation with variation oxidants and determined the active microbial community composition and their abundances. I conducted this work in the source pools of the two paired hot springs, Roadside East and Roadside West that are separated by ~20 meters, and also extended this analysis to the outflow channels to examine the effects of surface processes (i.e. infusion of O_2) on microbial communities and H_2 cycling.

Chapter 4 summarizes the findings discussed in Chapters 2 and 3. In this section, the major conclusions from each study are examined within the context of H_2 as a source of reductant supporting diverse microbial primary producers in hydrothermal systems as analogs for those that might have existed on early Earth. Furthermore, I discuss future research directions that could further our understanding of the role of H_2 in supporting primary production in these environments. This future work is specifically aimed at providing a direct link between H_2 oxidation activity, CO_2 fixation activity, and the identity and metabolism of primary producers in YNP hot springs.

References

Adam N, Perner M (2018). Microbially mediated hydrogen cycling in deep-sea hydrothermal vents. *Frontiers in Microbiology* **9:** 1-17.

Adams MWW, Eccleston E, Howard JB (1989). Iron-sulfur clusters of hydrogenase I and hydrogenase II of *Clostridium pasteurianum*. *Proceedings of the National Academy of Sciences of the United States of America* **86:** 4932-4936.

Allen ET, Day AL (1935). *Hot springs of the Yellowstone national park*. Carnegie Institution of Washington: Washington.

Amenabar MJ, Colman DR, Poudel S, Roden EE, Boyd ES (2018). Electron acceptor availability alters carbon and energy metabolism in a thermoacidophile. *Environmental Microbiology*. **20:** 2523-2537.

Anbar AD (2008). Elements and Evolution. *Science* **322:** 1481-1483.

Atkins PW (1990). Physical Chemistry. *Physical Chemistry*. Oxford University Press.

Ball JW, McCleskey RB, Nordstrom DK, Holloway JM (2006). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, 2003-2005. *United States Geological Survey Open-File Report.*

Ball JW, McCleskey RB, Nordstrom DK (2010). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, 2006-2008. *United States Geological Survey Open-File Report.*

Bergfeld D, Lowenstern JB, Hunt AG, Pat Shanks III WC, Evans WC (2014). Gas and Isotope Chemistry of Thermal Features in Yellowstone National Park, Wyoming. *USGS Scientific Investigations Report* **2011-5012**.

Boga HI, Brune A (2003). Hydrogen-dependent oxygen reduction by homoacetogenic bacteria isolated from termite guts. *Applied and Environmental Microbiology* **69:** 779-786.

Bothe H, Tennigkeit J, Eisbrenner G, Yates MG (1977). The hydrogenase-nitrogenase relationship in the blue-green alga *Anabaena cylindrica*. *Planta* **133:** 237-242.

Boyd ES, Jackson RA, Encarnacion G, Zahn JA, Beard T, Leavitt WD *et al* (2007). Isolation, characterization, and ecology of sulfur-respiring *Crenarchaea* inhabiting acid-sulfate-chloride-containing geothermal springs in Yellowstone National Park. *Applied and Environmental Microbiology* **73:** 6669-6677.

Boyd ES, Hamilton TL, Spear JR, Lavin M, Peters JW (2010). [FeFe]-hydrogenase in Yellowstone National Park: evidence for dispersal limitation and phylogenetic niche conservatism. *The ISME Journal* **4:** 1485-1495.

Boyd ES, Anbar AD, Miller SL, Hamilton TL, Lavin M, Peters JW (2011). A late methanogen origin for molybdenum-dependent nitrogenase. *Geobiology* **9:** 221-232.

Boyd ES, Fecteau KM, Havig JR, Shock EL, Peters JW (2012). Modeling the habitat range of phototrophs in yellowstone national park: toward the development of a comprehensive fitness landscape. *Frontiers in Microbiology* **3:** 221.

Boyd ES, Peters JW (2013). New insights into the evolutionary history of biological nitrogen fixation. *Frontiers in Microbiology* **4:** 1-12.

Boyd ES, Schut GJ, Adams MWW, Peters JW (2014). Hydrogen metabolism and the evolution of biological respiration. *Microbe* **9:** 361-367.

Brock TD (1967). Life at high temperatures. *Science* **158:** 1012-1019.

Buckel W, Thauer RK (2013). Energy conservation via electron bifurcating ferredoxin reduction and proton/Na^+ translocating ferredoxin oxidation. *Biochimica et Biophysica Acta* **1827:** 94-113.

Canfield DE, Farquhar J (2009). Animal evolution, bioturbation, and the sulfate concentrations of the oceans. *Proceedings of the National Academy of Science of the United States of America* **106:** 8123-8127.

Canovas III PA (2016). Energy transfer between the geosphere and biosphere, Arizona State University.

Catling DC, Zahnle KJ, McKay CP (2001). Biogenic methane, hydrogen escape, and the irreversible oxidation of early Earth. *Science* **293:** 839-843.

Christiansen RL (2001). The Quarternary and Pliocene Yellowstone Plateau Volcanic Field of Wyoming, Idaho, and Montana. USGS Professional Paper. USGS Information Services: Denver, CO.

Colman DR, Poudel S, Hamilton TL, Havig JR, Selensky MJ, Shock EL *et al* (2018). Geobiological feedbacks and the evolution of thermoacidophiles. *The ISME Journal* **12:** 225-236.

D'Imperio S, Leht CR, Oduro H, Druschel G, Kuhl M, McDermott TR (2008). Relative importance of H_2 and H_2S as energy sources for primary production in geothermal springs. *Applied and Environmental Microbiology* **74:** 5802-5808.

de Poorter LM, Geerts WJ, Keltjens JT (2005). Hydrogen concentrations in methane-forming cells probed by the ratios of reduced and oxidized coenzyme F_{420}. *Microbiology* **151:** 1697-1705.

Djokic T, Van Kranendonk MJ, Campbell KM, Walter MR, Ward CR (2016). Earliest signs of life on land preserved in ca. 3.5 Ga hot spring deposits. *Nature Communications* **8:** 15263.

Elderfield H, Schultz A (1996). Mid-ocean ridge hydrothermal fluxes and the chemical composition of the ocean. *Annual Review of Earth and Planetary Science* **24:** 191-224.

Garvin J, Buick R, Anbar AD, Arnold GL, Kaufman AJ (2009). Isotopic evidence for an aerobic nitrogen cycle in the latest Archaean. *Science* **323:** 1045-1048.

Godfrey LV, Falkowski PG (2009). The cycling and redox state of nitrogen in the Archaean ocean. *Nature Geoscience* **2:** 725-729.

Greening C, Biswas A, Carere CR, Jackson CJ, Taylor MC, Stott MB *et al* (2016). Genomic and metagenomic surveys of hydrogenase distribution indicate H_2 is a widely utilised energy source for microbial growth and survival. *The ISME Journal* **10:** 761-777.

Hamilton TL, Vogl K, Bryant DA, Boyd ES, Peters JW (2012). Environmental constraints defining the distribution, composition, and evolution of chlorophototrophs in thermal features of Yellowstone National Park. *Geobiology* **10:** 236-249.

Hedderich R (2004). Energy-converting [NiFe] hydrogenases from archaea and extremophiles: ancestors of complex I. *Journal of Bioenergetics and Biomembranes* **36:** 65-75.

Hoehler TM, Albert DB, Alperin MJ, Bebout BM, Martens CS, Des Marais DJ (2002). Comparative ecology of H_2 cycling in sedimentary and phototrophic ecosystems. *Antonie van Leeuwenhoek* **81:** 575-585.

Hoehler TM (2005). Biogeochemistry of dihydrogen. In: Sigel A, Sigel H, Sigel RKO (eds). *Biogeochemical Cycles of Elements*. Marcel Dekker: New York. pp 9-48.

Hunten DM (1973). The escape of light gases from planetary atmospheres. *Journal of the Atmospheric Sciences* **30:** 1481-1494.

Inskeep WP, Ackerman GG, Taylor WP, Kozubal M, Korf S, Macur RE (2005). On the energetics of chemolithotrophy in nonequilibrium systems: case studies of geothermal springs in Yellowstone National Park. *Geobiology* **3:** 297-317.

Inskeep WP, Jay ZJ, Tringe SG, Herrgard MJ, Rusch DB, YNP Metagenome Project Steering Committee and Working Group Members (2013). The YNP metagenome

project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem. *Frontiers in Microbiology* **4:** 1-15.

Kaster A, Moll J, Parey K, Thauer RK (2011). Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea. *Proceedings of the National Academy of Science of the United States of America* **108:** 2981-2986.

Kelley DS, Karson JA, Fruh-Green GL, Yoerger DR, Shank TM, Butterfield DA *et al* (2005). A sepentinite-hosted ecosystem: the lost city hydrothermal field. *Science* **307:** 1428-1434.

Kita I, Matsuo S, Wakita H (1982). H_2 generation by reaction between H_2O and crushed rock: an experimental study on H_2 degassing from the active fault zone. *Journal of Geophysical Research* **87:** 10789-10795.

Kozubal M, Macur RE, Korf S, Taylor WP, Ackerman GG, Nagy A *et al* (2008). Isolation and distribution of a novel iron-oxidizing Crenarchaeon from acidic geothermal springs in Yellowstone National Park. *Applied and Environmental Microbiology* **74:** 942-949.

Kozubal M, Macur RE, Jay ZJ, Beam JP, Malfatti SA, Tringe SG *et al* (2012). Microbial iron cycling in acidic geothermal springs of Yellowstone National Park: integrating molecular surveys, geochemical processes, and isolation of novel Fe-active microorganisms. *Frontiers in Microbiology* **3**.

Kral TA, Brink KM, Miller SL, McKay CP (1998). Hydrogen consumption by methanogens on the early Earth. *Origins of life and evolution of the biosphere* **28:** 311-319.

Lenz O, Friedrich B (1998). A novel multicomponent regulatory system mediates H_2 sensing in *Alcaligenes eutrophus*. *Proceedings of the National Academy of Science of the United States of America* **95:** 12474-12479.

Lie TJ, Costa KC, Lupa B, Korpole S, Whitman WB, Leigh JA (2012). Essential anaplerotic role for the energy-converting hydrogenase Eha in hydrogenotrophic methanogenesis. *Proceedings of the National Academy of Science of the United States of America* **109:** 15473-15478.

Lin LH, Hall JR, Lippmann-Pipke J, Ward JA, Lollar BS, DeFlaun M *et al* (2005). Radiolytic H_2 in continental crust: nuclear power for deep subsurface microbial communities. *Geochemistry, Geophysics, Geosystems* **6**.

Lineweaver CH, Schwartzman D (2004). Cosmic Thermobiology. In: Seckbach J (ed). *Origins. Cellular Origin, Life in Extreme Habitates and Astrobiology*. Springer: Dordrecht.

Lowenstern JB, Bergfeld D, Evans WC, Hunt AG (2015). Origins of geothermal gases at Yellowstone. *Journal of Volcanology and Geothermal Research* **302:** 87-101.

Lyons TW, Reinhard CT, Planavsky NJ (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **506:** 307-315.

Marrink SJ, Berendsen HJC (1996). Permeation process of small molecules across lipid membranes studied by molecular dynamic simulations. *Journal of Physical Chemistry* **100:** 16729-16738.

Martin W, Müller M (1998). The hydrogen hypothesis for the first eukaryote. *Nature* **392:** 37-41.

Martin WF (2012). Hydrogen, metals, bifurcating electrons, and proton gradients: The early evolution of biological energy conservation. *FEBS Letters* **586:** 485-493.

Masukawa H, Mochimaru M, Sakurai H (2002). Disruption of the uptake hydrogenase gene, but not of the bidirectional hydrogenase gene, leads to enhanced photobiological hydrogen production by the nitrogen-fixing cyanobacterium *Anabaena* sp. PCC 7120. *Applied Microbiology and Biotechnology* **58:** 618-624.

McCollom TM, Shock EL (1997). Geochemical constrains on chemolithoautotrophic metabolism by microorganisms in seafloor hydrothermal systems. *Geochimica et Cosmochimica Acta* **61:** 4375-4391.

Meyer J (2007). [FeFe] hydrogenases and their evolution: a genomic perspective. *Cellular and Molecular Life Sciences* **64:** 1063-1084.

Moore EK, Jelen BI, Giovannelli D, Raanan H, Falkowski PG (2017). Metal availability and the expanding network of microbial metabolisms in the Archaean eon. *Nature Geoscience* **10:** 629-636.

Mulder DW, Boyd ES, Sarma R, Lange RK, Endrizzi JA, Broderick JB *et al* (2010). Stepwise [FeFe]-hydrogenase H-cluster assembly revealed in the structure of HydA^DEFG. *Nature* **465:** 248-251.

Neal C, Stanger G (1983). Hydrogen generation from mantle source rocks in Oman. *Earth and Planetary Science Letters* **66:** 315-320.

Nicolet Y, Piras C, Legrand P, Hatchikian CE, Fontecilla-Camps JC (1999). *Desulfovibrio desulfuricans* iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure* **7:** 13-23.

Nordstrom DK, McCleskey RB, Ball JW (2009). Sulfur geochemistry of hydrothermal waters in Yellowstone National Park: IV Acid–sulfate waters. *Applied Geochemistry* **24:** 191-207.

Pace NR (1991). Origin of life - facing up to the physical setting. *Cell* **65:** 531-533.

Peters JW, Lanzilotta WN, Lemon BJ, Seefeldt LC (1998). X-ray crystal structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 angstrom resolution. *Science* **282:** 1853-1858.

Peters JW, Schut GJ, Boyd ES, Mulder DW, Shepard ES, Broderick JB *et al* (2015). [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. *Biochimica et Biophysica Acta - Molecular Cell Research* **1853:** 1350-1369.

Porat I, Kim E, Hendrickson EL, Xia Q, Zhang Y, Wang T *et al* (2006). Disruption of the operon encoding Ehb hydrogenase limits anabolic CO_2 assimilation in the archaeon *Methanococcus maripaludis*. *Journal of Bacteriology* **188:** 1373-1380.

Schut GJ, Boyd ES, Peters JW, Adams MWW (2013). The modular respiratory complexes involved in hydrogen and sulfur metabolism by heterotrophic hyperthermophilic archaea and their evolutionary implications. *FEMS Microbiology Reviews* **37:** 182-203.

Schut GJ, Zadvornyy O, Wu C-H, Peters JW, Boyd ES, Adams MWW (2016). The role of geochemistry and energetics in the evolution of modern respiratory complexes from a proton-reducing ancestor. *Biochimica et Biophysica Acta* **1857:** 958-970.

Shafaat HS, Rüdiger O, Ogata H, Lubitz W (2013). [NiFe] hydrogenases: A common active site for hydrogen metabolism under diverse conditions. *Biochimica et Biophysica Acta - Bioenergetics* **1827:** 986-1002.

Shock EL, Canovas P (2010). The potential for abiotic organic synthesis and biosynthesis at seafloor hydrothermal systems. *Geofluids* **10:** 161-192.

Shock EL, Holland M, Meyer-Dombard DA, Amend JP, Osburn GR, Fischer TP (2010). Quantifying inorganic sources of geochemical energy in hydrothermal ecosystems, Yellowstone National Park, USA. *Geochimica et Cosmochimica Acta* **74:** 4005-4043.

Spear JR, Walker JJ, McCollom TM, Pace NR (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 2555-2560.

Stevens TO, McKinley JP (2000). Abiotic controls on H_2 production from basalt - water reactions and implications for aquifer biogeochemistry. *Environmental Science and Technology* **34:** 826-831.

Telling J, Boyd ES, Bone N, Jones EL, Tranter M, MacFarlane JW *et al* (2015). Rock comminution as a source of hydrogen for subglacial ecosystems. *Nature Geoscience* **851-855**.

Thauer RK, Jungermann K, Decker K (1977). Energy conservation in chemotrophic anaerobic bacteria. *Bacteriological Reviews* **41:** 100-180.

Thauer RK (1998). Biochemistry of methanogenesis: a tribute to Marjory Stephenson. *Microbiology* **144:** 2377-2406.

Thauer RK, Kaster AK, Goenrich M, Schick M, Hiromoto T, Shima S (2010). Hydrogenases from methanogenic archaea, nickel, a novel cofactor, and H_2 storage. *Annual Review of Biochemistry* **79:** 507-536.

Therien JB, Artz JH, Poudel S, Hamilton TL, Liu Z, Noone SM *et al* (2017). The physiological functions and structural determinants of catalytic bias in the [FeFe]-hydrogenases CpI and CpII of *Clostridium pasteurianum* strain W5. *Frontiers in Microbiology* **8:** 1-11.

Tian F, Toon OB, Pavlov AA, De Sterck H (2005). A hydrogen-rich early Earth atmosphere. *Science* **308:** 1014-1017.

Trembley P, Lovley DR (2012). Role of the NiFe hydrogenase Hya in oxidative stress defense in *Geobacter sulfurreducens*. *Journal of Bacteriology* **194:** 2248-2253.

Vignais PM, Billoud B, Meyer J (2001). Classification and phylogeny of hydrogenases. *FEMS Microbiology Reviews* **25:** 455-501.

Vignais PM, Colbeau A (2004). Molecular Biology of Microbial Hydrogenases. *Current Issues in Molecular Biology* **6:** 159-188.

Vignais PM, Billoud B (2007). Occurence, classification, and biological function of hydrogenases: an overview. *Chemical Reviews* **107:** 4206-4272.

Wächtershäuser G (1990). Evolution of the first metabolic cycles. *Proceedings of the National Academy of Sciences of the United States of America* **87:** 200-204.

Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S *et al* (2016). The physiology and habitat of the last universal common ancestor. *Nature Microbiology* **1:** 1-8.

Windman T, Zolotova N, Schwander R, Shock EL (2007). Formate as an energy source for microbial metabolism in chemosynthetic zones of hydrothermal ecosystems. *Astrobiology* **7:** 873-890.

CHAPTER TWO

GEOLOGICAL SOURCE AND BIOLOGICAL FATE OF HYDROGEN IN YELLOWSTONE HOT SPRINGS

Contribution of Authors and Co-Authors

Manuscript in Chapter 2

Author: Melody R Lindsay

Contributions: Designed and conducted research, developed the model of hydrogen generation in YNP, analyzed samples, analyzed metagenomic data, conducted transcriptomic analyses, and contributed to the writing of the manuscript.

Co-Author: Daniel R Colman

Contributions: Analyzed metagenomic data from YNP springs, conducted analyses on [NiFe]-hydrogenases, contributed to the writing of the manuscript.

Co-Author: Maximiliano J Amenabar

Contributions: Field sampling and provided feedback on drafts of the manuscript.

Co-Author: Kirsten E Fristad

Contributions: Field sampling, analyzed gas samples, and contributed to the initial study design.

Co-Author: Kristopher M Fecteau

Contributions: Conducted geochemical analyses and provided feedback on drafts of the manuscript.

Co-Author: Randall V Debes II

Contributions: Conducted geochemical analyses and provided feedback on drafts of the manuscript.

Co-Author: John R Spear

Contributions: Analyzed SSU 16S rRNA sequences and provided feedback on drafts of the manuscript.

Co-Author: Everett L Shock

Contributions: Provided feedback on geochemical analyses, contributed to the hydrogen generation model, and provided feedback on drafts of the manuscript.

Co-Author: Tori M Hoehler

Contributions: Contributed to the study design and sampling locations, provided feedback on drafts of the manuscript.

Co-Author: Eric S Boyd

Contributions: Contributed to the study design and sampling locations, field sampling, contributed to the development of the hydrogen generation model, and contributed to the writing of the manuscript.

Manuscript Information

Melody R. Lindsay, Daniel R. Colman, Maximiliano J. Amenabar, Kirsten E. Fristad, Kristopher M. Fecteau, Randall V. Debes, John R. Spear, Everett L. Shock, Tori M. Hoehler, and Eric S. Boyd*

*Corresponding Author

Status of Manuscript:
____ Prepared for submission to a peer-reviewed journal
_X_ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Originally submitted on October 17, 2018; Resubmitted on February 20th, 2019

GEOLOGICAL SOURCE AND BIOLOGICAL FATE OF HYDROGEN IN YELLOWSTONE HOT SPRINGS

Abstract

Hydrogen (H_2) is enriched in hot springs (relative to atmospheric H_2) and can support microbial primary production. Using a series of geochemical proxies, a model to describe variable H_2 concentrations in Yellowstone National Park (YNP) hot springs is presented. Interaction between water and crustal iron minerals yields H_2 that partitions into the vapor phase during decompressional boiling of ascending hydrothermal fluids. Variable input of vapor leads to differences in H_2 concentration among springs. Analysis of 50 metagenomes from YNP springs reveals that genes encoding H_2 oxidation capability are enriched in communities inhabiting springs sourced with vapor-phase gas. Three springs in an area of YNP that are sourced with vapor-phase gas, termed Smokejumper (SJ), and with the highest [H_2] in YNP were examined to determine the fate of H_2. The highest concentration of H_2, the most biomass, and the highest abundance of culturable hydrogenotrophic cells were detected in SJ3 spring. Metagenomics and transcriptomics of SJ3 reveal a diverse sediment community comprised of abundant populations expressing genes involved in H_2 oxidation and carbon dioxide fixation. These observations suggest a link between the geologic processes that generate and source H_2 to hot spring environments and the distribution of organisms that utilize H_2 to support their energy metabolisms.

Introduction

Hydrogen (H_2) connects the geosphere and biosphere and has likely done so since early in the history of Earth (Boyd et al 2014, Hoehler 2005). Hot springs are often considered as analogs of early Earth environments (Djokic et al 2016) and can host waters with elevated H_2 concentrations (Bergfeld et al 2014, Lowenstern et al 2015, Spear et al 2005, Windman et al 2007). H_2 oxidation generally releases the most energy per mol of electrons transferred among inorganic compounds in hot spring ecosystems regardless of electron acceptor (i.e., oxidant) (Shock et al 2010). Therefore, variation in H_2 availability within hydrothermal hot spring systems could affect the distribution and activity of hydrogenotrophs, which are important primary producers in these environments (Lindsay et al 2018, Spear et al 2005).

A range of processes can generate H_2 including interaction between water and ferrous iron minerals (Kelley et al 2005, Stevens and McKinley 2000), the radiolysis of water (Lin et al 2005), reactions between mineral surface silicate radicals and water (Kita et al 1982, Telling et al 2015), or the activity of several biological enzymes (Boyd et al 2010, Peters et al 2015). The most likely sources of H_2 in high temperature Yellowstone National Park (YNP) hot springs are reactions between water and ferrous iron in rhyolitic or basaltic bedrock, although biological activity could contribute some H_2 to these springs (Boyd et al 2010). Additionally, during periods of extensive seismic activity that are commonplace in YNP (Christiansen 2001), silicate minerals can be sheared thereby exposing surface silica radicals or fresh unweathered ferrous iron-bearing minerals to water that could further contribute H_2 through the mechanisms described above.

H_2 is variable across YNP hot springs, ranging from 0 to ~7 mol % of total gases (Bergfeld et al 2014, Spear et al 2005, Windman et al 2007). This indicates spatial variability in the mechanisms that lead to H_2 generation in geographically distinct regions of YNP that may also correspond with various subsurface processes leading to spring formation. The current model for YNP hot spring formation begins with injection of magmatic gases [e.g., carbon dioxide (CO_2), sulfur dioxide (SO_2)] into a deeply seated hydrothermal aquifer(s) (Fournier et al 1976, Hurwitz and Lowenstern 2014, Rye and Truesdell 1993, Truesdell and Fournier 1976, Truesdell et al 1977, White et al 1971). When SO_2 condenses with hydrothermal water (at temperatures between ~200-400˚C), it is thought to disproportionate to yield sulfate (SO_4^{2-}) and sulfide (H_2S) (Nordstrom et al 2009). As the hydrothermal water ascends and infiltrates the crust, it can be further altered by processes such as water-rock interaction, decompressional boiling, and mixing with near-surface groundwater (Lowenstern et al 2012, Nordstrom et al 2009). Of particular importance is decompressional boiling that is thought to separate subsurface hydrothermal water into a liquid phase enriched in non-volatile constituents [e.g., chloride (Cl^-)] and a vapor phase enriched in volatile constituents [e.g., H_2S] (Fournier 1989, Nordstrom et al 2009). The vapor is then able to migrate toward the surface where it can condense and interact with oxygen (O_2)-rich meteoric fluids. Near-surface oxidation of H_2S with O_2 to SO_4^{2-} and protons can result in the acidification of waters (Fournier 1989, Nordstrom et al 2004, Nordstrom et al 2005), a process suggested to be mediated primarily by microbial activity (Colman et al 2018) as rates of biotic H_2S oxidation are estimated to be three or more orders of magnitude faster than rates of

abiotic H_2S oxidation (Luther et al 2011). Thus, in this model of hot spring formation, springs with vapor-phase-influenced source waters tend to be enriched in SO_4^{2-} and are often acidic to moderately acidic pH, while hot springs with liquid-phase-influenced source waters tend to have a lower concentration of SO_4^{2-}, are enriched in Cl^-, and are circumneutral to alkaline in pH.

In this study, we aimed to define the subsurface processes that lead to variable H_2 concentrations in YNP hot springs. Further, we sought to determine how these processes affect the distribution and abundance of hydrogenotrophic, chemoautotrophic populations in hot springs. The results are integrated to define links between processes that source and enrich springs in H_2 and the distribution and abundance of microorganisms that can metabolize H_2.

Materials and Methods

Determination of hydrogenase protein encoding gene homolog abundance in chemosynthetic community metagenomes.

Metagenomes deposited within the Integrated Microbial Genomes (IMG) database (Markowitz et al 2012) that were ascribed to chemosynthetic communities in YNP (as of 09/17/2018) were compiled. Geochemical data (e.g., pH and temperature) reported in the IMG database or otherwise in published reports associated with the metagenomes were also compiled. A total of 50 metagenomes [including Smokejumper 3 (SJ3), as described below; Dataset S1] were recovered.

The Kyoto Encyclopedia of Genes and Genomes (KEGG) KEGG Orthology (KO) annotations for each assembled metagenome were compiled from IMG source data files. From the annotations, total hydrogenase KO abundances were determined for large subunit (catalytic subunit) homologs of [NiFe]-hydrogenase isoforms inferred to be physiologically biased towards H_2 oxidation (K06281, K05922, K00437, and K14068 for group 1 homologs; K00440, K17993, K00436, and K14126 for group 3a and 3c homologs) (Peters et al 2015). KOs for group 2 [NiFe]-hydrogenase isoforms are not available in KEGG and thus were not included in this analysis. KO abundances were also compiled for catalytic subunits of [NiFe] isoforms inferred to exhibit minimal directional bias (bidirectional) (K17993 for group 3b; K00436 for group 3d) (Peters et al 2015). While many group 4 [NiFe]-hydrogenases exhibit minimal bias or are biased toward H_2 oxidation (Peters et al 2015), others are biased towards H_2 production. Given the difficulty of predicting directionality of these homologs via sequence motifs or homology, we did not further classify group 4 [NiFe]-hydrogenase homologs with respect to directionality. Similarly, we also did not further classify [FeFe]-hydrogenase homologs with respect to directionality in this calculation. KO abundances for all hydrogenase homologs were also compiled, including group 4 [NiFe]-hydrogenase and [FeFe]-hydrogenases (K12142, K14090, K14106, K14123, K15830, and K18016 for group 4 [NiFe]-hydrogenase homologs; K18332, K00533, and K17997 for [FeFe]-hydrogenase homologs). Homologs corresponding to [NiFe]-hydrogenase groups were further classified as previously described (Greening et al 2015, Søndergaard et al 2016). The summed totals of the KO annotations were then normalized by the total number of KO annotations for each metagenome.

Physical and Chemical Measurements.

Three springs within the Smokejumper hydrothermal area, an area previously shown to be enriched in H_2 (Bergfeld et al 2014) (Fig. 2.1), were selected for detailed chemical and biological characterization. These are referred to as SJ1 (N44°24'57.60"; W-110°57'20.46"), SJ2 (N44°24'58.68"; W-110°57'19.50"), and SJ3 (N44°24'57.42"; W-110°57'20.76") (Fig. 2.1B). Samples were collected on July 22, 2014. Temperature, pH, and conductivity of spring waters were measured on-site as described previously (Lindsay et al 2018). Ferrous iron (Fe^{2+}), total sulfide (S^{2-}), and dissolved silica (SiO_2) were quantified with Hach reagents and a field spectrophotometer. Samples for other cations and anions were filtered through a 0.22 μM membrane filter, acidified in Nalgene bottles, and stored at -20°C. Anion concentrations were determined on a Dionex DX-600 ion chromatography system according to previously published protocols (Lindsay et al 2018), except that columns were equilibrated for 10 min. prior to each injection. Samples for determining stable isotope ratios of water (δ^2H and $\delta^{18}O$) were collected in 30 mL Qorpak glass bottles and were subjected to cavity ring-down spectroscopy with a Los Gatos Research DLT-100 isotope analyzer. Data processing was completed as previously described (van Geldern and Barth 2012). Isotope ratios are reported relative to Vienna Standard Mean Ocean Water.

Triplicate samples for determining concentrations of dissolved gases [H_2, CH_4, and carbon monoxide (CO)] were collected using a modified bubble-strip method, as described previously (Chapelle et al 1997, Spear et al 2005). Briefly, a syringe was filled

with 55 mL of spring water. Five mL of ultra-high purity N_2 gas was then introduced into the syringe followed by shaking for 1 min. allowing for gas equilibration. The gas bubbles were then transferred to glass vials containing a saturated sodium chloride solution for storage until analysis by gas chromatography, as previously described (Lindsay et al 2018).

DNA Extraction.

Triplicate sediment subsamples were collected from the source of each spring (SJ1, SJ2, and SJ3) using a sterilized spatula. These sediments were collected from the sediment-spring water interface and were placed in 2 mL cryotubes containing 500 µL of RNALater. Tubes and their contents were immediately frozen on dry ice for transport to the lab where they were then stored at -80°C until used for molecular analyses. DNA was extracted and quantified according to our previously published protocols, using the FastDNA SPIN Kit for Soil (MP Biomedicals, Irvine, CA) with bead-beating (Hamilton et al 2013).

PCR, Quantitative PCR, and Sequencing.

DNA was subjected to amplification of small subunit rRNA genes using the PCR primers 515F-806R. Thirty-five cycles of PCR were conducted as described previously (Colman et al 2016). Quantitative PCR (qPCR) of 16S rRNA genes was performed on DNA extracts using the 515F-806R primers and the SsoAdvanced Universal SYBR Green Supermix in a reaction volume of 20 µL, using our described methods (Lindsay et al 2018). Plasmid standards for use in relating template copy number to threshold

amplification signals were prepared as described previously (Boyd et al 2011). Negative control reactions were performed in the absence of added DNA.

PCR amplicons were sequenced via multiplexed, paired-end Illumina MiSeq tag sequencing according to previously published protocols (Bradley et al 2017). Sequence was processed with Mothur (v.1.36.1) following previously described protocols (Lindsay et al 2018, Schloss et al 2009). Sequences were trimmed to a length of 225 bases and operational taxonomic units (OTUs) were assigned at a sequence identity of >97% using the nearest-neighbor method. Collectively, all steps resulted in a normalized size of 1,692 16S rRNA gene sequences for each library. Sequences specific for each OTU were classified using the Bayesian classifier and the Ribosomal Database Project (Wang et al 2007), with manual verification using the nucleotide basic local alignment search tool (BLASTn). Raw reads, quality scores, and mapping files for the 16S rRNA gene libraries have been deposited in the National Center for Biotechnology Information (NCBI) sequence reads archive under BioProject number PRJNA48349.

Dilution-to-Extinction Cultivation Assays.

Single tube dilution-to-extinction cultivation assays were used to estimate the number of viable autotrophic cells associated with hot spring sediments capable of coupling H_2 oxidation to the reduction of one of several electron acceptors. Base salts medium was prepared as described previously (Boyd et al 2007) but modified to exclude peptone and elemental sulfur (S°). For each experiment, the pH of the cultivation medium was adjusted to that measured at each of the springs. Tubes containing base salts medium,

Wolfes vitamins (1 mL L^{-1}), and SL10 trace elements (1 mL L^{-1}) were amended with electron acceptors ($MgSO_4 \bullet 7H_2O$, $NaNO_3$, and $FeCl_3 \bullet 6H_2O$) to final concentrations of 10 mM; S° was added to a final concentration of 5 g L^{-1}. Tubes and contents were purged with sterile N_2 and the headspace of all vials was then replaced with a 20%/80% CO_2/H_2 sterile gas mixture. In the case of aerobic assays, air (and therefore O_2) was added to the headspace to achieve final concentrations of 4% v/v O_2 with a balance of 12% N_2 v/v, 64% H_2 v/v, and 16% CO_2 v/v. A slight overpressure (5 p.s.i.) was introduced to each vial using 100% H_2.

Dilution-to-extinction assays were inoculated with 1 mL of a slurry of sediment and spring water collected from the source of each spring, followed by 10-fold serial dilution series that spanned a range from 10^{-1} to 10^{-8}. Following 8 weeks of incubation, the production of total S^{2-} or Fe^{2+} and/or cells was determined for each condition for each spring sample according to previously published protocols (Boyd et al 2007, Lindsay et al 2018). Dilutions that contained more than $1x10^5$ cells mL^{-1} after incubation (three orders of magnitude higher than background) were considered to have been inoculated with at least one viable H_2 oxidizing, autotrophic cell. Estimates of H_2 oxidizing autotrophic cells in the initial sediment and water slurries were calculated based on this assumption and were normalized to grams dry mass (gdm) in the original inoculum.

Metagenomic Sequencing and Analysis.

Total genomic DNA (~5 ng) from SJ3 was subjected to shotgun metagenomic sequencing using the paired-end [2x250 base pairs (bp)] Illumina HiSeq 2500 Rapid

Platform with Illumina Nextera DNA library preparation. Reads were quality trimmed and Illumina adapters removed using Trimmomatic v.0.35 (Bolger et al 2014). MEGAHIT v.1.1.1 was used to assemble contigs and remove short and low-depth contigs <75 bp long (Li et al 2015). Paired-end reads were aligned to contigs and locally assembled to generate missing *k-mers* in low depth regions using Bowtie2 (Langmead and Salzberg 2012). MetaQUAST v3.2 (Mikheenko et al 2015) was used to assess the assemblies. Contigs (>2.5 kbp) from the highest quality assembly (*k-mer* size=141) were binned by tetranucleotide word frequency distribution patterns and a window size of 4 kbp with MetaBAT v.0.26.3 using the "very sensitive parameters" defined by (Kang et al 2015). Binned contigs were then assessed for quality, contamination, and completeness using CheckM v1.0.5 (Parks et al 2015). Outlier contigs, defined as being outside the distribution of 95% of the contig tetranucleotide distance of each bin, were removed. Assembled contigs have been uploaded to the IMG database under the genome ID 3300029625.

Curated bins were then reassessed using the same CheckM quality control methods outlined above. Bins not meeting quality metrics were subjected to re-binning with even higher stringency until contamination in each bin was <5%. In some cases, genome bins that were estimated to be abundant (>1.0% relative abundance) exhibited low completeness. Improvement in bin quality was conducted by recruiting contigs to publicly available genomes, as described in (Colman et al. 2018) and extracted reads were then reassembled using the Spades v.3.10.0 assembler (Nurk et al 2013). All bins were confirmed by visualization with emergent self-organizing map binning (ESOM),

using methods to bin contigs with *k*-mer frequencies as previously described (Dick et al 2009). Gene predictions and annotations were performed using Prokka v1.11 (Seemann 2014). Annotations of genes encoding proteins that were of specific interest to this study (defined below) were further scrutinized using homology searches against the non-redundant database at the NCBI.

The abundance of genomic bins was determined by mapping individual reads to the assembled contigs, determining coverage, and normalizing by genome size, as described previously (Schut et al 2016). Paired reads were mapped to assembled contigs using Bowtie2 allowing for 1 mismatched bp (Langmead and Salzberg 2012). The total number of mapped reads was divided by the total contig length for each bin to assess coverage, and the coverage was multiplied by the estimated genome size to calculate normalized number of reads. Genome size was estimated based on genome completeness and total length of the binned genomes. The normalized number of reads was divided by the total number of reads for all populations to estimate relative abundances.

Identification and classification of [NiFe]- and [FeFe]-hydrogenase homologs in the SJ3 metagenome.

Hydrogenase protein homologs were extracted from the SJ3 metagenome using protein BLAST (BLASTp) searches for group 1, group 2, group 3, and group 4 [NiFe]-hydrogenases using the large subunit protein from *Aquifex aeolicus* (WP_010880393) for group 1, *Sulfurihydrogenibium azorense* (WP_01273603) for group 2, *Methanothermobacter marburgensis* (WP_013296467) for group 3, and *Pyrococcus abyssi* (WP_010867842) for group 4 as queries. E-value cutoffs for defining homologs

were empirically defined as $E < 7e^{-4}$ for group 1 [NiFe]-hydrogenase, $E < 7e^{-4}$ for group 2 [NiFe]-hydrogenase, $E < 5e^{-4}$ for group 3 [NiFe]-hydrogenases, and $E < 9e^{-52}$ for group 4 [NiFe]-hydrogenases. A BLASTp search for the large subunit of [FeFe]-hydrogenase (HydA) was conducted using a homolog from *Chlamydomonas reinhardtii* (XP_001693376) as the query specifying an E-value cutoff of $E < 1e^{-4}$. Reference homologs from each [NiFe]-hydrogenase homolog subgroup (Greening et al 2015) were also compiled. Screened [NiFe]-hydrogenase homologs were aligned using Clustal Omega (Sievers et al 2011), and then subjected to a phylogenetic reconstruction using RAxML v.8.2.9 and the PROTGAMMALG substitution model (Stamatakis 2014). HydA homologs identified in the SJ3 metagenome were aligned with a CpI-like HydA (i.e. biased towards the production of H_2) and a CpII-like HydA (i.e. biased towards the oxidation of H_2) as identified previously (Adams et al 1989), using Clustal Omega (Sievers et al 2011). HydA sequence alignments were visualized with SeaView (version 4.7) (Gouy et al 2010) and analyzed for conservation in key residues in the H-cluster (active site) and accessory iron-sulfur (Fe-S) cluster secondary coordination spheres that have previously been suggested to confer catalytic bias (Therien et al 2017).

RNA Extraction, cDNA Synthesis, and Quantification of Gene Expression.

The FastRNA Pro Soil-Direct kit, modified as previously described (Lindsay et al 2018), was used to extract RNA from SJ3 sediments. RNA extracts were treated with DNAse, checked for residual DNA, and cDNA was synthesized from 1 ng of RNA using the iScript cDNA Synthesis Kit as described previously (Lindsay et al 2018).

Genome bins (Dataset S2) were screened for homologs of genes encoding the large subunits of the two primary enzymes capable of activating H_2: [NiFe]- and [FeFe]-hydrogenase, as defined above (Peters et al 2015). Individual primer sets were designed to amplify fragments of homologs of the large subunit of hydrogenase genes found in the 10 most abundant genome bins using Primer-BLAST (Ye et al 2012), and these are reported in Table 2.1/S1 (Supplemental Table 1). Gene-specific PCRs were performed as described above at empirically optimized annealing temperatures as determined using gradient PCRs (from 48 to 65°C) (Table 2.1/S1). PCR amplicons were purified, ligated into pGEM plasmids, and cloned as described previously (Boyd et al 2011). Primer specificity was verified with PCRs using plasmids containing inserts corresponding to closely related gene homologs as negative controls.

Translated proteins from the 10 most abundant population-level genome bins in SJ3 were also subjected to a GHOSTX homology search against the KEGG database (Suzuki et al 2014). Bins were screened for genes encoding key proteins that demarcate known CO_2-fixation pathways (Berg et al 2010). These include genes encoding the β subunit of ATP-citrate lyase (*aclB*; reductive citric acid cycle), formate-tetrahydrofolate ligase (*fhs*; reductive acetyl-CoA pathway), the δ subunit of 4-hydroxybutyryl-CoA dehydratase (*abfD*; 3-hydroxypropionate-4-hydroxybutyrate cycle and the dicarboxylate-4-hydroxybutyrate cycle), the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (*cbbL*; reductive pentose phosphate cycle) and malonyl-CoA reductase (*mcr*; 3-hydroxypropionate bicycle). Individual primer sets were designed to amplify fragments of *aclB*, *fhs*, *abfD*, *cbbL*, and *mcr* genes from each genome bin using

Primer-BLAST (Ye et al 2012), and these are reported in Table 2.1/S1. Gene specific PCRs were performed at the optimized PCR annealing temperatures as determined experimentally with gradient PCRs (from 48 to 65°C). Amplicons were purified, ligated into pGEM plasmids, and cloned as described previously (Boyd et al 2011). Quantification with qPCR of gene transcript abundances of proteins involved in H_2 metabolism and CO_2-fixation was conducted as described above, using cDNA and the primers described in Table 2.1/S1, and plasmid standards.

RESULTS AND DISCUSSION

Geological source of H_2 in Yellowstone hot springs.

Previously compiled gaseous and aqueous chemical and isotopic data, together with geospatial data, from 128 geothermal features spanning YNP (Bergfeld et al 2014) were used to assess potential geological sources of H_2. H_2 concentrations ranged from <0.00003 to 6.969 mol% of total gases in geothermal features sampled from across YNP (Fig. 2.1). Spatial mapping of H_2 concentrations reveal distinct areas with geothermal features that have elevated H_2. These include the Smokejumper (SJ), Washburn Hot Springs (Washburn), and Hot Spring Basin (HSB) areas (Fig. 2.1), each of which are located within 10 km of the caldera rim formed from the most recent major volcanic eruption. These geothermal areas are also located near the edge of the ring-fracture zone (Fig. 2.1A), where bedrock is extensively fractured and fissured due to increased seismic activity in this area (Christiansen 2001). These features allow for gases to rise to the surface more easily (Lowenstern et al 2015). The high elevation of these spring areas

(Fig. 2.7/S1) likely enhances separation of vapor and liquid following decompressional boiling, further allowing gases to concentrate in springs in these areas.

Geological processes that may be involved in production of H_2 in YNP hot springs were investigated using a series of proxies for deep versus shallow fluid-rock interactions. CH_4 is commonly used as a proxy for crustal input of abiogenic gases (Bergfeld et al 2014, Giggenbach and Poreda 1993, Lowenstern et al 2015). In contrast, ^3He is sourced from the mantle while ^4He is produced in the crust through radioactive decay (Bergfeld et al 2014, Giggenbach and Poreda 1993, Lowenstern et al 2015). Thus, the ratio of ^3He to ^4He [given as the ratio (R) of ^3He to ^4He to the present atmospheric ratio (R_a)] is used as a proxy to evaluate input of mantle-derived abiogenic gases. YNP gases with significant mantle input exhibit R/R_a ratios above 5 (Lowenstern et al 2014) (Fig. 2.8A/S2A). A plot of CH_4 mol% as a function of R/R_a reveals increased CH_4 when R/R_a values decrease below 5 (Fig. 2.8B/S2B), consistent with increasing crustal inputs of both CH_4 and ^4He (Giggenbach et al 1993, Lowenstern and Janik 2003). As such, the fraction of total gas as CH_4 is positively correlated with the fraction of total gas as He, and both CH_4 and He increase in fractional abundance in gases that also have low R/R_a values (<5) (Fig. 2.8A and B/Fig. S2A and B).

Hot spring waters with elevated concentrations of total He and CH_4 also have elevated concentrations of H_2 (Fig. 2.8C/S2C) and the highest H_2 is observed in springs with R/R_a values between 5 and 9 (Fig. 2.8D/S2D) Together, these observations suggest that springs with elevated H_2 are sourced by mantle-influenced fluids (indicated by R/R_a

values) that have undergone extended reactions with crustal minerals (as indicated by high total He and CH_4). This includes minerals that contain ferrous iron, the putative source of reductant for reducing water to H_2.

To further investigate geological processes that may lead to H_2 enrichment in YNP thermal features, H_2 mol% was compared to the Cl^-/SO_4^{2-} ratio in corresponding waters. Cl^-/SO_4^{2-} ratios are used to infer the source of fluids to hot springs (Nordstrom et al 2009). Springs sourced with vapor phase gases are enriched with H_2S that, when oxidized, contributes SO_4^{2-} and acidity (Nordstrom et al 2009). Springs that are sourced with meteoric water infiltrated with vapor phase gases therefore have low Cl^-/SO_4^{2-} ratios and moderately acidic pH (~4 to 6.5). This spring type is also often enriched in H_2, He, and CH_4 (Fig. 2.8E and F/Fig. S2E and F). Based on these observations, it is proposed that H_2 (along with He, CH_4) produced by interaction of hydrothermal fluids and crustal minerals are concentrated in springs with moderately acidic pH that are sourced by meteoric water infiltrated with vapor phase gas (Fig. 2.2). While biological production of H_2 could also contribute to variation in H_2 concentrations in hot springs (Peters et al 2015), the broad pattern of H_2 enrichment in springs that have a combination of low Cl^-/SO_4^{2-}, high He, and high CH_4 concentrations suggest that variable input of vapor phase gas is the primary control on H_2 in this spring type (Fig. 2.2).

Biological fate of H_2 in Yellowstone hot springs.

To begin to investigate the biological fate of H_2 in hot springs across YNP, the abundance of hydrogenase homologs in metagenomes from 50 chemosynthetic hot spring

communities was determined. Specifically, the relative abundance of homologs of hydrogenase catalytic subunits predicted to be biased towards the oxidation of H_2 or that exhibit minimal bias (i.e., groups 1 and 3 of [NiFe]-hydrogenases) was determined. In addition, the relative abundance of group 4 [NiFe]-hydrogenases and [FeFe]-hydrogenases were determined. Results indicate that hydrogenase protein homologs predicted to be biased towards H_2 oxidation or that are bidirectional are enriched in metagenomes from springs that have moderately acidic pH and that are likely sourced with meteoric water infused with vapor phase gases (Fig. 2.3). This includes SJ3 and springs in Washburn and HSB, where the highest concentrations of H_2 in YNP have been reported. Importantly, although some proteins that are phylogenetically classified as group 4 have been shown to exhibit minimal catalytic bias (bidirectional), most exhibit bias towards H_2 production (Peters et al 2015). It is not possible to confidently predict the directionality of these enzymes based on homology alone. The same is true for [FeFe]-hydrogenases, and only in rare cases can enzyme directionality be inferred based on homology at the primary sequence level [described below, (Therien et al 2017)]. Nonetheless, inclusion of group 4 [NiFe]-hydrogenase and [FeFe]-hydrogenase homologs in this calculation does not change the overall pattern of enrichment of hydrogenase protein homologs in this spring type (Fig. 2.3). Collectively, these observations suggest a link between the geological processes that generate H_2 and lead to its enrichment in certain springs and the distribution of microorganisms that are adapted to take advantage of H_2 as an electron donor to fuel metabolism in these springs.

To further investigate the link between H_2 and microbial function, three chemosynthetic hot springs in the SJ hydrothermal area (i.e., SJ1, SJ2, and SJ3; Fig. 2.1B) were sampled since springs in this area have the highest average mol% H_2 abundances measured in YNP (Bergfeld et al 2014). Waters from SJ1, SJ2, and SJ3 spanned a range of temperatures (81.7°C, 81.5°C, and 61.9°C, respectively) and pH (6.5, 7.6, and 5.4, respectively). Ratios of Cl^- to SO_4^{2-} were low in the three SJ hot springs (Fig. 2.9A/S3A), suggesting that they comprise meteoric water that has been infused with vapor phase gas (Nordstrom et al 2009) (Table 2.2/S2). Consistent with this interpretation, the δD (δ^2H-H_2O) and $\delta^{18}O$ values for the three SJ hot spring waters (Fig. 2.9B/S3B) indicate that they are more negative than most YNP springs (Nordstrom et al 2009) and plot closer to hydrothermal vapor condensates (Bergfeld et al 2014). The concentration of dissolved H_2 in waters from all three SJ springs are among the highest reported for YNP, ranging from 0.47 to 1.87 μM (Fig. 2.4A). For perspective, cultivars of several organisms that metabolize H_2 as an electron donor tend to require a minimum of 5-10 nM H_2 (Lovley 1985). The highest concentration of H_2 (1.87 μM) was detected in SJ3, which also had the lowest concentration of Cl^- and the highest concentration of SO_4^{2-} among the three SJ springs sampled (Fig. 2.9A/S3A). Thus, H_2 in SJ springs is interpreted to reflect an extreme example of near surface mixing of meteoric water with abundant vapor phase gas, the latter of which is enriched in H_2.

The total abundance of archaeal and bacterial 16S rRNA gene templates associated with sediments from SJ hot springs was determined as a proxy for microbial biomass in each spring. Gene template abundance was highest in sediments from SJ3

($1.32x10^9$ copies gdm^{-1}) when compared to sediments from SJ1 and SJ2 ($1.28x10^7$ and $4.57x10^6$ copies gdm^{-1}, respectively) (Fig. 2.4B, Table 2.2/S2). The abundance of 16S rRNA gene templates in SJ3 is an order of magnitude greater than the median abundance of $1.08x10^8$ gene copies gdm^{-1} in 16 chemosynthetic springs in YNP measured previously (Colman et al 2016), suggesting this to be a productive hot spring community. The concentration of H_2, a low-potential reductant capable of fueling microbial productivity (Boyd et al 2014, Greening et al 2015), in SJ3 was 2- and 4-fold higher than SJ1 and SJ2, respectively. Moreover, the SJ springs exhibit dissolved H_2 concentrations that are 1.5 to 5.8-fold greater than the 12 YNP hot springs measured previously where H_2 was initially hypothesized to have a key role in supporting chemosynthetic hot spring productivity (Spear et al 2005).

To assess the role of H_2 in supporting autotrophic metabolism in SJ springs, the approximate minimum number of viable autotrophic hydrogenotrophic cells capable of reducing a variety of supplied oxidants common in hot spring environments (O_2, Fe^{3+}, NO_3^-, $S_2O_3^{2-}$, SO_4^{2-}, S°) (Shock et al 2010) was determined using serial dilution-to-extinction cultivation assays. Importantly, these assays are likely to underestimate the number of hydrogenotrophs in a given sample, since the specified growth conditions may not support all H_2 oxidizing autotrophic cell types. The abundance of hydrogenotrophic autotrophic cells, as indicated by growth in the most dilute assay supplied with any of the specified oxidants, was nearly two orders of magnitude higher at SJ3 ($1.0x10^6$ cells $gdws^{-1}$) than SJ1 and SJ2 ($3.9x10^4$ and $4.3x10^4$ cells $gdws^{-1}$, respectively) (Fig. 2.4C, Table 2.3/S3). The use of oxidants that supported H_2 dependent autotrophic growth in the SJ

communities also varied between communities. Populations comprising the SJ3 community were capable of coupling H_2 oxidation to reduction of all oxidants tested whereas the communities at SJ1 and SJ2 could couple H_2 oxidation to only a subset of the provided oxidants. Comparison to a spring (Roadside West (RSW), Nymph Lake area, YNP) with a similar pH and temperature (pH 7.0, 68°C) but with a lower concentration of H_2 (0.175 μM) revealed a slightly lower abundance of culturable hydrogenotrophic and autotrophic cells (4.6×10^5 $gdws^{-1}$) (Lindsay et al 2018). However, unlike the SJ3 populations, those in RSW could only couple H_2 oxidation with a narrow spectrum of oxidants. Thus, within the limitations of the dilution to extinction cultivation-based approach, these data suggest that of the three SJ springs, SJ3 hosts the most diverse and abundant assemblage of autotrophic hydrogenotrophic cells. Moreover, these data suggest that the diversity of hydrogenotrophs in SJ3 is likely greater than in springs in YNP with lower H_2 concentrations.

To further characterize the SJ3 community, community genomic DNA from sediments from SJ3 was subjected to metagenomic sequencing. Overall, the composition and percent abundances of genome bins identified in the SJ3 metagenome (Fig. 2.5A) were like those of 16S rRNA gene OTUs (Fig. 2.10/S4) indicating that community diversity measured at the 16S rRNA gene level translates to whole genome level community diversity for SJ3. A total of ~61 Gbp of high-quality assembled data comprising over 290 million reads were generated (Table 2.4/S4). Binning of assembled contigs based on tetranucleotide word frequencies and coverage profiles resulted in 109 draft genome bins (confirmed by ESOM binning visualization, Fig. 2.11/S5), with 82 of

those bins estimated to be >50% complete (Dataset S2), indicating that a significant portion of genes are estimated to be present in these bins.

The 82 genome bins were probed for homologs of genes encoding the large subunits of the two primary enzymes capable of activating H_2: [NiFe]- and [FeFe]-hydrogenase (Peters et al 2015). Homologs of genes coding for one or both enzymes were identified in 59 of the 82 (72%) of genome bins. Considering that only ~26% of all archaeal and bacterial organisms with available genome sequences in public databases code for homologs of either [NiFe]- or [FeFe]-hydrogenase (Peters et al 2015), these data indicate that populations inhabiting SJ3 are adapted to metabolize H_2.

The hydrogenases identified in the SJ3 bins were further classified based on homology to biochemically characterized hydrogenases and using previous classification schemes (Greening et al 2015, Poudel et al 2016, Therien et al 2017, Vignais et al 2001). The large subunits of the [NiFe]-hydrogenase were designated as belonging to subgroups of group 1, 2, 3, or 4 [NiFe]-hydrogenases (Greening et al 2015), based on phylogenetic relationships (Fig. 2.5B). [NiFe]-hydrogenase homologs affiliated with groups 1, 2, and 3 are physiologically biased toward H_2 oxidation or are bidirectional (Greening et al 2015, Vignais et al 2001). For [FeFe]-hydrogenases, catalytic bias was predicted based on conservation of residues in the H-cluster domain of the catalytic subunit, HydA, with homologs from *C. pasteurianum* which have previously been shown to be catalytically biased towards H_2 oxidation (CpII) or production (CpI) (Adams and Mortenson 1984, Therien et al 2017). The primary ligands for the H cluster and accessory iron (Fe)- sulfur

(S) cluster clusters in HydA are conserved (Meyer 2007), indicating that these are not responsible for catalytic bias. However, key residues in the secondary coordination sphere of the H cluster and accessory Fe-S clusters of HydA can be variable and have been suggested to contribute to catalytic bias (Therien et al 2017). For example, 13 of the 14 HydA in SJ3 bins (all except from Bin 98) exhibit substitutions in key secondary coordination sphere residues, including position 163 (Fe sub-cluster of the H cluster), 268 (proximal 4Fe-4S cluster), and 357/505 (4Fe-4S-subcluster of H cluster), that have been suggested to catalytically bias CpII towards H_2 oxidation (Fig. 2.12/S6).

Based on predictions of the catalytic bias for [NiFe]- and [FeFe]-hydrogenase homologs, 53 of the 59 (90%) genome bins encode at least one isoform predicted to be involved in H_2 oxidation (Dataset S2, Fig. 2.5A). This suggests that the SJ3 community is adapted to use H_2 as an electron donor to support metabolism. Furthermore, the [NiFe]-hydrogenase homologs in the SJ3 metagenome (including both binned and unbinned homologs) span the known diversity of these enzymes and are predicted to couple metabolically with a wide variety of oxidants (Fig. 2.5B), consistent with the presence of a taxonomically and functionally diverse hydrogenotroph community in this H_2 rich spring (Fig. 2.3). This conclusion is supported by results from a companion paper that revealed the SJ3 community to be hyperdiverse and host representatives of >50% of the higher order archaeal and bacterial lineages (Colman et al 2019).

The ten most abundant genome bins, accounting for 62% of the total reads, had relative abundances of 2.0% of total read abundances or greater. Sequence analysis of the

most abundant bin revealed close affiliation [RNA polymerase subunit β (RpoB) is 96% identical] to *Caldisericum exile*, an anaerobic, non-fermentative, chemoheterotrophic S^0 reducer isolated from a hot spring in Japan (Mori et al 2009). This bin represented 15.4% of total reads (Fig. 2.5A) and was estimated to be 80% complete (Dataset S2). The next most abundant bin was nearly as abundant (15.1% of total reads, 93% complete) and was mostly closely related (RpoB 97% identical) to the S^0 and H_2S oxidizing aerobic autotroph *Sulfurihydrogenibium yellowstonensis* isolated from Calcite Springs in YNP (Nakagawa et al 2005). This bin encoded for homologs of *sox*B and *sqr* genes (locus tags Ga0311297_10021208 and Ga0211297_10197871, respectively), indicating that this organism is likely supported by oxidation of reduced sulfur compounds, including H_2S. The presence of a dominant SJ3 population with this physiological activity is consistent with infiltration of meteoric water with vapor phase gas that is enriched in H_2S. The most abundant archaeal bin and seventh most abundant bin overall (genome is 98% complete) represented 2.7% of the total reads and was distantly related to the anaerobic genus *Aciduliprofundum* (RpoB is 64% identical) within the Deep Sea Hydrothermal Vent Euryarchaeota 2 (DHVE2) group, which are common in deep sea hydrothermal vents (Reysenbach et al 2006) (Fig. 2.5A).

To further establish the role of H_2 in supporting the SJ3 community, the 10 most abundant bins were screened for transcripts of the catalytic subunits of encoded hydrogenase homologs. Among the 10 most abundant bins, three did not code for hydrogenase homologs and these were most closely affiliated with *Sulfurihydrogenibium yellowstonense*, *Geobacter* sp. GB1 and *Thermus aquaticus* (Fig. 2.6A). Homologs in the

other seven bins represent a mix of hydrogenases with putative roles in H_2 oxidation or production (Fig. 2.5A & 2.6A, Dataset S2). The expression of hydrogenases in these 7 genome bins was examined using qPCR with cDNA as template and primers specific to each hydrogenase catalytic subunit (Table 2.1/S1). Hydrogenase transcripts affiliated with 6 out of the 7 most abundant bins were detected. Transcripts of [NiFe]-hydrogenases predicted to be involved in H_2 oxidation and that were affiliated with candidate division Acetothermia, a DHVE2 archaeon, an unclassified crenarchaeote, an unclassified bacterium (termed GXS-1, likely of the Spirochaete phylum), and an Elusimicrobia bacterium were detected. Transcripts of a second [NiFe]-hydrogenase in an unclassified crenarchaeote bin, predicted to exhibit bias towards the production of H_2, were also detected. Finally, transcripts of a [FeFe]-hydrogenase (Hyd) that is predicted to be involved in H_2 oxidation, based on the conservation of key residues thought to confer bias towards H_2 oxidation in CpII (Fig. 2.12/S6), were detected and these were affiliated with the most abundant bin, *Caldisericum exile* (Fig. 2.6A). If this Hyd is involved in H_2 oxidation, this finding may suggest that *Caldisericum* is supplementing organotrophic S^0 dependent growth with electrons from H_2 in a manner similar to that described for several other thermophilic S^0 reducers when grown organotrophically (Amenabar et al 2018, Jochimsen et al 1997). Of the hydrogenase transcripts detected, the majority were attributed to expression of homologs of [NiFe]-hydrogenases predicted to be capable of H_2 oxidation (Fig. 2.6A), consistent with the hypothesized role of H_2 as an electron donor supporting the SJ3 community.

To further assess the role of H_2 in supporting primary production in SJ3, the most abundant bins (>2.0% relative abundance) were screened for homologs of genes encoding enzymes with key functional roles in characterized CO_2-fixation pathways (Berg et al 2010). The genome bin closely affiliated with *Sulfurihydrogenibium* codes for proteins involved in the reductive citric acid cycle, including *aclB* (Fig. 2.6B). Genome bins most closely related to the candidate division Acetothermia, DHVE2 Archaea, Elusimicrobia, and to *Thermus aquaticus* coded for homologs of enzymes involved in the reductive acetyl-CoA pathway including *fhs* (Fig. 2.6B). The genome bin most closely affiliated with *Geobacter* sp. GB1 coded for homologs of enzymes involved in the 3-hydroxypropionate-4-hydroxybutyrate cycle and the dicarboxylate-4-hydroxybutyrate cycle, including *abfD* (Fig. 2.6B), which is implicated in both aforementioned cycles.

Transcripts of genes that are diagnostic for each of these pathways (*aclB*, *fhs*, *abfD*) were detected in five of the six abundant genome bins that encoded CO_2 fixation pathways (Fig. 2.6B). Among these, the most abundantly expressed gene was a *fhs* homolog from the bin closely affiliated with the candidate division Acetothermia, which also exhibited high levels of expression of a putative oxidative [NiFe]-hydrogenase (Fig. 2.6A). Transcripts of *fhs* from bins affiliated with the DHVE2 Archaea and Elusimicrobia were also detected, as were transcripts of putative oxidative [NiFe]-hydrogenases in these bins. Overall, transcripts for genes involved in CO_2 fixation and H_2 oxidation were detected in 5 of the same bins. This suggests that H_2 supports primary production in the gas-rich SJ3 hot spring, and presumably other hot springs where lithogenic H_2 is common.

Conclusions

Data presented here suggest a model for H_2 production in the YNP subsurface that involves interaction of water with crustal ferrous iron-bearing minerals in the subsurface, with subsequent partitioning of H_2 into the vapor phase following decompressional boiling of ascending hydrothermal fluids (Fig. 2.2). Variable input of vapor into springs yields substantial differences in H_2 concentration, with the highest concentrations in springs located in highly fractured bedrock and at high elevation that together facilitate gas exhalation and migration. Thus, springs that are sourced with gas resulting from subsurface phase separation, which are concentrated near the caldera boundary and at high elevation (e.g., SJ, HSB, Washburn), are more likely to host abundant H_2-dependent populations. Indeed, SJ springs, which exhibit H_2 concentrations that are the highest reported in YNP, are located near the caldera boundary, at a topographic high, and host populations that appear to use H_2 as a reductant. Similarly, microbial populations in springs in the high elevation areas of Washburn and HSB, as well as other geologically similar areas of the Park, are also enriched in homologs of hydrogenase proteins predicted to be capable of H_2 oxidation (Fig. 2.3).

The presence and expression of genes that encode oxidative or bidirectional hydrogenases in dominant SJ3 populations suggest they are adapted to use H_2 as an electron donor. The same is true for populations that inhabit springs that are similarly sourced with vapor phase gases that have mixed with near surface meteoric water (e.g., Washburn and HSB), since they too exhibit high concentration of H_2 and are enriched in genes that encode hydrogenases with similar predicted physiological biases. Reducing

equivalents from H_2 oxidation are likely to support primary production as indicated by cultivation assays and by expression of genes that encode key enzymes involved in CO_2 fixation pathways in the SJ3 community. Taken together, we suggest that lithogenic H_2 is sourced from hydrothermal water-rock interactions in the subsurface of YNP and is concentrated in certain geothermal areas by the geological process of phase separation. Springs that are sourced by vapor phase gases have higher concentrations H_2 that, in turn, select for populations capable of using H_2 as a reductant during the assembly of resident microbial communities. It is likely that the links defined between the geological processes that generate and source springs with H_2 and the distribution of H_2-dependent microbial populations in those springs extend to other hydrothermal systems where similar geological processes are thought to take place.

Acknowledgements

This work was supported by grants from the NASA Exobiology and Evolutionary Biology program (NNX13AI11G) and by the NASA Early Career Fellowship program (NNX17AH10G) to ESB. The NASA Astrobiology Institute is supported by grant number NNA15BB02A to ESB, JRS, ELS, and TMH. MRL acknowledges support from the NASA Earth and Space Science Fellowship program (NNX16AP51H). The authors declare that they have no conflict of interest.

Figures

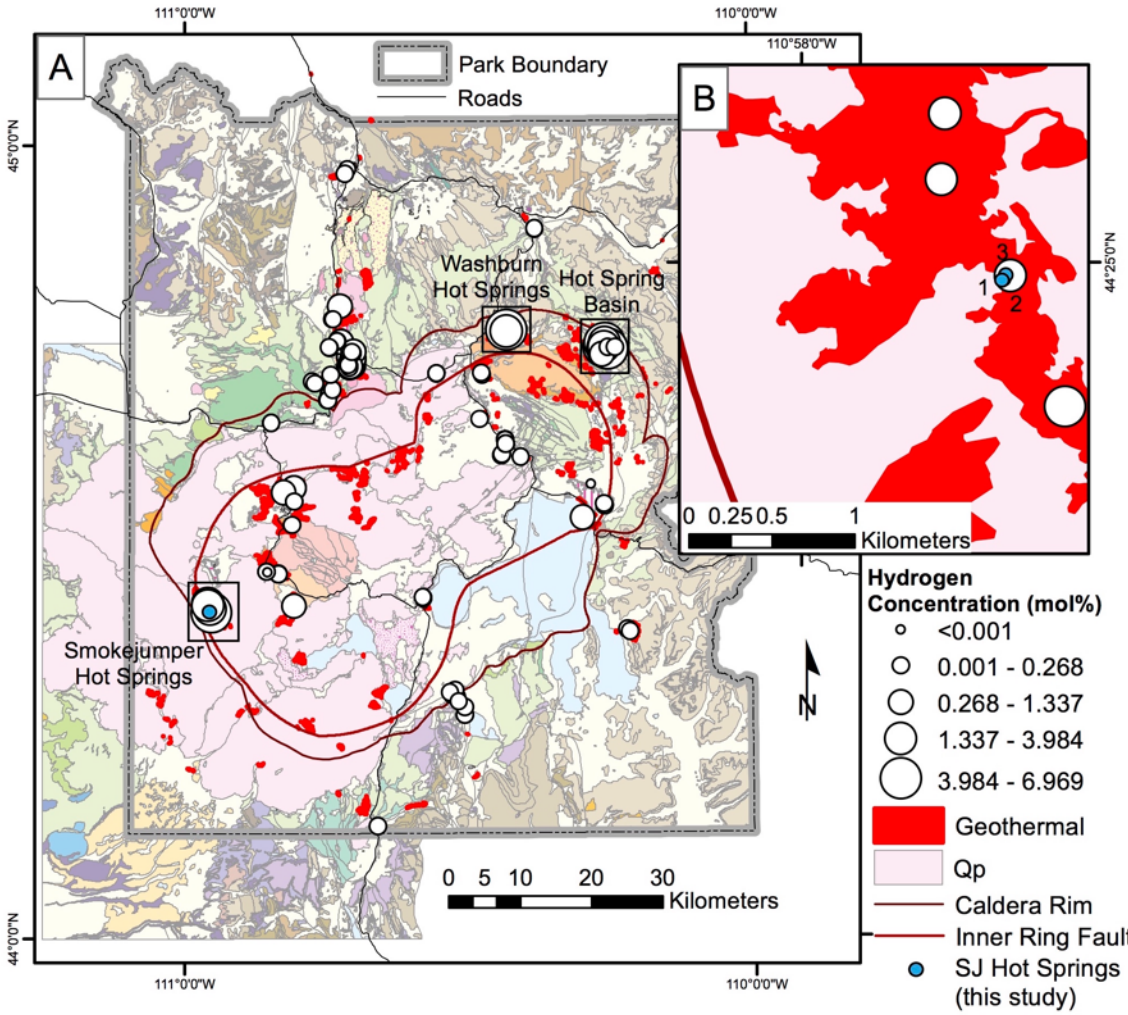


Fig. 2.1. Map of Yellowstone National Park (YNP). A) Map of YNP constructed with geothermal, road, geology, and Yellowstone lake (light blue) reference layers. Quaternary plateau rhyolite (Qp) is indicated. Other lava flows and rock types corresponding to the geological map are reported previously (Christiansen 2001) and are not defined here. The probable inner and outer caldera rim delineating the margins of the ring-fracture zone are indicated. Geology, caldera rim, and road reference layers from the United States Geological Survey Geographical Information Systems (USGS GIS) database. Other geology reference legend and rock type information is from the USGS GIS database (Christiansen 2001). Dot size indicates mole percent hydrogen [H_2 (mol%)] relative to all other gases (H_2O free). Gas data and layers modified from USGS GIS database (Bergfeld et al 2014). B) Inset map of partial Madison Plateau geothermal area including the Smokejumper hot springs area. Geology reference layers are the same as in Fig. 2.1A (Christiansen 2001). H_2 (mol %) concentrations plotted from previous USGS data (Bergfeld et al 2014). Location of SJ sites 1, 2, and 3 are plotted as small blue circles.

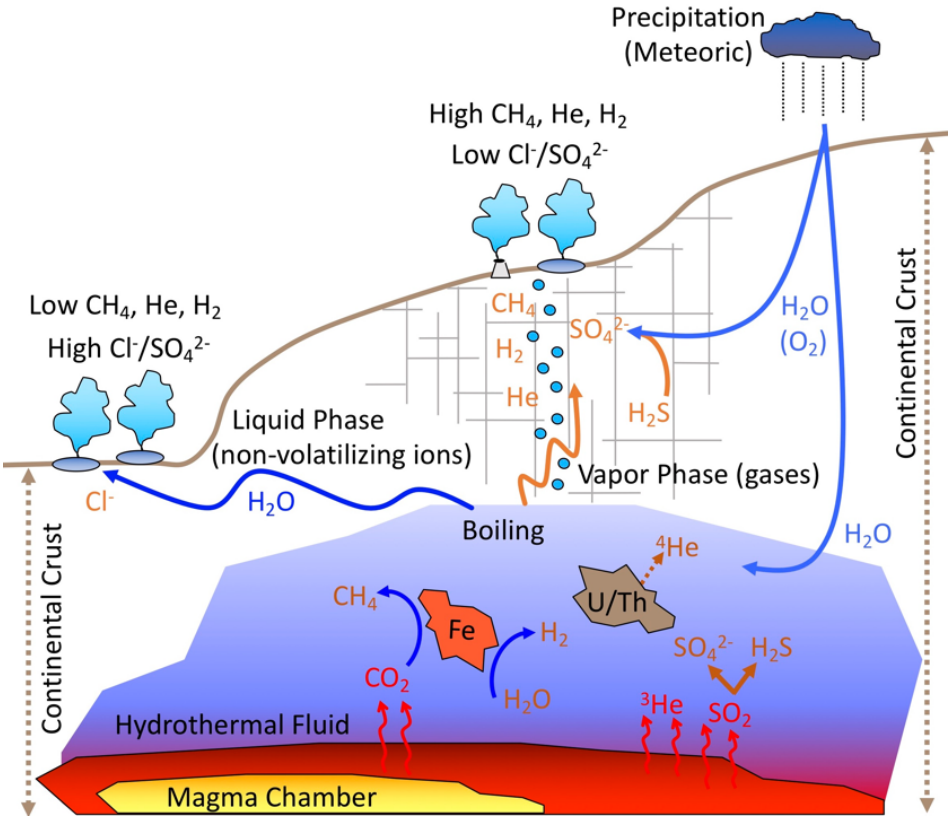


Fig. 2.2. Schematic summarizing a conceptual model of interactions between subsurface crustal minerals and water leading to production of gas in hydrothermal fluids, developed based on geochemical and isotopic proxies (Fig. 2.8/S2). In this model, interaction between water and crustal minerals yields hydrogen (H_2) and interaction between H_2 and carbon dioxide (CO_2) yield methane (CH_4). ^{3}Helium (^{3}He) is sourced from the mantle while ^{4}He is produced by α-decay of uranium and thorium that can occur anywhere in the crust. Subsurface hydrothermal fluids (recharged by meteoric water) undergo decompressional boiling during their ascent to the surface and partition into a vapor and a liquid phase. Vapor phase influenced springs have low concentrations of non-volatilizing ions such as chloride (Cl^-) but are enriched in hydrogen sulfide (H_2S), sourced from disproportionation of volcanic sulfur dioxide (SO_2). The oxidation of H_2S with oxygen (O_2) in near-surface waters leads to increased sulfate (SO_4^{2-}) concentrations (low Cl^-/SO_4^{2-} ratio) in vapor phase influenced springs. Liquid phase influenced springs are enriched in non-volatilizing ions (e.g., Cl^-) but not in H_2S, the primary source of SO_4^{2-} in spring waters, leading to high Cl^-/SO_4^{2-} ratios in spring waters. He, H_2, and CH_4 gases sourced from the mantle and the crust (Fig. 2.8/S2) partition into the vapor phase during boiling and thus are enriched in springs in hydrothermal areas that are at higher elevation (Fig. 2.7/S1) and in areas with highly fractured bedrock (as indicated by grey hashes) that allow for vapor to rise more easily, characteristics that enhance separation and concentration of gases, respectively. Colors of text and arrows correspond with hypothesized sources of geochemical constituents (i.e. red for magmatic-sourced gases, blue for H_2O or reactions with H_2O, orange for dissolved ions or gases in phase separated fluids).

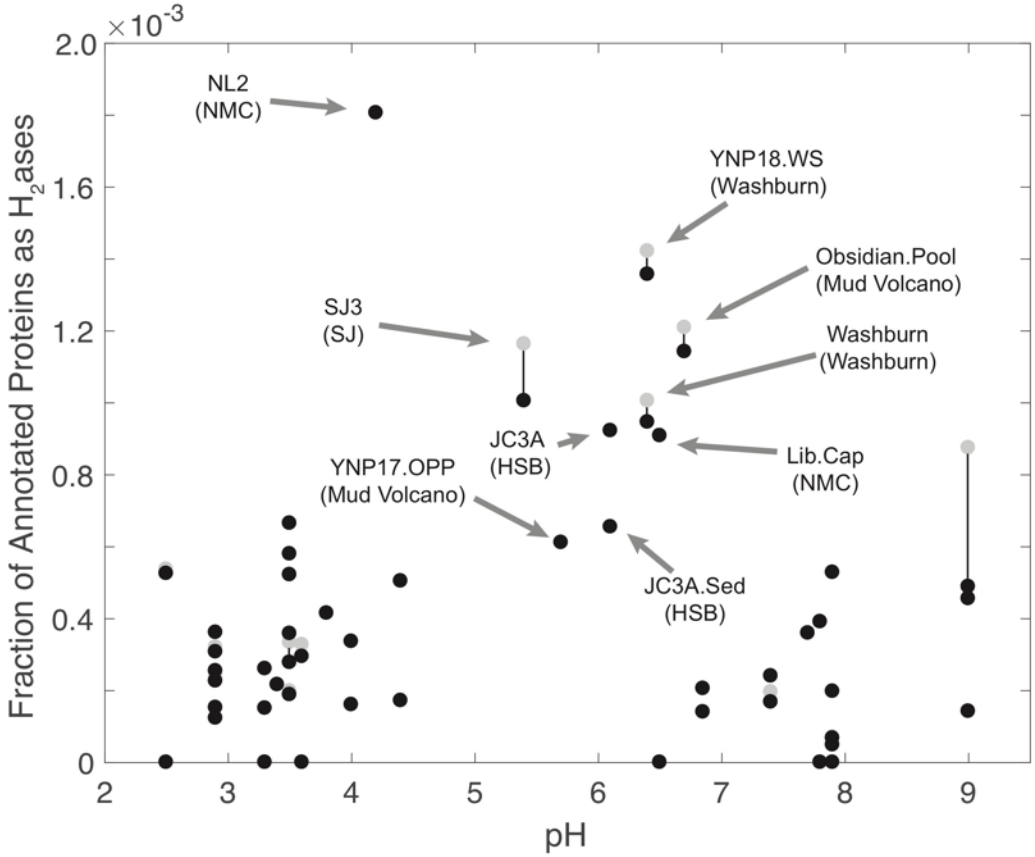


Fig. 2.3. Prevalence of hydrogenase protein encoding gene homologs among 50 chemosynthetic community metagenomes from YNP hot springs plotted as a function of spring pH. Black dots depict the prevalence of homologs of the catalytic subunits of group 1 and 3 [NiFe] hydrogenase that are inferred to be physiologically biased toward H_2 uptake or that are bidirectional (exhibit minimal bias). Total homologs were summed from total Kyoto Encyclopedia of Genes and Genomes (KEGG) KEGG Orthology (KO) annotations across all available hydrogenase KEGG KOs for a given metagenome, followed by normalization to the total number of KEGG KO annotations within each metagenome. KO annotations for group 2 [NiFe]-hydrogenases are not in the KEGG database and homologs of these protein encoding genes were therefore not included in this calculation. Grey dots indicate the same data as the black dots, with the addition of homologs of catalytic subunits of group 4 [NiFe]-hydrogenases and [FeFe]-hydrogenases as a metric of total hydrogenase abundance in metagenomes. Data points for each spring are connected by a line; springs where grey dots are not visible indicate that those metagenomes did not encode a substantial number of group 4 [NiFe]-hydrogenases and [FeFe]-hydrogenases homologs. Additional details for the metagenome dataset are provided in Supplemental Dataset S1. Abbreviations: H_2ases, hydrogenases; NMC, Norris-Mammoth Corridor; SJ, Smokejumper hot springs, HSB, Hot Spring Basin.

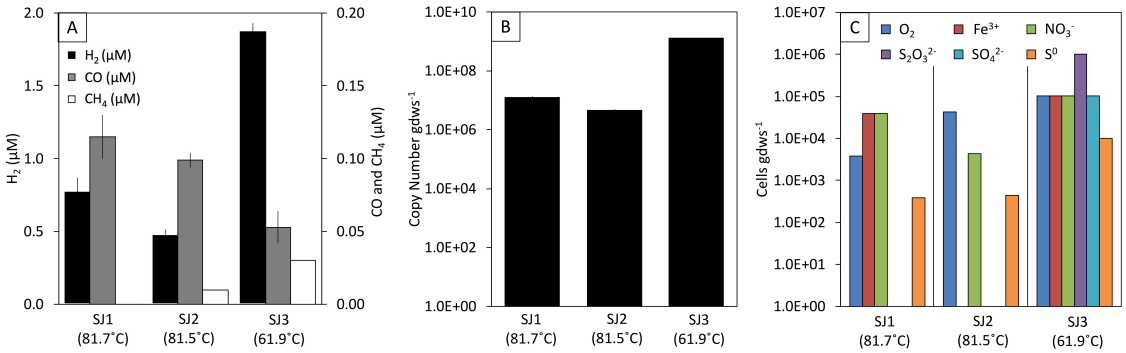
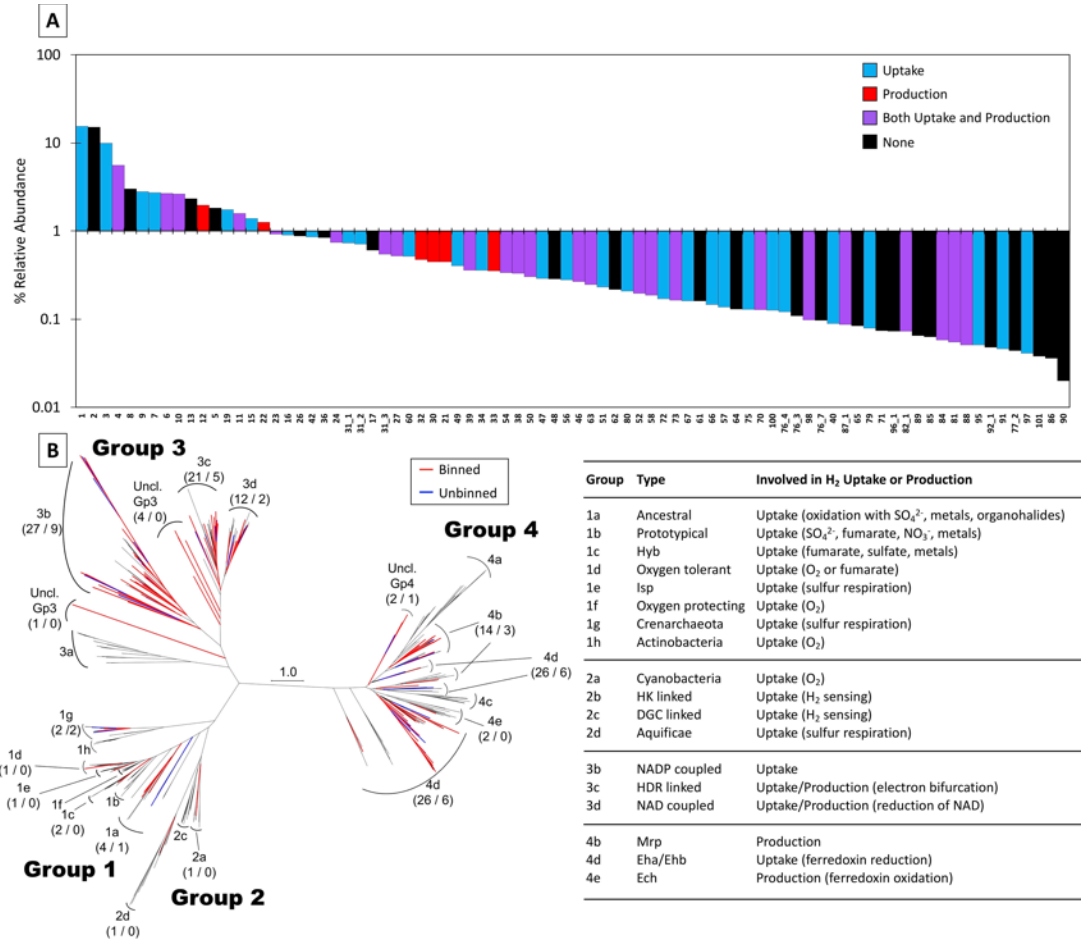


Fig. 2.4. Concentrations of dissolved gases, 16S rRNA genes, and cultivatable hydrogenotrophic and autotrophic cells in SJ hot springs. A) Dissolved hydrogen (H_2), carbon monoxide (CO), and methane (CH_4) in SJ hot springs. CH_4 in SJ1 is below the limits of detection (0.005 μM). B) Abundance of 16S rRNA gene templates associated with sediments sampled from SJ hot springs as determined using qPCR on extracted community DNA. Results are presented as the mean of triplicate qPCR assays; error bars represent the standard error of triplicate assays with some error bars too small to depict. C) Approximate number of autotrophic cells [per gram dry weight sediment (gdws)] capable of coupling H_2 oxidation to the reduction of specified oxidants (electron acceptors) as measured by dilution series cultivation experiments. A single dilution series cultivation assay was conducted for each condition for each spring.



Group	Type	Involved in H_2 Uptake or Production
1a	Ancestral	Uptake (oxidation with SO_4^{2-}, metals, organohalides)
1b	Prototypical	Uptake (SO_4^{2-}, fumarate, NO_3^-, metals)
1c	Hyb	Uptake (fumarate, sulfate, metals)
1d	Oxygen tolerant	Uptake (O_2 or fumarate)
1e	Isp	Uptake (sulfur respiration)
1f	Oxygen protecting	Uptake (O_2)
1g	Crenarchaeota	Uptake (sulfur respiration)
1h	Actinobacteria	Uptake (O_2)
2a	Cyanobacteria	Uptake (O_2)
2b	HK linked	Uptake (H_2 sensing)
2c	DGC linked	Uptake (H_2 sensing)
2d	Aquificae	Uptake (sulfur respiration)
3b	NADP coupled	Uptake
3c	HDR linked	Uptake/Production (electron bifurcation)
3d	NAD coupled	Uptake/Production (reduction of NAD)
4b	Mrp	Production
4d	Eha/Ehb	Uptake (ferredoxin reduction)
4e	Ech	Production (ferredoxin oxidation)

Fig. 2.5. Rank abundance plot (log transformed) of genome bins in SJ3 and the distribution of hydrogenase isoforms in these bins. A) Rank abundance plot for all genome bins reconstructed from the SJ3 metagenome. Genome bins are color coded to reflect the presence of genes coding for catalytic subunits of [NiFe]-hydrogenases or [FeFe]-hydrogenases (HydA) predicted to be physiologically or catalytically biased, respectively, towards the uptake of hydrogen (H_2), the production of H_2, or that encode isoforms of both enzyme types, as predicted based on phylogenetic affiliations (for [NiFe]-hydrogenases) or presence of key catalytic residues thought to confer catalytic bias (for HydA) as reported previously (Greening et al 2015, Therien et al 2017). Many bins have multiple hydrogenases as detailed in Dataset S2. Black bars depict genome bins that lack hydrogenase homologs. B) Phylogenetic distribution of [NiFe]-hydrogenase catalytic subunit homologs in the metagenome of SJ3 depicted with representatives of [NiFe]-hydrogenase groups, as defined previously (Greening et al 2015). Representatives with at least one homolog of each enzyme from the metagenome are overlaid in red if it was identified in a genome bin or blue if the homolog was unbinned. The scale bar shows the expected number of substitutions per site. Bootstrap values are provided in Dataset S3. The table corresponds with the groups and types of [NiFe]-hydrogenases in SJ3 (also reported in Dataset S2), their putative bias towards either H_2 uptake or production of H_2, and their overall predicted role in specified metabolisms, as reported previously (Greening et al 2015).

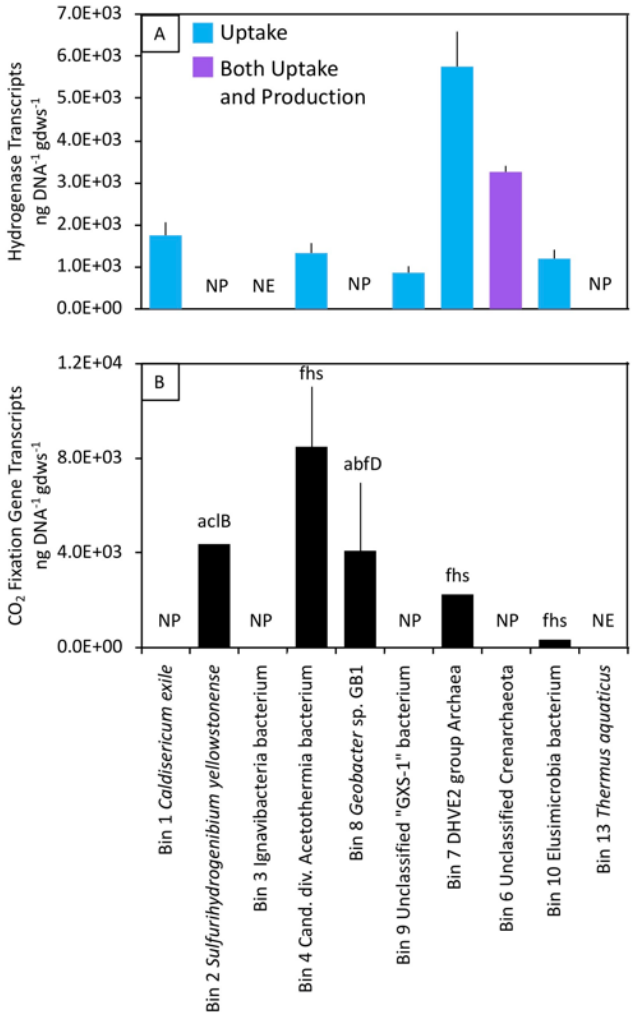


Fig. 2.6. Abundance of gene transcripts associated with dominant SJ3 metagenomic bins. A) Abundance of transcripts of genes encoding catalytic subunits of [NiFe]- or [FeFe]-hydrogenase associated with dominant SJ3 bins. Results are presented as the mean of triplicate qPCR assays; error bars represent the standard error of triplicate assays. Primers were not designed or tested for homologs in genome bins that were <2.0% of total abundance. Bars are colored blue if the transcript is predicted to be involved in H_2 uptake or purple if the transcripts are from multiple hydrogenases; with one putatively involved in H_2 uptake and one involved in H_2 production, as in Fig. 2.5. A few bins that encoded for multiple hydrogenases did not exhibit expression of more than one hydrogenase, and the bars for each bin are therefore colored to reflect the putative directionality of the single transcribed hydrogenase. B) Abundance of transcripts of genes encoding key proteins involved in CO_2-fixation pathways associated with dominant SJ3 bins. Results are presented as the mean of triplicate qPCR assays; error bars represent the standard error of triplicate assays. Gene transcripts involved in carbon fixation that were quantified in Fig. 2.6B include *aclB* (beta subunit of ATP-citrate lyase), *fhs* (formate-tetrahydrofolate ligase), and *abfD* (delta subunit of 4-hydroxybutyryl-CoA dehydratase). Abbreviations: NE, not expressed. NP, not present.

Supplemental Tables

Table 2.1/S1. Table of primers designed and used in the present study to quantify transcript abundances for target gene homologs. The sequence of the primers, the predicted length of the PCR amplicon, and optimized PCR annealing temperatures (determined with gradient PCRs) are provided. Abbreviations: *hydA* (α subunit of [FeFe]-hydrogenase), *hyhL* (large subunit of NADP coupled [NiFe]-hydrogenase), *hysA* **(**α subunit of [NiFe]-hydrogenase), *hoxH* (large subunit of NAD-coupled hydrogenase), *mvhA* (large subunit of HDR-linked hydrogenase), *hycE* (large subunit of formate hydrogenlyase), *aclB* (β subunit of ATP-citrate lyase), *fhs* (formate-tetrahydrofolate ligase), *abfD* (δ subunit of 4-hydroxybutyryl-CoA dehydratase).

Bin	Most closely related species	Gene	Forward Primer	Reverse Primer	Length (bp)	Annealing Temp. (˚C)
1	*Caldisericum exile*	*hydA*	TGCGAGGGCT CTTAGATTGC	TCGCACCCATC ATTTGCTGA	233	60
3	*Ignavibacteria* bacterium	*hyhL*	CCACGCAATG GAAACTGCAT	AGCGGATAGTC GTTTGGGTG	441	60
4	Cand. div Acetothermia bacterium	*hysA*	TTCGGTTTCTC ACCCCATCG	TCGCCTTCCGC CTTCTTTAG	289	60
9	Unclassified "GXS-1" bacterium (Spirochaetes)	*hoxH*	TACCGGGTGG CGTTAGAGA	TTGCAAGTGGT CCAACCCTG	381	61
7	DHVE2 group Archaea	*mvhA*	ACAAGATTGG AGGGGCATGG	TCTCTGGGAGC CTTAGGACC	347	60
6	Unclassified Crenarchaeota	*hyhL*	AAAGCTGGAT ACAGGGTGCC	CCGCTCAGGAT AAGGTTCCC	491	59
6	Unclassified Crenarchaeota	*hycE*	GCTATGCCAG GGTGAAGGTT	GTTCGCCTTCT CCTAGCCTC	491	59
10	*Elusimicrobia* bacterium	*hyhL*	TGTTGGTGGT GTGAGAAGGG	ATGGCTTTCCA TGCCCTGTT	329	59
2	*Sulfurihydrog-enibium yellowstonense*	*aclB*	GGGATAAAGT GGTAGAAGTC CATA	TTACCAACAAC AAAACCGGCTG	256	53
4	Ca. div. Acetothermia	*fhs*	TCATTCCCTAT GGCCGCATC	AGTTGAACTCG TGCTGGAGG	372	57
7	DHVE2 group Archaea	*abfD*	TGCGAAAATG GCAGGACTTG	CCCGTGGCACC TCCTTTTAT	273	56
8	*Geobacter* sp. GB1	*fhs*	TGCTGGCGCA ATGACAGATA	CTTTCTCGTATC GCTCGCCT	283	55
10	*Elusimicrobia* bacterium	*fhs*	ATAAAGGCAA GCGGGGCAAT	AAGAACTGCTG CATCTGGGA	270	57
13	*Thermus aquaticus*	*fhs*	GAGAAGGTGC TGGAGGCTTT	GGCCTCGCATC CTTCTTTCT	180	57

Table 2.2/S2. Geochemical measurements and 16S rRNA gene abundances (copy number $gdws^{-1}$) in waters and sediments, respectively, from Smokejumper springs. Data correspond with those presented in Fig. 2.4A, Fig. 2.4B, and Fig. 2.9/S3. Abbreviations: $gdws^{-1}$, per grams dry weight sediment; BDL, Below detection limit (0.005 µM for CH_4 and 0.011 mg L^{-1} for Fe^{2+}).

Site	pH	Temp. (°C)	SO_4^{2-} (mg L^{-1})	Cl^- (mg L^{-1})	SiO_2 (mg L^{-1})	S^{2-} (mg L^{-1})	Fe^{2+} (mg L^{-1})	$\delta^{18}O$	δD	H_2 (µM)	CO (µM)	CH_4 (µM)	Copy Number $gdws^{-1}$
SJ1	6.5	81.7	49.4	1.55	80.1	0.51	BDL	-16.2	-132.4	0.77	0.12	BDL	1.28×10^7
SJ2	7.6	81.5	36.0	1.53	137	0.78	BDL	-15.7	-130.5	0.47	0.10	0.01	4.57×10^6
SJ3	5.4	61.9	70.4	1.05	140	0.18	0.179	-17.0	-133.9	1.87	0.05	0.03	1.32×10^9

Table 2.3/S3. Approximate number of autotrophic cells (per gram dry mass sediment) capable of coupling H_2 oxidation to the reduction of specified oxidants (electron acceptors) as measured by dilution series cultivation experiments (one dilution series conducted for each condition for each spring). Data correspond with those presented in Fig. 2.4C.

Electron Acceptor	**SJ1 (pH 6.5, 82.5° C)**	**SJ2 (pH 7.5, 81.5° C)**	**SJ3 (pH 5.4, 61.9° C)**
O_2	3.9×10^3	4.3×10^4	1.0×10^5
Fe^{3+}	3.9×10^4	ND	1.0×10^5
NO_3^-	3.9×10^4	4.3×10^3	1.0×10^5
$S_2O_3^{2-}$	ND	ND	1.0×10^6
SO_4^{2-}	ND	ND	1.0×10^5
S°	3.9×10^2	4.3×10^2	1.0×10^4
Max.	3.9×10^4	4.3×10^4	1.0×10^6

Abbreviations: ND, not detected. Max., maximum number of H_2 oxidizing cells.

Table 2.4/S4. Metagenome sequence statistics for the SJ3 community.

Statistical Analyte	**Value**
Yield (Mbase)	60,754
Reads	290,691,162
Total bins	109
Total bins >50% completeness	82
N50	3907
Contigs >1000 bp	58,965
Contigs >2.5 kbp	931
Largest contig (bp)	438,391

Supplemental Figures

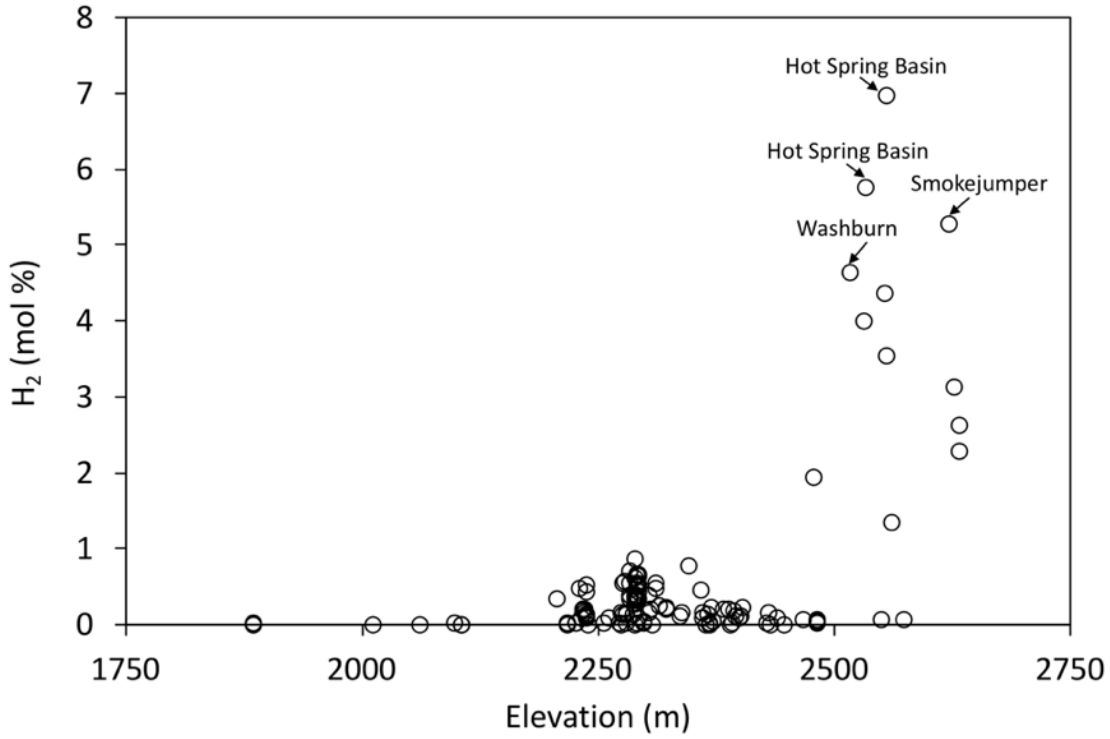


Fig. 2.7/S1. H_2 (mol %) in gas samples from 130 springs in YNP plotted as a function of the elevation (m) of those springs. Gaseous chemical and geospatial data for springs are reported previously (Bergfeld et al 2014). The geospatial data was used to calculate elevations of hot springs using ArcGIS and a 10 m digital elevation raster.

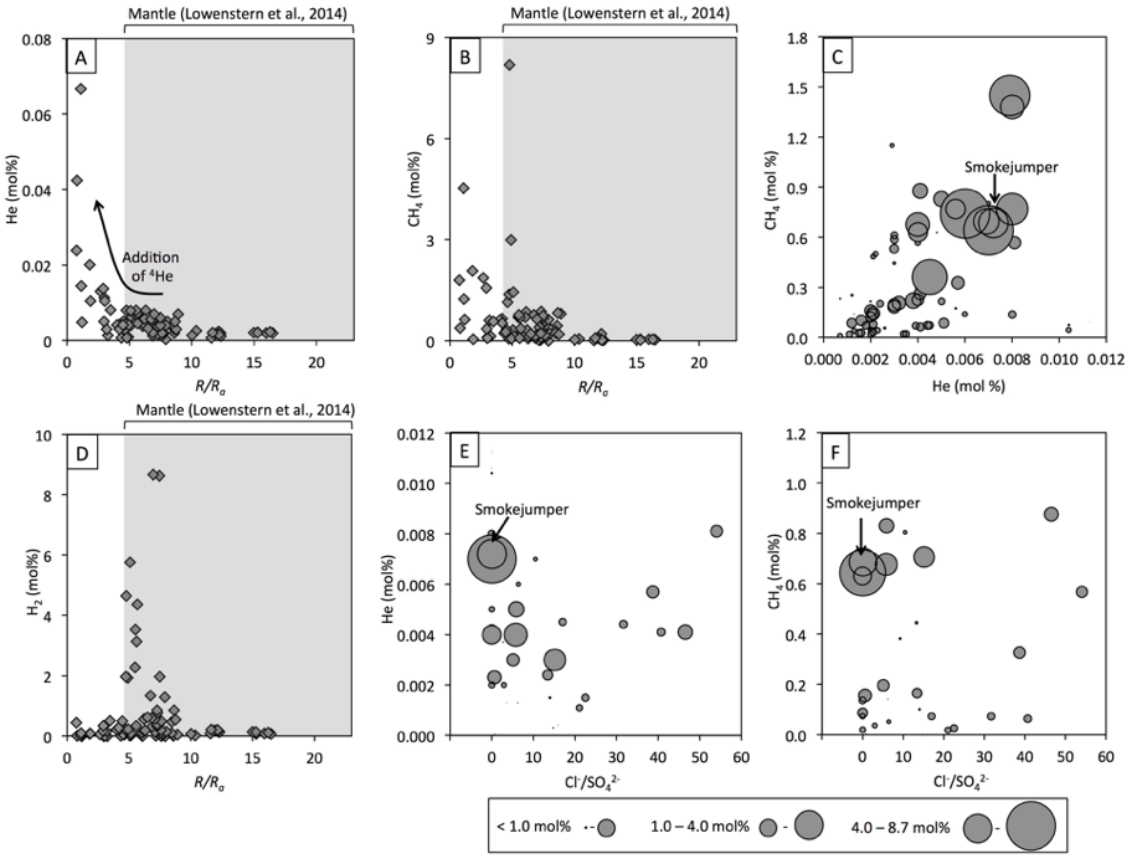


Fig. 2.8/S2. Helium (He), methane (CH_4), and hydrogen (H_2) in springs across YNP. A) He (mol %) plotted as a function of the ratio of the $^3He/^4He$ isotopes (R/R_a), relative to total He, in gas samples from 92 hot springs in YNP from (Bergfeld et al 2014). These 92 hot springs sampled do not overlap with the hot springs in Fig. 2.3. Grey shading delineates the range of R/R_a values in springs indicative of significant mantle input. End-member values are adapted from (Lowenstern et al 2014). B) Fractional abundances of CH_4 (mol % of total gas) plotted as a function of R/R_a ratios in the same gas samples from Fig. 2.8A/S2A. C) Fractional abundances of CH_4 (mol % of total gas) plotted as a function of the fractional abundance of He (mol % of total gas) in the same gas samples from Fig. 2.8A/S2A and 2.8A/S2B. He (mol % of total gas) is a proxy for subsurface input of gas, while CH_4 (mol % of total gas) is a proxy for crustal input of gas (see description in text). The size of the bubbles corresponds to the fractional abundance of H_2 in total gas (mol %). The sample corresponding to a site within the Smokejumper hydrothermal area is delineated. D) H_2 (mol %) plotted as a function of the ratio of the R/R_a in the same gas samples from Fig. 2.8A/S2A, 2.8B/S2B, and 2.8C/S2C. E) He (mol %) plotted as a function of the ratio of chloride (Cl^-) (mol L^{-1}) to sulfate (SO_4^{2-}) (mol L^{-1}) in the waters of a subset of samples presented in Fig. 2.8A-D/S2A-D. The size of the bubbles corresponds to the fractional abundance of H_2 in total gas (mol %). The sample corresponding to a site within the Smokejumper hydrothermal area is delineated (the same Smokejumper site in Fig. 2.8D/S2D). F) CH_4 (mol %) plotted as a function of the ratio of Cl^- to SO_4^{2-} in the waters of the same subset of samples as Fig. 2.8E/S2E. The size of the bubbles corresponds to the fractional abundance of H_2 in total gas (mol %).

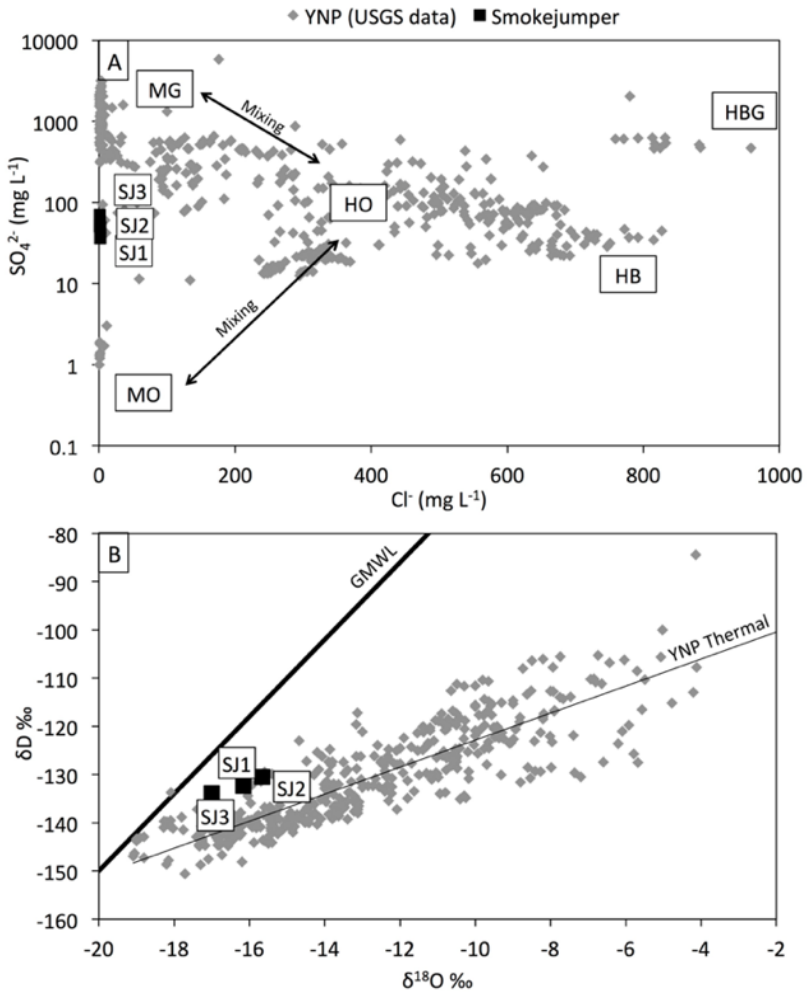


Fig. 2.9/S3. Sulfate (SO_4^{2-})/chloride (Cl^-) concentrations (mg L^{-1}), water isotope composition (‰), and spring classifications for three Smokejumper hot springs. A) SO_4^{2-} concentrations (mg L^{-1}) in hot spring waters plotted as a function of Cl^- concentrations (mg L^{-1}) in the same waters. The three samples in this study are plotted as black squares along with previous data collected from YNP springs between the years 2003-2013 (grey diamonds) (Ball et al 2006, McCleskey et al 2014). The classifications of hot spring fluid types, as described previously (Nordstrom et al 2009), are plotted for reference: hydrothermal water only (HO), hydrothermal water with subsurface boiling (HB), meteoric water only (MO), meteoric water with hot gas discharge (MG), and hydrothermal water with subsurface boiling and hot gas discharge (HBG). B) The isotopic composition of hydrogen (δ^2H) plotted relative to the isotopic composition of oxygen ($\delta^{18}O$) in hot spring waters. The three samples in this study are plotted as black squares along with previous data collected from YNP springs between the years 2003-2013 (grey diamonds) (Ball et al 2006, McCleskey et al 2014). The global meteoric water line (GMWL) is plotted for reference (Nordstrom et al 2009), and the slope of fractionated Yellowstone thermal waters (YNP thermal) is also plotted. The slope of the GMWL is 8; the slope of the YNP Thermal water line is 2.8.

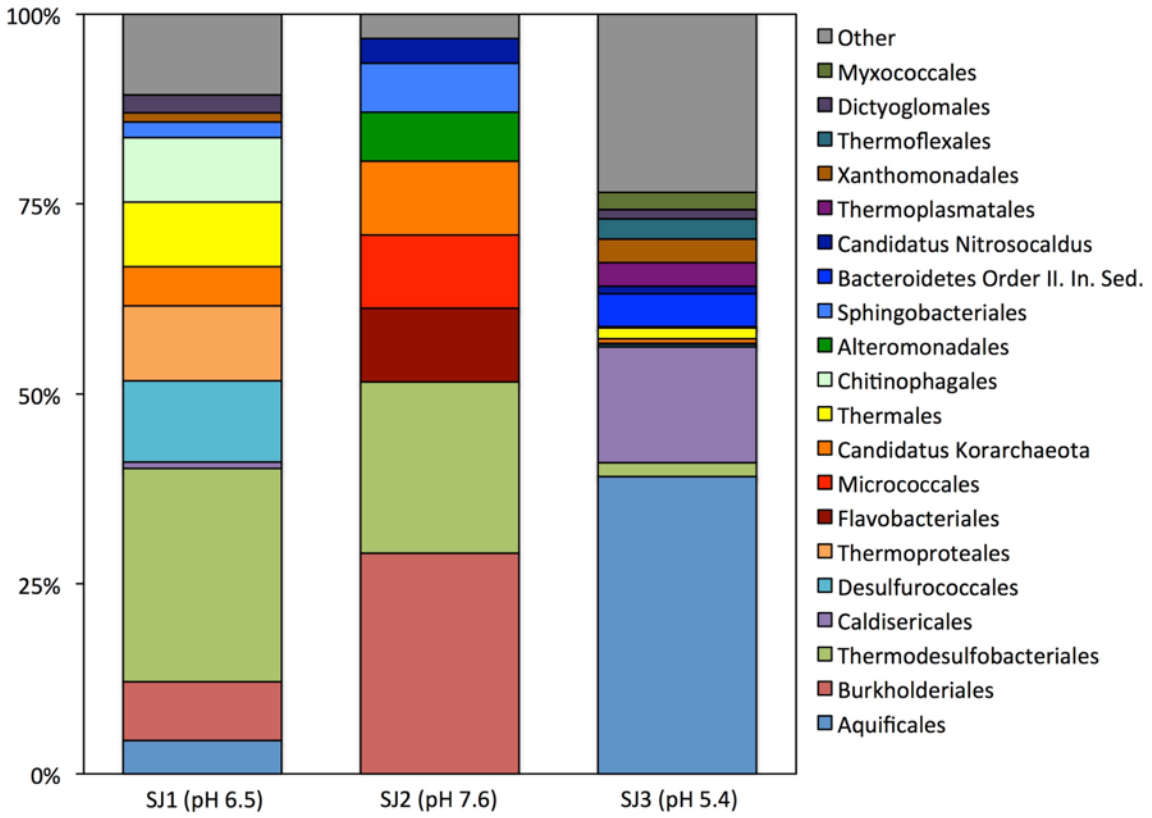


Fig. 2.10/S4. Taxonomic composition of 16S rRNA gene OTUs (as represented by their closest cultivated representatives) in sediments from SJ hot springs. Representative OTUs were binned at the order level, with orders that represent >1.0% of total sequences in a given sampling site included in the legend. Taxonomic bins that represented <1.0% of the total sequences from each assemblage were pooled and depicted as “Other”.

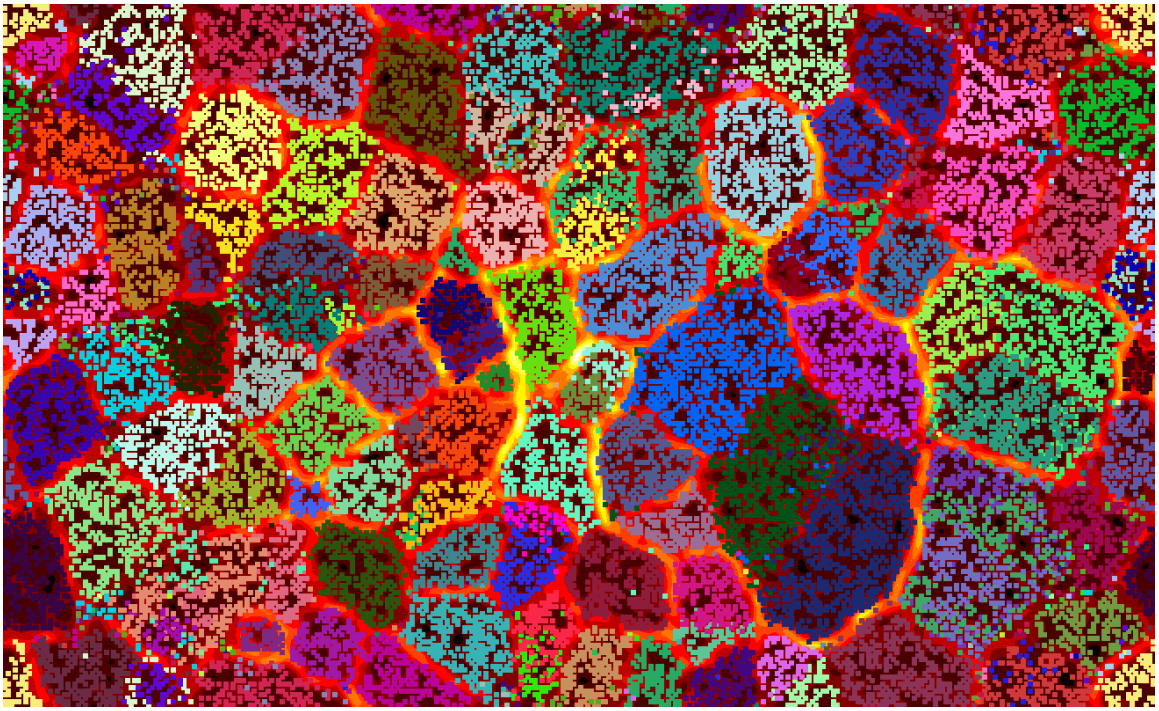


Fig. 2.11/S5. Emergent self-organizing map (ESOM) of assembled genome contigs. The tetranucleotide signatures were used to verify binning of contigs. Contigs are colored to reflect the bin they belong to, as assembled by MEGAHIT and as binned by MetaBAT.

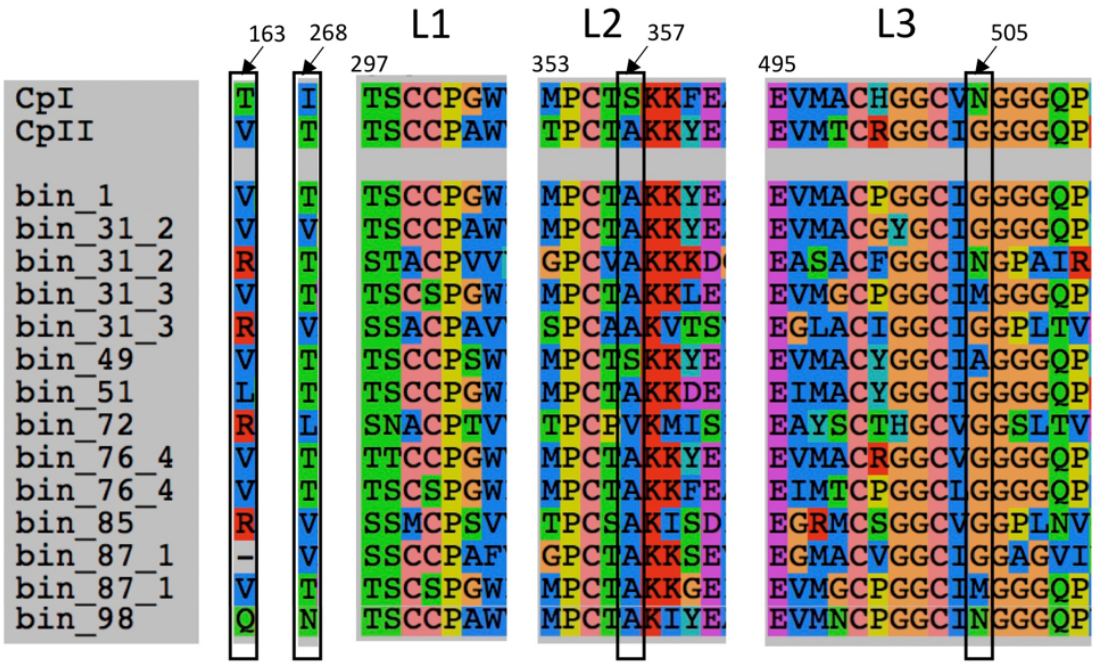


Fig. 2.12/S6. Multiple sequence alignment of H-cluster binding motifs for the large (catalytic) subunit of [FeFe]-hydrogenase (HydA) homologs from SJ3 bins. Multiple sequence alignment of HydA homologs identified in SJ3 bins (>50% abundance) and CpI and CpII HydA from *Clostridium pasteurianum* strain W5. CpI is catalytically biased towards H_2 production whereas CpII is catalytically biased towards H_2 uptake (Adams and Mortenson 1984, Therien et al 2017). The H-cluster binding motifs are designated L1, L2, and L3 and other key residues in the secondary coordination spheres of the H cluster thought to confer catalytic bias are indicated by black boxes (Therien et al 2017). Different colors are used to distinguish amino acids families (with amino acids in the same family colored the same color).

Additional Supplemental Datasets

Dataset S1 (separate excel file). Site and spring names, pH, hydrogenase abundances, Integrated Microbial Genome (IMG) identifier, and citations for geochemical data for 50 YNP hot spring chemosynthetic community metagenomes as presented in Fig. 2.3.

Dataset S2 (separate excel file). Estimated completeness, average coverage, total mapped reads, % relative abundance, taxonomic identity [determined by Maximum-likelihood phylogenetic reconstruction of a concatenation of ribosomal and housekeeping genes, as reported previously (Colman et al 2019)], and hydrogenase presence, type, and group designation for all bins that were >50% complete in the Smokejumper hot spring 3 (SJ3) metagenome. Data correspond with those presented in Fig. 2.5A and 2.5B. Hydrogenase homologs were identified using BLASTp at a cutoff of $E < 7e^{-4}$ for group 1 [NiFe]-hydrogenases, $E < 5e^{-4}$ for group 3 [NiFe]-hydrogenases, $E < 9e^{-52}$ for group 4 [NiFe]-hydrogenases, and $E < 1e^{-4}$ for [FeFe]-hydrogenases. BLASTp homologs used as queries are defined in the methods section of the primary text. Homologs were further classified according to previous classification schemes based on phylogenetic analyses (Greening et al 2015, Søndergaard et al 2016).

Dataset S3 (separate newick file). Phylogenetic reconstruction of homologs of the catalytic subunits of [NiFe]-hydrogenase proteins inferred from the metagenome of SJ3, depicted with representatives of groups of [NiFe]-hydrogenases as defined previously (Greening et al 2015). Bootstrap values are provided here for Fig. 2.5B. Homologs of both binned and unbinned [NiFe]-hydrogenase catalytic subunits were included.

References

Adams MWW, Mortenson LE (1984). The physical and catalytic properties of hydrogenase II of *Clostridium pasteurianum*. *The Journal of Biological Chemistry* **259:** 7045-7055.

Adams MWW, Eccleston E, Howard JB (1989). Iron-sulfur clusters of hydrogenase I and hydrogenase II of *Clostridium pasteurianum*. *Proceedings of the National Academy of Sciences of the United States of America* **86:** 4932-4936.

Amenabar MJ, Colman DR, Poudel S, Roden EE, Boyd ES (2018). Electron acceptor availability alters carbon and energy metabolism in a thermoacidophile. *Environmental Microbiology*. **20**: 2523-2537.

Ball JW, McCleskey RB, Nordstrom DK, Holloway JM (2006). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, 2003-2005. *USGS Open File Report*.

Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hugler M *et al* (2010). Autotrophic carbon fixation in archaea. *Nature Reviews Microbiology* **8:** 447-460.

Bergfeld D, Lowenstern JB, Hunt AG, Pat Shanks III WC, Evans WC (2014). Gas and Isotope Chemistry of Thermal Features in Yellowstone National Park, Wyoming. *USGS Scientific Investigations Report* **2011-5012**.

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30:** 2114-2120.

Boyd ES, Jackson RA, Encarnacion G, Zahn JA, Beard T, Leavitt WD *et al* (2007). Isolation, characterization, and ecology of sulfur-respiring *Crenarchaea* inhabiting acid-sulfate-chloride-containing geothermal springs in Yellowstone National Park. *Applied and Environmental Microbiology* **73:** 6669-6677.

Boyd ES, Hamilton TL, Spear JR, Lavin M, Peters JW (2010). [FeFe]-hydrogenase in Yellowstone National Park: evidence for dispersal limitation and phylogenetic niche conservatism. *The ISME Journal* **4:** 1485-1495.

Boyd ES, Lange RK, Mitchell AC, Havig JR, Hamilton TL, Lafreniere MJ *et al* (2011). Diversity, abundance, and potential activity of nitrifying and nitrate-reducing microbial assemblages in a subglacial ecosystem. *Applied and Environmental Microbiology* **77:** 4778-4787.

Boyd ES, Schut GJ, Adams MWW, Peters JW (2014). Hydrogen metabolism and the evolution of biological respiration. *Microbe* **9:** 361-367.

Bradley JA, Daille LK, Trivedi CB, Bojanowski CL, Stamps BW, Stevenson BS *et al* (2017). Carbonate-rich dendrolitic cones: insights into a modern analog for incipient microbialite formation, Little Hot Creek, Long Valley Caldera, California. *NPJ Biofilms* **32**.

Chapelle FH, Vroblesky DA, Woodward JC, Lovley DR (1997). Practical considerations for measuring hydrogen concentrations in groundwater. *Environmental Science and Technology* **31:** 2873-2877.

Christiansen RL (2001). The Quarternary and Pliocene Yellowstone Plateau Volcanic Field of Wyoming, Idaho, and Montana. USGS Professional Paper 729-G.

Colman DR, Feyhl-Buska J, Robinson KJ, Fecteau KM, Xu H, Shock EL *et al* (2016). Ecological differentiation in planktonic and sediment-associated chemotrophic microbial populations in Yellowstone hot springs. *FEMS* **92:** 1-13.

Colman DR, Poudel S, Hamilton TL, Havig JR, Selensky MJ, Shock EL *et al* (2018). Geobiological feedbacks and the evolution of thermoacidophiles. *The ISME Journal* **12:** 225-236.

Colman DR, Lindsay MR, Boyd ES (2019). Mixing of meteoric and geothermal fluids supports hyperdiverse chemosynthetic hydrothermal communities. *Nature Communications*. **10:** 1-13.

Dick GJ, Andersson AF, Baker BJ, Simmons SF, Thomas BC, Yelton AP *et al* (2009). Community-wide analysis of microbial genome sequence signatures. *Genome Biology* **10:** 1-16.

Djokic T, Van Kranendonk MJ, Campbell KM, Walter MR, Ward CR (2016). Earliest signs of life on land preserved in ca. 3.5 Ga hot spring deposits. *Nature Communications* **8:** 15263.

Fournier RO, White DE, Truesdell AH (1976). Convective heat flow in Yellowstone National Park. *Proceedings of the U N Symposium on Development and Use of Geothermal Resources* **1:** 731-739.

Fournier RO (1989). Geochemistry and dynamics of the Yellowstone National Park hydrothermal system. *Annual Review of Earth and Planetary Science* **17:** 13-53.

Giggenbach WF, Poreda RJ (1993). Helium isotopic and chemical composition of gases from volcanic-hydrothermal systems in the Phillipines. *Geothermics* **22:** 369-380.

Giggenbach WF, Sano Y, Wakita H (1993). Isotopic composition of helium, and CO_2 and CH_4 contents in gases produced along the New Zealand part of a convergent plate boundary. *Geochimica et Cosmochimica Acta* **57:** 3427-3455.

Gouy M, Guindon S, Gascuel O (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27:** 221-224.

Greening C, Biswas A, Carere CR, Jackson CJ, Taylor MC, Stott MB *et al* (2015). Genomic and metagenomic surveys of hydrogenase distribution indicate H_2 is a widely utilised energy source for microbial growth and survival. *The ISME Journal* **10:** 761-777.

Hamilton TL, Peters JW, Skidmore ML, Boyd ES (2013). Molecular evidence for an active endogenous microbiome beneath glacial ice. *The ISME Journal* **7:** 1402-1412.

Hoehler TM (2005). Biogeochemistry of dihydrogen. In: Sigel A, Sigel H, Sigel RKO (eds). *Biogeochemical Cycles of Elements*. Marcel Dekker: New York. pp 9-48.

Hurwitz S, Lowenstern JB (2014). Dynamics of the Yellowstone hydrothermal system. *Reviews of Geophysics* **52:** 375-411.

Jochimsen B, Peinemann-Simon S, Völker H, Stüben D, Botz R, Stoffers P *et al* (1997). *Stetteria hydrogenophila* gen. nov. and sp. nov., a novel mixotrophic sulfur-dependent crenarchaeote isolated from Milos, Greece. *Extremophiles* **1:** 67-73.

Kang DD, Froula J, Egan R, Wang Z (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*. **3:** 1-15.

Kelley DS, Karson JA, Fruh-Green GL, Yoerger DR, Shank TM, Butterfield DA *et al* (2005). A sepentinite-hosted ecosystem: the lost city hydrothermal field. *Science* **307:** 1428-1434.

Kita I, Matsuo S, Wakita H (1982). H_2 generation by reaction between H_2O and crushed rock: an experimental study on H_2 degassing from the active fault zone. *Journal of Geophysical Research* **87:** 10789-10795.

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9:** 357-359.

Li D, Liu CM, Luo R, Sadakane K, Lam TW (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Brujin graph. *Bioinformatics* **31:** 1674-1676.

Lin LH, Hall JR, Lippmann-Pipke J, Ward JA, Lollar BS, DeFlaun M *et al* (2005). Radiolytic H_2 in continental crust: nuclear power for deep subsurface microbial communities. *Geochemistry, Geophysics, Geosystems* **6**: 1-13.

Lindsay MR, Amenabar MJ, Fecteau KM, Debes RV, Fernandes MC, Urschel MR *et al* (2018). Subsurface Processes Influence Oxidant Availability and Chemoautotrophic Hydrogen Metabolism in Yellowstone Hot Springs. *Geobiology* **16:** 674-692.

Lovley DR (1985). Minimum threshold for hydrogen metabolism in methanogenic bacteria. *Applied and Environmental Microbiology* **49:** 1530-1531.

Lowenstern JB, Janik CJ (2003). The origins of reservoir liquids and vapors from The Geysers geothermal field, California (USA). In: Simmons SF, Graham I (eds). *Volcanic, geothermal and ore-forming fluids; rulers and witnesses of processes within the Earth*. Society of Economic Geologists Special Publication. pp 181-195.

Lowenstern JB, Bergfeld D, Evans WC, Hurwitz S (2012). Generation and evolution of hydrothermal fluids at Yellowstone: Insights from the Heart Lake Geyser Basin. *Geochemistry, Geophysics, Geosystems* **13:** 1-20.

Lowenstern JB, Evans WC, Bergfeld D, Hunt AG (2014). Prodigious degassing of a billion years of accumulated radiogenic helium at Yellowstone. *Nature* **506:** 355-358.

Lowenstern JB, Bergfeld D, Evans WC, Hunt AG (2015). Origins of geothermal gases at Yellowstone. *Journal of Volcanology and Geothermal Research* **302:** 87-101.

Luther GW 3rd, Findlay AJ, Macdonald DJ, Owings SM, Hanson TE, Beinart RA, Girguis PR (2011). Thermodynamics and kinetics of sulfide oxidation by oxygen: a look at inorganically controlled reactions and biologically mediated processes in the environment. *Frontiers in Microbiology* **2**: 1-9.

Markowitz VM, Chen IM, Chu K, Szeto E, Palaniappan K, Grechkin Y *et al* (2012). IMG/M: the integrated metagenome data management and comparative analysis system. *Nucleic Acids Research* **40:** D123-129.

McCleskey RB, Chiu RB, Nordstrom DK, Holloway JM (2014). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, beginning 2009. USGS Water Resources Dataset: doi:10.5066/F7M043FS.

Meyer J (2007). [FeFe] hydrogenases and their evolution: a genomic perspective. *Cellular and Molecular Life Sciences* **64:** 1063-1084.

Mikheenko A, Saveliev V, Gurevich A (2015). MetaQUAST: evaluation of metagenome assemblies. *Bioinformatics* **32:** 1088-1090.

Mori K, Yamaguchi K, Sakiyama Y, Urabe T, Suzuki K (2009). *Caldisericum exile* gen. nov., sp. nov., an anaerobic, thermophilic, filamentous bacterium of a novel bacteriual phylum, *Caldiserica* phyl. nov. originally called the candidate phylum OP5, and description of *Caldisericaceae* fam. nov., *Caldisericales* ord. nov. and *Caldisericia* classis nov. *International Journal of Systemic and Evolutionary Microbiology* **59:** 2894-2898.

Nakagawa S, Shtaih Z, Banta A, Beveridge TJ, Sako Y, Reysenbach A-L (2005). *Sulfurihydrogenibium yellowstonense* sp. nov., an extremely thermophilic, facultatively heterotrophic, sulfur-oxidizing bacterium from Yellowstone National Park, and emended descriptions of the genus *Sulfurihydrogenibium, Sulfurihydrogenibium subterraneum* and *Sulfurihydrogenibium azorense*. *International Journal of Systemic and Evolutionary Microbiology* **55:** 2263-2268.

Nordstrom DK, Ball JW, McCleskey RB (2004). Oxidation reactions for reduced Fe, As, and S in thermal outflows of Yellowstone National Park: biotic or abiotic? In: Wanty RB, Seal II RR (eds). *Water-Rock Interaction*. Taylor & Francis Group: London, UK.

Nordstrom DK, Ball JW, McCleskey RB (2005). Ground water to surface water: chemistry of thermal outflows in Yellowstone National Park In: Inskeep WP, McDermott (eds). *Geothermal Biology and Geochemistry in Yellowstone National Park*. Thermal Biology Institute: Montana State University, Bozeman, MT. pp 73-94.

Nordstrom DK, McCleskey RB, Ball JW (2009). Sulfur geochemistry of hydrothermal waters in Yellowstone National Park: IV Acid–sulfate waters. *Applied Geochemistry* **24:** 191-207.

Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A *et al* (2013). Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *Journal of Computational Biology* **20:** 714-737.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research* **25:** 1043-1055.

Peters JW, Schut GJ, Boyd ES, Mulder DW, Shepard ES, Broderick JB *et al* (2015). [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. *Biochimica et Biophysica Acta - Molecular Cell Research* **1853:** 1350-1369.

Poudel S, Tokmina-Lukaszewska M, Colman DR, Refai M, Schut GJ, King PW *et al* (2016). Unification of [FeFe]-hydrogenases into three structural and functional groups. *Biochimica et Biophysica Acta* **1860:** 1910-1921.

Reysenbach A-L, Liu Y, Banta AB, Beveridge TJ, Kirshtein JS, Schouten S *et al* (2006). A ubiquitous thermoacidophilic archaeon from deep-sea hydrothermal vents. *Nature* **442:** 444-447.

Rye RO, Truesdell AH (1993). The question of recharge to the geysers and hot springs of Yellowstone National Park. US Geological Survey Professional Paper**:** 93-384.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: open-source, platform-independent, community-supported software

for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75:** 7537-7541.

Schut GJ, Zadvornyy O, Wu C-H, Peters JW, Boyd ES, Adams MWW (2016). The role of geochemistry and energetics in the evolution of modern respiratory complexes from a proton-reducing ancestor. *Biochimica et Biophysica Acta* **1857:** 958-970.

Seemann T (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30:** 2068-2069.

Shock EL, Holland M, Meyer-Dombard DA, Amend JP, Osburn GR, Fischer TP (2010). Quantifying inorganic sources of geochemical energy in hydrothermal ecosystems, Yellowstone National Park, USA. *Geochimica et Cosmochimica Acta* **74:** 4005-4043.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W *et al* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7:** 1-6.

Søndergaard D, Pedersen CNS, Greening C (2016). HydDB: A web tool for hydrogenase classification and analysis. *Scientific Reports* **6:** 34212.

Spear JR, Walker JJ, McCollom TM, Pace NR (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 2555-2560.

Stamatakis A (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312-1313.

Stevens TO, McKinley JP (2000). Abiotic controls on H_2 production from basalt - water reactions and implications for aquifer biogeochemistry. *Environmental Science and Technology* **34:** 826-831.

Suzuki S, Kakuta M, Ishida T, Akiyama Y (2014). GHOSTX: An improved sequence homology search algorithm using a query suffix array and a database suffix array. *PLOS One* **9:** e103833.

Telling J, Boyd ES, Bone N, Jones EL, Tranter M, MacFarlane JW *et al* (2015). Rock comminution as a source of hydrogen for subglacial ecosystems. *Nature Geoscience* **8:** 851-855.

Therien JB, Artz JH, Poudel S, Hamilton TL, Liu Z, Noone SM *et al* (2017). The physiological functions and structural determinants of catalytic bias in the [FeFe]-hydrogenases CpI and CpII of *Clostridium pasteurianum* strain W5. *Frontiers in Microbiology* **8:** 1-11.

Truesdell AH, Fournier RO (1976). Conditions in the deeper parts of hot spring systems of Yellowstone National Park, Wyoming. US Geological Survey Open-File Report 76-428: Reston, Virginia.

Truesdell AH, Nathenson M, Rye RO (1977). The effects of subsurface boiling and diultion on the isotopic compositions of Yellowstone thermal waters. *Journal of Geophysical Research* **82:** 3694-3704.

van Geldern R, Barth JAC (2012). Optimization of instrument setup and post-run corrections for oxygen and hydrogen stable isotope measurements of water by isotope ratio infrared spectroscopy (IRIS). *Limnology and Oceanography: Methods* **10:** 1024-1036.

Vignais PM, Billoud B, Meyer J (2001). Classification and phylogeny of hydrogenases. *FEMS Microbiology Reviews* **25:** 455-501.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology* **73:** 5261-5267.

White DE, Muffler LJP, Truesdell AH (1971). Vapor-dominated hydrothermal systems compared with hot-water systems. *Economic Geology: Bulletin of the Society of Economic Geologists* **66:** 75-97.

Windman T, Zolotova N, Schwander R, Shock EL (2007). Formate as an energy source for microbial metabolism in chemosynthetic zones of hydrothermal ecosystems. *Astrobiology* **7:** 873-890.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL (2012). Primer-BLAST: a tool to design target specific primers for polymerase chain reaction. *BMC Bioinformatics* **13:** 1-11.

CHAPTER THREE

SUBSURFACE PROCESSES INFLUENCE OXIDANT AVAILABILITY AND CHEMOAUTOTROPHIC HYDROGEN METABOLISM IN YELLOWSTONE HOT SPRINGS

Contribution of Authors and Co-Authors

Manuscript in Chapter 3

Author: Melody R Lindsay

Contributions: Designed and conducted field and lab experiments, analyzed data, and contributed to the writing of the manuscript.

Co-Author: Maximiliano J Amenabar

Contributions: Contributed to field experiments and provided feedback on drafts of the manuscript.

Co-Author: Kristopher M Fecteau

Contributions: Contributed to geochemical analyses and provided feedback on drafts of the manuscript.

Co-Author: Randall V Debes II

Contributions: Contributed to geochemical analyses and provided feedback on drafts of the manuscript.

Co-Author: Maria C Fernandes Martins

Contributions: Contributed to microbiological analyses.

Co-Author: Kirsten E Fristad

Contributions: Contributed to geochemical analyses.

Co-Author: Huifang Xu

Contributions: Contributed to mineralogical analyses.

Co-Author: Tori M Hoehler

Contributions: Contributed to study design and provided feedback on drafts of the manuscript.

Co-Author: Everett L Shock

Contributions: Provided feedback on the design and implementation of geochemical analyses and provided feedback on drafts of the manuscript.

Co-Author: Eric S Boyd

Contributions: Designed field and lab experiments, contributed to the study design and sampling location selection, and contributed to the writing of the manuscript.

Manuscript Information

Melody R. Lindsay, Maximiliano J. Amenabar, Kristopher M. Fecteau, R. Vincent Debes II, Maria Clara Fernandes, Kirsten E. Fristad, Huifang Xu, Tori M. Hoehler, Everett L. Shock, and Eric S. Boyd*

*Corresponding Author

Geobiology

Status of Manuscript:
____ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
_ X_ Published in a peer-reviewed journal

Published July 2018, doi:10.1111/gbi.12308

SUBSURFACE PROCESSES INFLUENCE OXIDANT AVAILABILITY AND CHEMOAUTOTROPHIC HYDROGEN METABOLISM IN YELLOWSTONE HOT SPRINGS

Abstract

The geochemistry of hot springs and the availability of oxidants capable of supporting microbial metabolic processes are influenced by subsurface processes including the separation of hydrothermal fluids into vapor and liquid phases. Here, we characterized the influence of geochemical variation and oxidant availability on the abundance, composition, and activity of hydrogen (H_2)-dependent chemoautotrophs along the outflow channels of two paired hot springs in Yellowstone National Park. The hydrothermal fluid at Roadside East (RSE; 82.4°C, pH 3.0) is acidic due to vapor phase input while the fluid at Roadside West (RSW; 68.1°C, pH 7.0) is circumneutral due to liquid phase input. Most chemotrophic communities exhibited net rates of H_2 oxidation, consistent with H_2 support of primary productivity, with one chemotrophic community exhibiting a net rate of H_2 production. Abundant H_2-oxidizing chemoautotrophs were supported by reduction of oxygen, elemental sulfur, sulfate, and nitrate in RSW and oxygen and ferric iron in RSE; O_2 utilizing hydrogenotrophs increased in abundance down both outflow channels. Sequencing of 16S rRNA transcripts or genes from native sediments and dilution series incubations, respectively, suggest that members of the archaeal orders Sulfolobales, Desulfurococcales, and Thermoproteales are likely responsible for H_2 oxidation in RSE, whereas members of the bacterial order Thermoflexales and the archaeal order Thermoproteales are likely responsible for H_2 oxidation in RSW. These observations suggest that subsurface processes strongly influence spring chemistry and oxidant availability, which in turn select for unique

assemblages of H_2 oxidizing microorganisms. Therefore, these data point to the role of oxidant availability in shaping the ecology and evolution of hydrogenotrophic organisms.

Introduction

Primary production in environments with temperatures exceeding 73°C is driven by inorganic sources of chemical energy (Boyd et al 2010, Boyd et al 2012, Brock 1967, Cox et al 2011, Hamilton et al 2012). Inorganic electron donors commonly detected in hydrothermal systems include ferrous iron (Fe^{2+}), thiosulfate ($S_2O_3^{2-}$), elemental sulfur (S°), hydrogen sulfide (H_2S) and hydrogen (H_2) (Amend and Shock 2001, Shock et al 2010). Among these, the oxidation of H_2 generally releases the most energy regardless of whether it is coupled with reduction of oxygen (O_2), nitrate (NO_3^-), S°, sulfate (SO_4^{2-}), $S_2O_3^{2-}$, or ferric iron (Fe^{3+}) (Shock et al 2010, Spear et al 2005). This observation, when combined with the abundance of archaeal and bacterial 16S rRNA genes with affiliation to known H_2 oxidizing taxa in spring sediments and elevated concentrations of H_2 in spring waters, led to the hypothesis that primary production in high temperature hot springs in Yellowstone National Park (YNP) is driven by H_2 oxidation (Spear et al 2005). Further, it was suggested that the availability of oxidants, rather than the availability of H_2, is likely to dictate the distribution and ecology of hydrogenotrophs in high temperature hot spring environments where H_2 is detectable (Shock et al 2010, Spear et al 2005).

The availability and abundance of oxidants in hot springs are dependent on processes that take place deep in the subsurface, as well as those that take place near the

surface of these springs. The current model for the development of hot springs in YNP begins with injection of magmatic gases into a deeply seated hydrothermal aquifer(s) hosting fluids that have previously undergone and will continue to undergo water-rock reactions (Fournier et al 1976, Hurwitz and Lowenstern 2014, Rye and Truesdell 1993, Truesdell and Fournier 1976, Truesdell et al 1977, White et al 1971). As this hydrothermal water ascends and infiltrates through the crust, it is subsequently altered by processes such as water-rock interaction, boiling, and mixing with near-surface groundwater (Lowenstern et al 2012, Nordstrom et al 2009). Of particular importance to the generation of variation in hot spring geochemistry is subsurface boiling and resultant separation of subsurface fluid into a liquid phase and a vapor phase, the latter of which is thought to be enriched in H_2S relative to the liquid phase (Fournier 1989, Nordstrom et al 2009). The gas-rich vapor is free to migrate toward the surface where it can condense and interact with O_2-rich meteoric fluids. The near surface oxidation of H_2S/HS^- results in the formation of the intermediate $S_2O_3^{2-}$, which under acidic conditions (pH<4.0) can disproportionate to form solid-phase S° and sulfite (SO_3^{2-}) (Fournier 1989, Nordstrom et al 2004). In the presence of O_2, S° can be oxidized to form SO_4^{2-} and protons (H^+), and SO_3^{2-} can be oxidized to form SO_4^{2-} (Nordstrom et al 2005). Thus, hot springs with vapor-phase-influenced source waters often have a low pH (< 3.5) and tend to be enriched in SO_4^{2-} while hot springs with liquid-phase-dominated source water tend to have a lower concentration of SO_4^{2-} and are circumneutral to alkaline in pH (> 6.5) (Amenabar et al 2015). Protons generated from $H_2S/S°$ oxidation in vapor-phase-impacted environments can also contribute to the release of metals from altered host rock, including ferrous iron (Fe^{2+}) that can then be oxidized by thermophilic microorganisms leading to formation of

soluble ferric iron (Fe^{3+}) ions and Fe^{3+} containing solid phases (Inskeep et al 2005, Kozubal et al 2012). Thus, vapor-phase-influenced hot spring environments can also be enriched in both Fe^{2+} and Fe^{3+} (as soluble and solid phases) when compared to liquid-phase-dominated hot spring environments (Ball et al 2010).

The availability of oxidants can also vary spatially within a given spring. For example, the biologically-mediated oxidation of H_2S/HS^- could potentially be responsible for the decrease in H_2S/HS^- in spring outflow channels (Cox et al 2011, D'Imperio et al 2008, Nordstrom et al 2004). Alternatively, the volatilization of H_2S [pKa = 6.4 at 100°C (Amend and Shock 2001)] can decrease the concentration of dissolved H_2S/HS^- in acidic spring waters as they flow away from their sources (Nordstrom et al 2009). Removal of H_2S along spring outflow channels can affect the generation of S°, $S_2O_3^{2-}$, and SO_4^{2-}, since H_2S is the primary source for these compounds in hot springs (Nordstrom et al 2004). In contrast, the decrease in the temperature of hydrothermal waters along an outflow channel allows for higher dissolved O_2 concentrations since O_2 solubility increases with decreasing temperature (Shock et al 2010). Increased availability of O_2 can in turn affect the generation of oxidants such as S°, $S_2O_3^{2-}$, SO_4^{2-}, and Fe^{3+}, since O_2 is one of the few compounds with a redox potential that is positive enough to oxidize reduced forms of these compounds (Amend and Shock 2001).

In the present study, we hypothesized that geochemical variation and differences in the availability of oxidants resulting from subsurface boiling and phase separation, subsurface water-rock interactions, as well as surface degassing/ingassing exert strong

controls on rates of H_2 transformation and the abundance and composition of microbial communities. Since H_2 oxidizing populations are often autotrophic (Boyd et al 2014, Greening et al 2016, Vignais et al 2001), we also hypothesized that variations in oxidant availability would influence the productivity of chemosynthetic communities that were dependent on H_2. These interrelated hypotheses were examined in two springs in YNP that are separated by ~20 m and that represent examples of localized vapor-phase- and liquid-phase-influenced hot springs, Roadside East Spring (RSE; pH 3) and Roadside West Spring (RSW; pH 7), respectively. Here we present an integrated analysis of aqueous-, gas-, and solid-phase geochemical data, microbial activity and cultivation-based data, and quantitative RNA-based molecular data from samples spanning the outflow channels of RSE and RSW. The results provide new insights into the influence of subsurface processes on the geochemistry and availability of oxidants as well as their role in dictating the distribution and ecology of hydrogenotrophic and chemoautotrophic populations in high temperature ecosystems.

Materials and Methods

Description of sample sites and sample collection.

The two springs selected for this study are in the Nymph Lake area along the Norris-Mammoth Corridor (N 44°45'12.7"; W 110°43'30.8"). Five locations along the outflow channel at Roadside East hot spring (RSE) and 4 locations along the outflow channel at Roadside West hot spring (RSW) were examined (Fig. 3.1). Samples used for RNA-based analyses, dissolved inorganic carbon (DIC) assimilation assays, and geochemical analyses were collected on April 16, 2014. On March 31, 2015, H_2

transformation assays were conducted and samples for dilution to extinction cultivation assays were also collected from selected sample sites. Selected geochemical analyses were conducted on March 31, 2015 to identify appropriate locations along the outflow channels for targeted dilution to extinction and H_2 transformation assays to coordinate sampling campaigns.

Physical and chemical measurements.

The pH of spring waters at each sampling location was measured on-site with a YSI pH100CC-01 temperature-compensated pH meter (YSI Inc., Yellow Springs, OH). Conductivity and temperature were measured using a temperature-compensated YSI EC300 conductivity meter. Fe(II) and total sulfide (S^{2-}) were quantified in the field using Hach ferrozine pillows and Hach sulfide reagents 1 and 2, respectively, and a Hach DR/890 spectrophotometer (Hach Company, Loveland, CO). The detection limit for Fe(II) is 0.011 mg L^{-1} and the detection limit for S^{2-} is 0.01 mg L^{-1}.

Samples for cations and anions were filtered through 0.22 μM acrodisc Supor membrane filters, acidified, collected in plastic bottles, and stored at -20°C. Concentrations of the major anions Cl^-, NO_3^-, and SO_4^{2-}, and the major cations NH_4^+, Na^+, and K^+ were determined on separate Dionex DX-600 ion chromatography systems employing suppressed-conductivity detection. The anion system includes a potassium hydroxide eluent generator, carbonate removal device, and AG/AS11-HC columns. The anion system was equilibrated at 1 mM hydroxide for 6 min. prior to injection, after which it remained at 1 mM for the first 6 min. The hydroxide concentration was then

increased non-linearly to 51 mM from 6 to 37 min., followed by a linear increase to 55 mM from 37 to 45 min., after which it was returned to 1 mM. The flow rate was held constant at 1 mL $min.^{-1}$. The cations NH_4^+, Na^+, and K^+ were analyzed using CG/CS-16 columns under isocratic conditions (19 mM methanesulfonic acid, 0.5 mL $min.^{-1}$ flow rate) over 50 min. In both systems, samples were injected via AS40 autosamplers (5 mL vials, 2 injections per vial) onto 100 µL (anions) or 75 µL (cation) injection loops. Suppressor currents were 137 mA and 28 mA for anions and cations, respectively, and an external source of deionized water was used for regeneration. Uncertainties in reported concentrations are estimated to be <5%.

Samples for stable isotopic composition of H_2O (reported as δD and $\delta^{18}O$) were collected such that no air remained in the sample vials to minimize the possibility of isotopic fractionation between phases. Values of δD and $\delta^{18}O$ were determined using a Los Gatos Research DLT-100 isotope analyzer via cavity ring-down spectroscopy. Ten injections (0.93 µL injection volume) of each sample were made by a CTC PAL autosampler from 2 mL amber glass vials; the first 5 injections of each sample were discarded to remove sample memory effects. Data processing followed procedures previously reported (van Geldern and Barth 2012). Isotope ratios are reported relative to the Vienna Standard Mean Ocean Water (VSMOW) standard. Additional water isotope data (δD and $\delta^{18}O$) for YNP hot springs were compiled from USGS reports of samples collected between 2003-2005 and 2009-2013 (Ball et al 2006, McCleskey et al 2014).

DIC analyses were performed on filtered (0.2 μm) water samples stored in 40 mL amber vials that had been soaked in 10% nitric acid, rinsed with deionized water, and allowed to dry under laminar flow. The vials were capped with butyl rubber septa and stored at 4°C until processing. Analyses were completed using an OI Wet Oxidation TOC analyzer coupled to a Thermo Delta Plus Advantage mass spectrometer as previously described (Havig et al 2011). DIC in each sample was converted to CO_2 by addition of phosphoric acid. The ion chromatogram for the CO_2 molecular ion (44 m/z) was used for quantification of CO_2 by comparison to calibration curves constructed with sodium bicarbonate standards encompassing expected concentrations.

Samples for determining concentrations of dissolved gases (H_2, CO_2, and CH_4) were collected using the bubble strip method, as described by (Chapelle et al 1997, Spear et al 2005). Briefly, spring water was pumped at a flow rate of 100 mL $min.^{-1}$ through a 250-mL glass bottle bubble stripping device using a 12 V portable peristaltic pump and insulated, H_2-impermeable polyethylene tubing. A 20 mL "bubble" of ultra-high purity N_2 gas was introduced into the water-filled bottle and pumping was continued for an additional 20 min. to allow for gas equilibration. Following equilibration, the 20 mL bubble was collected with an air-tight syringe and transferred to a butyl rubber stoppered 24 mL serum bottle containing a 25% w/v sodium chloride (NaCl) solution. The gas was collected by displacement of the NaCl solution. A small volume (~ 4 mL) of NaCl was allowed to remain in the serum bottle and the bottle was inverted for storage until analysis via gas chromatography. Three replicate gas samples were collected from each sampling site in each spring.

At the lab, 5 mL of gas was injected into a SRI 8610C gas chromatograph with an injector temperature of 61°C and an oven temperature of 41°C (SRI instruments, Torrance, CA). Gases were separated on a 4.5 m x 0.125" OD Hayesep DB 100/120 packed column (Valco Instrument Company Inc., Houston, TX). H_2 was measured on a pulsed discharge helium ionization detector (PDHID) at a temperature of 100°C. CO_2 was measured on a thermal conductivity detector (TCD) at a temperature of 105°C while CH_4 was measured on a flame ionization detector (FID) at a temperature of 156°C. Helium was used as a carrier gas for all analyses. Gas concentrations were calculated using a standard curve (1 to 10,000 ppm calibration range) generated daily using certified (± 5%) standards (American Gas Group, Toledo, OH). Gas phase concentrations were converted to dissolved aqueous phase concentrations according to Henry's law, as previously described (Spear et al 2005).

X-Ray diffraction (XRD).

X-ray diffraction analyses were carried out on minerals in hot spring sediment samples with a Rigaku Rapid II X-ray diffraction system with a 2-D image plate (Mo *K*α radiation). Resultant 2-D patterns were converted to 1-D X-ray powder diffraction patterns using Rigaku's 2DP program. The samples were analyzed using reflection mode to detect non-crystalline silica phases.

RNA extraction and cDNA synthesis.

Triplicate subsamples of sediment were collected at each sample location using a sterilized spatula. Roughly 250 mg of sediment was placed in 2 mL cryotubes containing 500 µL of RNALater (Sigma-Aldrich, St. Louis, MO) and the tubes and their contents were immediately frozen on dry ice for transport back to the laboratory. The sediment samples were stored at -80°C until further processed. The FastRNA Pro Soil-Direct kit (MP Biomedicals, Irvine, CA) was used to extract RNA according to our previously published protocols (Hamilton et al 2013) with the following adjustment: the phenol-chloroform reagent was heated to 60°C prior to use. RNA extracts were treated with 2 units of DNAse (Sigma-Aldrich, St. Louis, MO) at room temperature (~21°C) for 10 min. Following DNAse treatment, the extract was checked for residual DNA by subjecting 1 ng of extract to 35 cycles of PCR using 'universal' 16S rRNA gene primers (505F/806R) following methods previously described (Hamilton et al 2013). Once determined to be free of PCR-amplifiable 16S rRNA genes (as a proxy for removal of total DNA), RNA extracts were quantified with the Qubit RNA HS Assay kit (Molecular Probes, Life Technologies, Madison, WI) and a Qubit fluorometer and aliquoted into 8 µL subsamples that were then frozen at -80°C. cDNA was synthesized from 1 ng of RNA in 20 µL reactions using the iScript cDNA Synthesis Kit (Bio-Rad Laboratories, Hercules, CA).

PCR, qPCR, and sequencing.

cDNA was subjected to amplification of archaeal, bacterial, and eukaryal small subunit rRNA gene transcripts using primers: 344F/915R (archaeal 16S rRNA; annealing temperature of 61°C), 1100F/1492R (bacterial 16S rRNA amplicons; annealing

temperature of 55°C), or A7F/570R (eukaryal 18S rRNA amplicons; annealing temperature of 42°C), as previously described (Hamilton et al 2013). Reactions were set up in a final volume of 50 µL, in proportions as previously described (Hamilton et al 2013). A master mix solution was added to approximately 1 ng of cDNA and reactions were subjected to the following cycling conditions in a Mastercycler Nexus gradient thermocycler (Eppendorf, Hamburg, Germany): initial denaturing at 95°C for 5 min. followed by 35 cycles of 95°C for 1 min., annealing at specified temperature for 1 min., and extension at 72°C for 1.5 min., with a final extension step at 72°C for 10 min. Positive control reactions were performed using 1 ng of plasmid DNA containing inserts corresponding to amplified regions of 16S rRNA genes. Amplicons used to generate plasmid standards were constructed with the primers and PCR protocols specified above using DNA from a high temperature hot spring environment in YNP (Hamilton et al 2013).

Quantitative PCR (qPCR) of 16S and 18S rRNA gene transcripts was performed using the same primers and concentrations described above and the SsoAdvanced Universal SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA, USA) in a final reaction volume of 20 µL. Roughly 1 ng of cDNA from each sample was subjected to the following reaction cycling conditions in a CFX Connect Real-Time system (Bio-Rad): an initial denaturing at 98°C for 0.5 min., followed by 35 cycles at 98°C for 0.5 min., annealing and elongation for 1 min., and eventually a melt curve of 65-95°C in 5 sec./step 0.5°C step increments. Plasmid standards for use in relating template copy number to threshold amplification signals were prepared as described previously (Boyd et

al 2011), and all standard curves were generated over 5 orders of magnitude from 5.0 x 10^1 to 5.0 x 10^5 templates (μL of extract)$^{-1}$. Negative control reactions were performed in the absence of added cDNA.

Archaeal, bacterial, and eukaryl small subunit (SSU) cDNA amplicons and rDNA amplicons from dilution cultivation experiments (described below) were submitted to MrDNA (Shallowater, TX) for barcoding and sequencing. Briefly, amplicons were barcoded using the aforementioned primers (forward primer barcoded) and the HotStarTaq Plus Master mix Kit (Qiagen, Valencia, CA) with the following conditions: 94°C for 3 min. followed by 5 cycles of 94°C for 30 sec., specified annealing temperature for each primer set (see above) for 40 sec., and 72°C for 1 min., after which a final elongation step at 72°C for 5 min. was performed. Illumina adapters were added following the confirmation of successful PCR amplification by electrophoresis in a 2% gel. Libraries were generated using the 2x300 MiSeq Reagent Kit v3 following the manufacturer's instructions (Illumina, San Diego, CA).

PCR amplicons were sequenced via multiplexed, paired-end Illumina MiSeq tag sequencing. Post sequence processing was performed with Mothur (ver. 1.36.1) (Schloss et al 2009) as previously described (Hamilton et al 2013), after merging the paired reads. Merging was only conducted for bacterial and eukaryal reads; archaeal reads were not merged due to poor overlap owing to amplicon length. Briefly, primers and adapters were removed from raw sequences and trimmed based on a Phred quality score of >25. The remaining sequences were trimmed to a minimum length of 250 bases for bacterial and

eukaryal sequences, and 200 bases for archaeal sequences. All sequences were subjected to a filtering step using the quality scores file to remove sequences with anomalous base calls. Unique sequences were aligned using domain specific SILVA databases (Quast et al 2013) and sequences were trimmed using defined start and end sites based on inclusion of 75% of the total sequences; sequences that started before or after these defined positions were removed without further consideration. The resultant unique sequences were pre-clustered to remove amplification and sequencing errors and chimeras were identified and removed using UCHIME (Edgar et al 2011). Operational taxonomic units (OTUs) were assigned at a sequence similarity of >97% using the nearest-neighbor method. The remaining sequences were randomly sub-sampled to normalize the total number of sequences in each library. Collectively, these steps resulted in a normalized size of 3,922, 21,802, and 1,023 SSU cDNA sequences for each archaeal, bacterial, and eukaryal library, respectively. Rarefaction curves were used to compute the percent coverage of the predicted taxonomic richness for each library. Sequences specific for each OTU were classified using the Bayesian classifier (Wang et al 2007) and the Ribosomal Database Project database, with manual verification using the nucleotide basic local alignment search tool BLASTn. The composition of the forward and reverse reads for archaeal 16S rRNA gene transcript libraries were similar and thus only the results of the forward reads are presented. Raw reads, quality scores, and mapping files for the archaeal, bacterial, and eukaryl SSU gene transcript libraries have been deposited in the National Center for Biotechnology Information short reads archive under accession number PRJNA353608.

Dissolved inorganic carbon assimilation assays.

Rates of dissolved inorganic carbon (DIC) assimilation were assessed using a microcosm-based approach, as previously described (Boyd et al 2009, Boyd et al 2012). In the case of RSE, sediments (~100 mg) were added to sterile 24 mL serum bottles, capped with a butyl rubber stopper, and purged with N_2 for ~5 min. in the field. In the case of RSW, a slurry was created with spring water and flocculent mat material and 1 mL of this slurry was added to sterile 24 mL serum bottles that were then sealed and purged with N_2 for ~5 min. Ten mL of spring water sampled from each location was added to each microcosm assay and the gas phase of each microcosm was equalized to atmospheric pressure using a sterile needle and syringe. Eight replicate microcosms were set up for each site, with 4 replicates being wrapped in aluminum foil to exclude light. The microcosm assays were initiated by injection of 10.0 μCi of [^{14}C] sodium bicarbonate ($NaH^{14}CO_3$) to each serum bottle to achieve a final concentration of 20 μM $NaH^{14}CO_3$. All microcosms were placed in a sealed bag (secondary containment) and were incubated in the outflow channel at each sampling site for 60 min. Assays were terminated in the field by freezing on dry ice and were stored at -20°C until processed.

In the laboratory, sealed microcosm assays were thawed at room temperature for approximately 1 hr. and were acidified to a pH of < 2 by injection of 1.0 mL of 1 N hydrochloric acid into the microcosm to volatize unreacted DIC. After acidification, the microcosms were unsealed and allowed to degas in a fume hood for an additional 2 hrs., with frequent manual stirring. Microcosm contents were filtered onto pre-weighed white 0.22 μm polycarbonate membranes, washed with 5 mL of sterile deionized water, and

dried overnight at 80°C. For RSE assays, dried filters and contents were weighed after drying to facilitate normalization of DIC assimilation rates to grams dry mass (gdm). For RSW assays, separate samples of each inoculum were dried and weighed to determine the gdm of sediment added to each assay. Dried filters were placed in scintillation vials and overlain with 10 mL of CytoScint ES liquid scintilliation fluid. The radioactivity measured in counts per minute (CPM) associated with each of the samples was measured on a Beckman LS 6500 liquid scintillation counter (Beckman Coulter, Inc., Indianapolis, IN) and converted to disintegrations per minute (DPM) using a quench curve. ^{14}C DPMs for DIC (RSW) or CO_2 (RSE) assays were converted to rates of uptake as previously described (Urschel et al 2015) using measured values of total DIC or dissolved CO_2, respectively. Dissolved CO_2 was used in place of DIC in RSE due to the rapid loss of DIC in acidic springs resulting from sample collection and filtration. The average and standard error of the mean of four replicate assimilation assays at each site are presented.

Hydrogen transformation assays.

Rates of net hydrogen (H_2) transformation were quantified using a microcosm-based approach. Triplicate microcosms were prepared in pre-sterilized capped (butyl rubber stoppers) 70-mL serum bottles that had been purged with N_2 for ~5 min. Thirty-one and a half mL of spring water sampled directly from each site was added to each serum bottle using a sterile syringe and needle, and 3.5 mL of a sediment and spring water slurry were added (total liquid plus sediment volume of 35 mL). Triplicate killed controls were prepared in the same manner and were amended with glutaraldehyde to a final concentration of 1% v/v and $HgCl_2$ to a final concentration of 500 μM. Microcosm

headspaces were amended with H_2 to a final concentration of 100 ppm, bottles were inverted (to prevent loss of H_2), and incubated in the spring at each site for 1 hr. Following incubation, 10 mL of the headspace of each microcosm was sampled for determination of the concentration of H_2. The 10 mL gas sample was collected in a serum bottle containing 25% NaCl by positive displacement as described above. The concentration of H_2 in a 5 mL subsample of gas was determined via gas chromatography, as described above. The average and standard error of triplicate biological assay headspace gas concentration determinations minus control assay (killed) headspace gas concentration determinations are presented.

Dilution to extinction cultivation assays.

Single tube dilution to extinction cultivation assays were used to estimate the number of viable autotrophic cells associated with hot spring sediments that are capable of coupling H_2 oxidation to the reduction of a suite of electron acceptors. Cultivation medium for assays contained NH_4Cl (0.33g L^{-1}), KCl (0.33g L^{-1}), $MgCl_2 \bullet 6H_2O$ (0.33g L^{-1}), KH_2PO_4 (0.33g L^{-1}), and $CaCl_2 \bullet 2H_2O$ (0.33g L^{-1}) in distilled water (Boyd et al 2007). The pH of the medium was adjusted to the pH measured at each of the sampling sites using dilute HCl. Ten mL of medium was dispensed into 20 mL Hungate tubes and sealed with butyl rubber stoppers with the exception of assays with S°. In this case, assays were left open during autoclaving and sterile S° (sterilized by baking at 100°C for 24 hrs) was then added to a final concentration of 5 g L^{-1}, followed by closure with sterile butyl rubber stoppers. All other assay tubes were sealed prior to autoclave sterilization.

Filter-sterilized and anaerobic Wolfe's vitamins and SL-10 trace elements were added to all vials to a final concentration of 1 mL L^{-1}. Electron acceptor amendment stock solutions of $MgSO_4 \bullet 7H_2O$, $NaNO_3$, and $FeCl_3 \bullet 6H_2O$ were prepared with distilled H_2O and were sterilized by filtration (0.22 μm). Solutions of $MgSO_4$, $NaNO_3$, and $FeCl_3$ were added to Hungate tubes to a final concentration of 10 mM. Hungate tubes and their contents (medium with vitamins, trace elements, and electron acceptor amendments) were purged for 15 min. with N_2 passed over heated (210°C) and H_2-reduced copper shavings to remove O_2. The headspaces of all tubes (anaerobic and aerobic) were purged for 5 min. with H_2. Two mL of the headspace gas (20% of the gas phase volume) was then replaced with CO_2 (final headspace concentration 80% H_2/20% CO_2). Air (and therefore O_2) was added to the headspace of aerobic tubes to achieve final concentrations of 16% v/v O_2 with a balance of 64% N_2 v/v, 16% H_2 v/v, and 4% CO_2 v/v for aerobic assays.

Assays were conducted on a select subsample of the sites. Assays were inoculated with 1 mL of a sediment and spring water slurry from each sampling site at RSE and from sites RSW1 and RSW3 (Fig. 3.1). This initial culture slurry was then immediately used to inoculate a dilution series that spanned a range from 10^{-1} to 10^{-8}. Each dilution series was incubated in the lab at a temperature corresponding to the temperature measured at the site from where the specific sediment slurry was sampled. Following 8 weeks of incubation, activities and/or cell numbers were determined for each condition at each site. Activities that were measured included the production of total S^{2-} (reduction of SO_4^{2-} or S°) as determined via the methylene blue reduction method (Fogo and Popowsky

1949) and the production of Fe^{2+} (reduction of Fe^{3+}) as determined using the ferrozine assay (Stookey 1970) in both inoculated assays as well as controls. Cell counts (used to evaluate all assays including O_2 and NO_3^- amendments) were conducted by subsampling 1 mL of each assay and incubating with SYBR gold nucleic acid gel stain (Life Technologies) at a final concentration of 1:4,000 (vol/vol). After incubating in the dark at room temperature with periodic mixing for 30 min., samples were filtered onto 0.2 μm black polycarbonate membranes and cells were counted using fluorescent microscopy. Dilutions that contained more than three orders of magnitude more cells than uninoculated controls were considered to have been inoculated with at least one viable cell capable of H_2 oxidation. Estimates of the number of H_2 oxidizing cells in the initial slurries were calculated using this assumption and were normalized to gdm in the original inoculum

One mL from the most dilute positive assay from each dilution series was transferred into 10 mL of fresh culture media, prepared as described above. Cultures were incubated for 3 weeks and growth was monitored via cell counting as described above. This transfer protocol was conducted a second time, and DNA was extracted from the final transferred enrichments following 3 weeks incubation using a modified phenol-chloroform extraction as previously described (Amenabar et al 2018). PCR amplicons for bacterial or archaeal specific 16S rRNA genes were generated and sequenced using the same protocols for the sediment SSU cDNA amplicons as described above. Sequence data for these enrichment assays is deposited in the NCBI short reads archive under accession number PRJNA430279.

Results

Geochemistry.

Concentrations of Cl^- and SO_4^{2-} from RSE and RSW were used to categorize these springs using a previously reported classification scheme (Fig. 3.8/S1, Table 3.2/S1) (Nordstrom et al 2009). When thus classified, both springs are interpreted to have input from deep-seated hydrothermal waters sourced from deep reservoirs. However, RSE waters are interpreted to be primarily influenced by subsurface vapor-phase, gas-rich fluids that have undergone mixing with precipitation-derived (meteoric) fluids (higher SO_4^{2-}/ Cl^- ratio), whereas RSW waters are interpreted to be primarily influenced by hydrothermal-only fluids (lower SO_4^{2-}/ Cl^- ratio) and to a lesser extent by subsurface vapor-phase fluids (Nordstrom et al 2009). The δD (‰) and $\delta^{18}O$ (‰) of waters sampled from both springs plot along a best fit line with a slope of close to 3 (Fig. 3.9/S2), the same value reported for heated (near-boiling) surface evaporation occurring under non-equilibrium conditions (Craig et al 1963, Giggenbach 1978, Nordstrom et al 2009). When considered in the context of the source of waters based on SO_4^{2-}/ Cl^- ratios, the spring water δD and $\delta^{18}O$ reveal differences that are consistent with mixing of RSE waters with precipitation-derived waters (i.e., isotopically more negative; Fig. 3.9/S2). Based on these geochemical data, RSE and RSW waters represent examples of springs that are influenced by phase separation, with vapor-phase fluids mixing with precipitation-derived waters (groundwater) in RSE and with liquid phase fluids primarily sourcing RSW. Importantly, these springs are separated by only ~20 meters, which may indicate that boiling, phase separation, and mixing with groundwater takes place in the near subsurface of these springs (Nordstrom et al 2005).

Sampling locations at RSE and RSW span ranges of temperatures (RSE 42.7°C to 82.4°C; RSW 44.1°C to 68.1°C) while the pH values of the two springs were nearly constant, 2.9-3.0 (RSE) and 7.0-7.2 (RSW) (Fig. 3.2A and 3.2B, Table 3.2/S1). Non-linear relationships between water temperature and distance from the spring source in RSE and RSW are likely due to differences in the geometry of the outflow channel (Fig. 3.1), which in turn likely influence water residence times and the extent of evaporative cooling. Cl^- concentrations can be used as a conservative tracer against which to evaluate the influence of evaporation on the composition of fluids as they flow downstream from their source. The concentration of Cl^- increased by 9.1% down the RSE outflow channel (from RSE1 to RSE5) and by 5.3% down the RSW outflow channel (from RSW1 to RSW4) (Fig. 3.2C and 3.2D; Table 3.2/S1). The increase in the concentration of Cl^- down the outflow channel of these two springs is interpreted to result from evaporative concentration, which is corroborated by similar changes in concentration of other conservative tracers such as Na^+, which increased by 8.3% in RSE and 4.0% in RSW. Likewise, the conservative tracer K^+ concentrations increased by 8.5% in RSE and 3.5% in RSW (Fig. 3.2C and 3.2D; Table 3.2/S1). and the $\delta^{18}O$ of hot spring waters decreased by 8.3% in RSE and 5.3% in RSW, while the slope of the δD as a function of $\delta^{18}O$ remained at 3 (Table 3.2/S1, Fig. 3.9/S2).

The concentration of SO_4^{2-}, after being normalized to account for evaporative concentrating effects, did not change significantly down the outflow channel of RSE but increased down the outflow channel of RSW by 6.7% (Fig. 3.2E and 3.2F; Table 3.3/S2).

In contrast, the concentration of HS^-/H_2S decreased down the outflow channels of both RSE and RSW by 95.3% and 38.6%, respectively. The decrease in HS^-/H_2S down the outflow channel of these springs may be due to a combination of volatilization and/or oxidation, but for RSE are hypothesized to reflect volatilization of uncharged H_2S [pKa of HS^-/H_2S is 6.4 at 80°C (Amend and Shock 2001)]. The decrease in HS^-/H_2S down the outflow channel of RSW (~ 4 μM) cannot account for the ~ 50 μM increase in the concentration of SO_4^{2-}, suggesting that oxidation of another sulfur compound (e.g., S°) may be contributing additional SO_4^{2-} to this spring. The concentration of NO_3^- decreased by 100.0% along the outflow channel of RSE and by 83.1% along the outflow channel of RSW, indicating that it is likely utilized by these communities as a nutrient source or as an electron acceptor. The concentration of Fe^{2+} decreased down the outflow channels of both springs (near lower detection limits) by 28.0% in RSE and 100.0% in RSW (Fig. 3.2E and 3.2F; Table 3.3/S2).

The concentrations of CH_4, H_2, and CO_2 in all measured spring waters at RSE and RSW were supersaturated with respect to atmospheric conditions. Concentrations of CO_2 and CH_4 were higher in source waters from RSW than in RSE, whereas H_2 was higher in RSE than in RSW. Source waters for each spring (RSE1 and RSW1) had the highest concentrations of dissolved CH_4, H_2, and CO_2 with concentrations decreasing down the outflow channels of both springs (Fig. 3.2G and 3.2H; Table 3.4/S3). A majority of the gas was removed from the water down the entire length of each outflow channel, with concentrations of CH_4, H_2, and CO_2 in RSE decreasing by 90.6%, 99.3%, and 95.9%, respectively, and in RSW by 65.5%, 72.6%, and 75.8% (Fig. 3.2G and 3.2H; Table

3.3/S2). Increases in the concentrations of dissolved H_2 were measured at the most downstream location at RSW (i.e., RSW4), likely due to H_2 production by photosynthetic or fermentative cells. This hypothesis is supported by the presence of putative phototrophic and fermentative taxa in the RSW4 community, as described below. Overall, the concentration of dissolved CO_2 was higher than the concentrations of dissolved H_2 or CH_4 at all sampling locations in both springs.

Mineralogy.

XRD analysis indicated that the assemblages of minerals present in sediments sampled along the outflow channels of RSE and RSW were similar, primarily consisting of quartz, alkali feldspar, tridymite, alunite, and kaolinite (Table 3.5/S4). While sediments lining the outflow channel of RSE visibly contained iron (hydr)oxides (Fig. 3.1A), minerals containing ferric iron were not detected by XRD indicating that these minerals comprised < 3% of the total mass of the sediments (Table 3.5/S4). Similarly, while sediments sampled from the source of RSW looked to be comprised of flocculent S° (visible as gray flocculent in the source pool, Fig. 3.1B), this mineral was not detected by XRD.

DIC and dissolved CO_2 assimilation.

Assimilation of DIC was detected in sediment-associated communities at all sampling locations in RSW and RSE (Fig. 3.3, Table 3.6/S5). While the difference in rates of DIC assimilation in light and dark assays for RSE locations 1 to 4 and for RSW locations 1 to 3 were minimal and generally not significantly different among these

sampling locations, rates of DIC assimilation in light incubations significantly exceeded those incubated in the dark by two orders of magnitude at RSE5 and RSW4, both of which comprised communities that included abundant phototrophs (discussed below).

Rates of DIC uptake in sediment communities sampled from RSE and RSW and that were incubated in the dark differed, and varied at different sites along the outflow channels of each spring. Rates of DIC uptake in sediment communities sampled along the outflow channel of RSW were greater than those along the outflow channel of RSE. The rates of DIC uptake in sediment communities sampled from RSE1 and RSW1, when incubated in the dark, were greater than that those measured in the downstream locations RSE3 and RSW2, respectively.

H_2 transformation.

Rates of net H_2 transformation in sediments sampled from RSE and RSW differed, and varied at different sites along the outflow channels of each spring (Fig. 3.4, Table 3.6/S5). Two RSE sediment-associated microbial communities (RSE1 and RSE5) were characterized by net increases in H_2 concentration during the incubation period, indicating net H_2 production. Headspace H_2 concentrations in microcosm assays containing sediments from the other three sites (RSE2, RSE3, RSE4) were characterized by a decrease in H_2 concentration during the incubation, indicating net H_2 consumption. The rates of net H_2 production at RSE1 and RSE5 were 4.20 and 3.61 nmol H_2 gdm^{-1} min^{-1}, respectively. The rates of net H_2 consumption in RSE2, RSE3, and RSE4 were 3.51, 1.69, and 4.34 nmol H_2 gdm^{-1} min^{-1}, respectively. Both RSW sediment-associated

communities examined (RSW1 and RSW3) exhibited decreases in the concentration of H_2 during the incubation period, indicating net H_2 consumption. The rates of net H_2 consumption were 5.34 and 24.85 nmol H_2 gdm^{-1} min^{-1}, respectively, in these communities.

Abundance of SSU rRNA gene transcripts.

The total abundance of SSU rRNA gene transcript templates (archaeal, bacterial, and eukaryal combined), as determined via qPCRs conducted on cDNA, was greater in RSW than RSE (Fig. 3.5). The abundance of archaeal 16S rRNA gene transcripts was similar in RSW and RSE, however, the abundance of bacterial 16S rRNA gene transcripts was 3 to 7 orders of magnitude greater at all sampling locations in RSW than RSE. Thus, active sediment communities in RSE are dominated by Archaea while those in RSW are dominated by Bacteria. In both RSW and RSE, the abundance of SSU rRNA gene transcripts, when considered together, increased down the outflow channels, with the highest copy numbers for each spring found at the lowest temperature sites: RSE5 ($1.2 \times 10^8 \pm 2.3 \times 10^7$ transcripts gdm^{-1}) and RSW4 ($2.3 \times 10^{11} \pm 8.8 \times 10^{10}$ transcripts gdm^{-1}), which are both within the pH/temperature/sulfide realm allowing for photosynthesis (Fig. 3.10/S3). As discussed below, these sites harbor bacterial and eukaryal populations inferred to be capable of photosynthesis, and both sites exhibited higher rates of DIC assimilation when incubated in the light when compared to the dark (Fig. 3.3), which are consistent with assimilation of DIC using light-energy.

Sediment community taxonomic diversity and composition.

Archaea. Archaeal 16S rRNA gene transcripts were detected in sediments sampled from all locations at RSE and RSW. A total of 3922 sequences were recovered from each individual site along each outflow channel with varying predicted coverages and diversity (Table 3.7/S6). The composition and structure of archaeal communities shifted down the outflow channels of both RSE and RSW (Fig. 3.6). The most abundant archaeal 16S rRNA gene transcript OTU in the RSE communities (maximum 73% of total sequences at RSE1, and found in smaller percentages in the other RSE sites) exhibited 100% sequence identity to the thermophilic, chemoautotrophic crenarchaeote *Metallosphaera yellowstonensis* MK1 within the order Sulfolobales (Fig. 3.6) (Kozubal et al 2011). The next most abundant OTU within the RSE communities (maximum 72% of total sequences at RSE4, and large percentages of 53% and 57% at RSE3 and RSE5, respectively) exhibited 84% sequence identity to the hyperthermophilic and chemoorganotrophic crenarcheote *Desulfurococcus mucosus* DSM2162 (Wirth et al 2011) and 99% sequence identity with YNPFFA108, an uncultured archaeal clone from YNP thermal soil (accession number AF391993) (Kozubal et al 2008).

The most abundant archaeal 16S rRNA gene transcript OTU in the RSW communities (maximum 78% of total sequences at site RSW2, and found in significant percentages at all four sites of RSW) exhibited 82% sequence identity to the hyperthermophilic heterotrophic crenarchaeon *Thermofilum* sp. 1807-2 (Anderson et al 2008, Toshchakov et al 2015) within the class Thermoprotei (Fig. 3.6). This OTU also exhibited 98% sequence identity to the uncultured archaeal clone AK8 recovered from

Bor Khlueng hot spring in Thailand (accession number AY555814) (Kanokratana et al 2004). Another OTU in these communities (maximum 37% of total sequences in the RSW4 community, and found in RSW3 at 3% of total sequences) exhibited 83% sequence identity to the thaumarchaeote *Candidatus* Nitrososphaera sp. N89-12 within the order Nitrososphaerales (Stieglmeier et al 2014), and also exhibited 99% sequence identity to the uncultured archaeal clone OPA-GC-16SrRNA-145 recovered from a terrestrial hydrothermal ecosystem [accession number KF836708, (Poret-Peterson et al 2013)]. An additional OTU in these communities (maximum 29% of total sequences in the RSW1 community) exhibited 100% sequence identity with *Pyrobaculum* strain WP30, a putative heterotroph recovered from a sulfur-rich Yellowstone hot spring [accession number CP012158; (Jay et al 2015)].

Bacteria. Sediments collected from site RSE5 and all sampling locations at RSW yielded bacterial 16S rRNA gene transcripts of sufficient quantity for sequencing. From the sites that yielded amplicons of sufficient abundance, a total of 38,696 sequences were sampled (Table 3.7/S6). The composition and structure of bacterial communities shifted down the outflow channels of RSW (Fig. 3.6).

The most abundant bacterial 16S rRNA gene transcript OTU in RSW communities (maximum 72.9% of total sequences at site RSW3) exhibited 91% sequence identity to the thermophilic and heterotrophic microaerophile *Thermoflexus hugenholtzii* (Dodsworth et al 2014) within the order Chloroflexi (Fig. 3.6). The next most abundant OTU in these communities (56% total sequences at site RSW4) exhibited 93% sequence

identity to the cyanobacterium *Cyanothece* sp. PCC7425 within the order Chroococcales (Bandyopaghyay et al 2011). Sequences most closely affiliated (100% sequence identity) with the arsenite oxidizing bacterium *Pseudomonas arsenicoxydans* within the order Pseudomonadales (Campos et al 2010) represented 34%, 17%, and 1% of total sequences in RSW1, RSW2, and RSW3, respectively. In the RSE5 community, the most abundant OTU (29% of total sequences) exhibited 99% sequence identity to the facultative anaerobe *Streptococcus pneumonia* 4V4 (Glover et al 2008, Martin et al 2007). Another abundant OTU in RSE5 at 12% of total sequences respectively, exhibited 100% sequence identity to the peptidoglycan-degrading bacteria *Delftia lacustris* strain R-547334 within the order Burkholderiales (Jorgensen et al 2009).

Eukarya. Sites RSE5 and RSW4 yielded eukaryal 18S rRNA gene transcripts of sufficient abundance for sequencing, with a total of 8,554 eukaryal sequences obtained from these sites (Fig. 3.6, Table 3.7/S6). The most abundant eukaryal 18S rRNA transcript OTU in both RSE5 and RSW4 communities (77% and 35% of total sequences, respectively) exhibited 100% sequence identity to the photoautotrophic red alga *Cyanidioschyzon merolae* within the order Cyanidiales (Matsuzaki et al 2004). The next most abundant OTU in RSW4 (16% of total sequences) exhibited 100% sequence identity to the freshwater diatom *Achnanthidium minutissimum* (Larras et al 2014). A small percentage of total sequences (1.2%, 0.6%, 0.5%, and 0.4%) in sites RSE5 and RSW4 exhibited 98%, 89%, 100% and 98% identity to a single-celled monad (order Chromulinales), an amoeba (order Entamoebidae), yeast (order Sporidiobolales), and lichen (order Teloschistales), respectively (Fig. 3.6).

Dilution to extinction cultivation assays.

The number of viable autotrophic hydrogenotrophic cells capable of reducing various oxidants was estimated in sediments sampled along the outflow channels of RSE and RSW using serial dilution to extinction cultivation assays. The maximum number of the hydrogenotrophic cells, as indicated by the highest dilution exhibiting activity for any given assay, increased down the outflow channel of RSE. In contrast, the maximum numbers of viable hydrogenotrophic cells was similar in samples collected from RSW1 and RSW3. Collectively, site RSE5 was characterized by the highest number of viable hydrogenotrophic cells (2.3×10^7 total cells) (Table 3.1).

The oxidants used to support autotrophic H_2 oxidation activity in communities from RSE and RSW varied between springs, as well as along the outflow channels of each spring. All sediment-associated communities sampled from RSE included autotrophic hydrogenotrophs capable of using Fe^{3+} and O_2 as electron acceptors. Sequencing of the most dilute enrichments indicates that the dominant OTUs changed down the outflow channel of RSE, and were most closely related to members of the genus *Metallosphaera* in RSE1 (O_2 and Fe^{3+} as oxidants), the genera *Nitrososphaera* and *Desulfurococcus* in RSE1 and RSE2 (O_2 and Fe^{3+} as oxidants), *Bacillus* in RSE3 (Fe^{3+} as oxidant), *Methylophilaceae* and *Sulfobacillus* in RSE4 (O_2 as oxidant), *Paracoccus* in RSE4 (Fe^{3+} as oxidant), and finally the genus *Methylobacterium* in RSE5 (Fe^{3+} as oxidant) based on 16S rRNA gene sequencing (Table 3.8/S7). Sediment-associated communities sampled from RSW1 included hydrogenotrophs most closely related to the

genera *Thermofilum* and *Fervidicoccus* (O_2 as oxidant). Unlike RSE, hydrogenotrophs capable of using Fe^{3+} were not detected in RSW (Table 3.1; Table 3.9/S8).

Hydrogenotrophs most closely related to *Methylobacterium* in RSW1 and *Sulfurihydrogenibium* and Ca. *Korarchaeum* in RSW3 were apparently supported by NO_3^- as the oxidant (Table 3.9/S8); hydrogenotrophs capable of coupling H_2 oxidation with NO_3^- were not detected in RSE (Table 3.8/S7). Hydrogenotrophic cells capable of reducing SO_4^{2-} were only detected in one site in RSE (RSE1), however, PCR amplifiable 16S rRNA genes could not be obtained from this culture. Cells capable of reducing S° were only detected in RSE1 and RSE2 (most closely related to *Desulfurococcus*) (Table 3.8/S7). In contrast, hydrogenotrophs capable of reducing S° and SO_4^{2-} were detected at both sampling locations in RSW (Table 3.1) and dominant OTUs in these enrichments were most closely related to *Thermodesulfobacterium* in RSW1 and *Caldiserica* and *Ca*. Caldiarchaeum in RSW3 (Table 3.9/S8).

Discussion

In the present study, we used geochemical proxies (ratios of SO_4^{2-}/Cl^- and $\delta D/\delta^{18}O$ in hot spring waters) to characterize the influence of subsurface and surface processes, specifically phase separation, on the geochemistry of a paired set of hot springs and how these geochemical differences, in turn, influence the activity and composition of chemoautotrophic hydrogenotrophs that inhabit these springs. Based on a previous classification scheme [(Nordstrom et al 2009); Fig. 3.7], we interpret that the acidity of RSE (source pH = 3.0) is due to the interaction of vapor-phase-dominated

fluids in the near-surface with O_2-rich precipitation-derived meteoric fluids, forming acid-sulfate-chloride waters (vapor-phase input). In contrast, RSW is circumneutral (source pH = 7.0) due to sourcing primarily from deep hydrothermal fluid (liquid-phase input). The presence of NO_3^- in RSE and RSW source fluids is interpretted as the result of mixing of vapor phase or deep hydrothermal fluid with near surface groundwaters, respectively, such as has been suggested as the source of NO_3^- in other geothermal fields areas such as Hot Springs, Arkansas (Yeatts 2006). Thus, both pools are likely influenced by mixing of hydrothermal fluids with near surface groundwater, but differ in the extent of vapor phase and liquid phase fluid input to each spring. The difference in the source of fluids supplied to these neighboring springs separated by ~20 meters leads to strong differences in spring geochemistry in their source pools and along their outflow channels.

Subsurface boiling and separation of hydrothermal fluids into a gas and a steam phase could occur as a single stage (steam remains mixed with liquid and separates near a single temperature) or could be continuous (steam is physically separated from liquid as it is formed) (Truesdell et al 1977). In Yellowstone, interpretations indicate that this subsurface process is multistage or intermediate between single stage and continuous models (Truesdell et al 1977), suggesting that phase separations are not typically complete. Incomplete separation of phases in fluids in the subsurface of RSE and RSW could explain the detection of H_2 in RSW, albeit at 50% of the concentration in RSE, which should otherwise partition with the vapor phase in a continuous steam separation. Alternatively, H_2 could be produced biologically in the near subsurface of RSW through fermentative biologically-mediated reactions. Like uncharged H_2, boiling and separation

of fluids into vapor and liquid phases would be expected to concentrate uncharged CH_4 in the vapor phase, which should lead to higher abundances of CH_4 in RSE when compared to RSW. However, dissolved CH_4 is 600% greater in concentration at RSW than RSE. The source of the CH_4 in RSW is not completely understood, however, it is possible that the CH_4 is biogenic in origin and could be produced in the shallow subsurface of RSW. Additional data (e.g., the isotopic composition of CH_4) are needed to further assess the source of CH_4 in RSW. Nonetheless, observations like these reveal that while subsurface phase separation is a primary driver of geochemical differences in hot springs of YNP (Amenabar et al 2015, Nordstrom et al 2009, Truesdell et al 1977), nuances exist in the chemistry of individual springs that oftentimes require additional data (i.e., isotopes of CH_4) to deconvolute.

The presence of H_2 in both RSE and RSW indicates that these proximal springs provide an opportunity to evaluate the influence of geochemical variation, specifically differences in oxidant availability, on the ecology of chemoautotrophic organisms involved in H_2 metabolism. However, prior to conducting extensive experimentation on chemoautotrophic H_2 cycling within the selected microbial communities, it was necessary to confirm which of the sampling sites in RSE and RSW were chemoautotrophic since all of the sampling locations in RSW (RSW1-4) as well as RSE5 lie within pH, temperature, and sulfide space that allows for photoautotrophic life (Boyd et al 2010, Boyd et al 2012, Cox et al 2011) (Fig. 3.10/S3). The exclusion of light from microcosms did not decrease rates of DIC assimilation in sediment communities sampled from RSW1-3, RSE1, or RSE3 when compared to those incubated in the light, indicating that DIC assimilation in

these communities is not dependent on light (i.e., phototrophic). This is consistent with a lack of detectable 16S/18S rRNA gene transcripts affiliated with known phototrophic lineages in sediments sampled from these sites (Fig. 3.6). RSE2 and RSE4 also lacked 16S/18S rRNA gene transcripts affiliated with known phototrophs. However, DIC assimilation assays were not conducted at these sites, which precludes a definitive assessment of the potential for photoautotrophy in these environments. Nevertheless, it is highly likely that photoautotrophic metabolism is absent in these two spring locations given that they are at a temperature that is close to or above those known to be tolerated by acidophilic phototrophs in hot springs [i.e., ~52°C (Boyd et al 2010, Boyd et al 2012, Cox et al 2011)].

Despite RSW1-3 sampling locations harboring pH, temperature, and sulfide conditions that should allow for photosynthesis (Boyd et al 2010, Boyd et al 2012, Cox et al 2011) (Fig. 3.10/S3), DIC assimilation assays and affiliations of 16S rRNA gene transcripts suggest they are not photoautotrophic. A previous analysis of 442 hot spring samples from Yellowstone suggest that pH, temperature, and sulfide concentration differences among these locations can only explain ~67% of the variation in the distribution of photosynthesis in these environments (Boyd et al 2012). This indicates that other unaccounted factors likely constrains phototrophic DIC assimilation in these RSW spring locations. In contrast to RSE1 and RSE3 and RSW1-3, sediments sampled from RSW4 and RSE5 exhibited rates of DIC assimilation in microcosms incubated in the light that were two to three orders of magnitude greater than those incubated in the dark, indicating that photoautotrophic DIC assimilation was occuring at these sites (Fig.

3.3). RSW4 and RSE5 microbial communities also exhibited large increases in 18S rRNA gene transcript numbers (Fig. 3.5), which is likely due to the much higher levels of productivity associated with phototrophic ecosystems (*i.e.*, Fig. 3.3) and the greater amount of biomass that this fixed carbon can support. Since chemoautotrophic H_2 metabolism is the emphasis of this study, subsequent discussion focuses on the non-phototrophic sites RSW1-3 and RSE1-4.

Dissolved H_2 was detected at all sampling locations and activity assays indicate that its oxidation is a potential source of chemical energy for all of the chemosynthetic communities in RSE and RSW examined with the possible exception of RSE1 (discussed below). Rates of net H_2 oxidation (RSE2-4, RSW1 and RSW3) ranged from 1.69 – 4.34 nmoles H_2 min^{-1} gdm^{-1} at RSE and from 5.34 – 24.85 nmoles H_2 min^{-1} gdm^{-1} at RSW (Fig. 3.4). These rates are slightly higher than the rates of H_2 oxidation (1.5 nmoles H_2 min^{-1} gdm^{-1}) measured previously in 'Dragon Spring' (D'Imperio et al 2008), an acidic (pH ~3) and high temperature (~70°C) sulfur rich spring in the Norris Geyser Basin of YNP. The variation in the rates of net H_2 oxidation measured in these studies may be attributable to methodological differences, in particular amendment with H_2 to a concentration of 100 ppm in the present study. This contrasts with the study at 'Dragon Spring' where rates were measured in unamended microcosms supplied with source water. The concentration of H_2 in source waters associated with the sediments of 'Dragon Spring' was ~ 13 nM (D'Imperio et al 2008), which is substantially lower than the 200 to 300 nM concentrations determined for RSE and RSW source waters. Since net rates of H_2 oxidation were determined in closed systems in both studies, it is possible that rates of

H_2 oxidation measured previously in 'Dragon Spring' were low due to depletion of the low amount of native H_2 early during the microcosm incubation.

The majority of hydrogenotrophic and autotrophic cells inhabiting RSE and supported by our cultivation medium reduced O_2 or Fe^{3+}, and the number of these cells increased down the outflow channel. The increase in the abundance of hydrogenotrophs capable of utilizing O_2 is likely due to the increased influx of atmospheric O_2 due to decreases in the temperature of spring water, and increasing contact with the atmosphere along the outflow channel of RSE. The decreased temperature would allow for increased infusion of O_2 (Shock et al 2010) that, in turn, may support a higher number of O_2-dependent hydrogenotrophic cells. The increase in the number of Fe^{3+}-dependent hydrogenotrophic cells down the outflow channel may also be due to increased input of O_2 along the outflow channel of RSE as the spring cools, thereby allowing for increased O_2-dependent biological oxidation of Fe^{2+} and more abundant Fe^{3+} to support Fe^{3+} dependent microorganisms. While mineralogical evidence for oxidized iron phases was not observed in RSE sediments via XRD analysis (Table 3.5/S4), iron (hydr)oxides were visually apparent along the outflow channel between sites 2 and 4 (Fig. 3.1). Moreover, previous studies have measured soluble Fe^{3+} in this same spring (Wilson et al 2000). These observations are consistent with the presence of 16S rRNA gene transcripts along the outflow of RSE that are closely affiliated with the aerobic Fe^{2+} oxidizing crenarchaeote *Metallosphaera yellowstonensis* strain MK1 (Sulfolobales) (Kozubal et al 2011). Additionally, the amount of Fe^{2+} decreased by 28.0% between RSE1 and RSE5. At pH 3.0, this decrease can be atrributed to biological oxidation with O_2 given the slow

abiotic Fe^{2+} oxidation kinetics in waters with pH less than 4.0 (Edwards et al., 2000; Edwards et al., 1998; Johnson and Hallberg, 2005).

Dilution to extinction cultivation assays indicate that site RSW1 was dominated by hydrogenotrophic and autotrophic cells capable of reducing S° (4.6×10^5 cells gdm^{-1}). S° was not detected via XRD analysis of these sediments (Table 3.5/S4) although visible S° flocs were observed as a thin coating on the sediments of RSW1, RSW2, and RSW3. These flocs were likely formed by incomplete sulfide oxidation, which is possibly consistent with the decrease in total sulfide concentration observed between measurements made at RSW1 (spring source) and RSW4 by 38.6%, although it is also likely that degassing plays a role in the decrease of sulfide (Fig. 3.2D; Table 3.2/S1). Based on available models for the development of hot springs, the source of the sulfide in RSW is not clear since this spring is unlikely to be influenced to a significant extent by vapor phase input, as delineated by SO_4^{2-}/Cl^- ratios (Fig. 3.8/S1). One possible explanation is that sulfide is not fully volatilized during repeated episodes of near subsurface boiling, in particular the fraction that is dissolved as bisulfide [pKa for H_2S/HS^- is 6.4 at 100°C (Amend and Shock 2001)]. This interpretation is consistent with the multi-stage steam separation model previously invoked for the majority of Yellowstone hot springs (Truesdell et al 1977). Alternatively, it is possible that the sulfide in RSW source waters is due to microbial SO_4^{2-} reduction in the subsurface or near subsurface of RSW. Evidence in potential support of this possibility comes from our dilution to extinction cultivation assays of H_2-dependent autotrophs from the source of RSW supplied with SO_4^{2-}, which indicated the presence of ~ 10^3-10^4 cells gdm^{-1} (most

closely related to *Thermodesulfobacterium*). Hydrogenotrophic *Thermodesulfobacterium* strains have been shown to respire SO_4^{2-} (Jeanthon et al 2002) and previous studies have documented high rates of sulfate reduction (Fishbain et al 2003) in a spring in YNP with similar pH (6.6), temperature (69°C), and SO_4^{2-} concentration (0.6 mM) to that measured in source waters of RSW. Further work is needed to determine the sources of sulfide in springs like RSW and other circumneutral to alkaline springs in YNP [*e.g.*, Boulder Spring (Cox et al 2011)].

Based on sequencing of transcripts of 16S rRNA genes from sediments and physiological inferences, populations putatively involved in H_2 oxidation in RSW1 include OTUs related to *Thermoflexus hugenholtzii* and *Thermofilum* sp. 1807-2 (members of the orders Thermoflexales and Thermoproteales, respectively) (Dodsworth et al 2014, Huber et al 2006), although neither of these strains was apparently enriched in our cultivation assays. *Thermoflexus hugenholtzii* is a facultative anaerobe and its genome encodes for at least one [NiFe]-hydrogenase, which is classified as a group 3c enzyme (Greening et al 2016). Group 3c enzymes are involved in H_2 oxidation and are typically found in methanogens (Kaster et al 2011) and some SO_4^{2-} reducers (Greening et al 2016). It is unclear whether *T. hugenholtzii* can reduce SO_4^{2-} since its characterization did not include an examination of SO_4^{2-} as a potential electron acceptor. That said, genes for the SO_4^{2-} dissimilatory reduction pathway were not detected in the genome of *T. hugenholtzii* (Dodsworth et al 2014). Nonetheless, it is still possible that the organism affiliated with *T. hugenholtzii* can respire SO_4^{2-} given that the 16S rRNA gene transcripts obtained from RSW were only 91% homologous to *T. hugenholtzii*. *Thermofilum* spp.

tend to be facultative anaerobes and the genome of strain 1807-2, which is the genome of the most closely related *Thermofilum* strain to the sequences obtained from RSW, encodes for a [NiFe]-hydrogenase classified as a group 1g enzyme (Greening et al 2016). Group 1g enzymes are proposed to function in H_2 oxidation when S° serves as a terminal electron acceptor (Greening et al 2016, Laska et al 2003, Pihl et al 1989). These findings are consistent with the high number of S°-dependent autotrophic cells capable of oxidizing H_2 in RSW1 (Table 3.1).

Net H_2 production (biological controls minus killed controls) was measured in microcosms containing sediments from RSE1, although it is possible that some H_2 oxidation may be occuring at RSE1 given the detection of cells capable of H_2 oxidation under autotrophic conditions (Table 3.1). Since these rates were determined using a closed system microcosm assay, it is possible that oxidants capable of driving H_2 oxidation (e.g., O_2 or Fe^{3+}) became limiting during the 1 hr. incubation period (i.e., a bottle effect), thereby shifting the local conditions of the experiment toward one that favors anaerobic or fermentative metabolism. The dominant archaeal 16S rRNA gene transcript in RSE1 was affiliated with the facultative anaerobe *Metallosphaera* sp. MK1 and this strain has been shown to catalyze the O_2-dependent oxidation of Fe^{2+} (Kozubal et al 2012). A net decrease in the concentration of Fe^{2+} (4 μM difference) was observed between sites RSE1 and RSE2 and this is attributable to biology since Fe^{2+} reacts extremely slowly in aqueous solutions with pH less than 4.0 (Edwards et al 1999, Edwards et al 2000, Johnson and Hallberg 2005). Fe^{2+} oxidation would consume a fraction of the already limited amount of O_2 that would be dissolved in the 82°C source

fluids, in a manner similar to what has been shown for other high-temperature Fe^{2+}-consuming *Metallosphaera* dominated hot spring communities (Bernstein et al 2013).

Consumption of an already limited supply of O_2 via biological Fe^{2+} oxidation may select for cells capable of anaerobic respiration or fermentation during the assembly of the RSE1 community. Indeed, net production of H_2 was observed in RSE1 implying the presence of anaerobic, fermentative cells. Consistent with this observation, an abundant 16S rRNA gene transcript OTU that is distantly (83%) related to *Desulfurococcus mucosus* (Desulfurococcales) and an abundant OTU distantly affiliated (82%) with *Thermofilum* sp. 1807-2 (Thermoproteales) were identified in RSE1 sediments. Members of both of these orders are capable of fermentative metabolism (Huber and Stetter 2015a, Huber and Stetter 2015b) and members of the Desulfurococcales order are known to fermentatively produce H_2 (Anderson et al 2009). Further evidence that the RSE1 community has very few cells capable of autotrophic growth with H_2 comes from the dilution to extinction assays where the number of cells capable of oxidizing H_2 was 3 to 4 orders of magnitude lower than in other sample locations in RSE.

Intriguingly, the most abundant hydrogenotroph identified in RSE1, either determined by serial dilution enrichment assays amended with Fe^{3+} or O_2 (Table 3.8/S7) or as inferred from 16S rRNA gene transcript sequences (Fig. 3.6), was affiliated with *Metallosphaera* sp. MK1. As outlined above and in previous studies (Kozubal et al 2008, Kozubal et al 2011), this strain is generally associated with aerobic Fe^{2+} oxidation and formation of Fe(III) oxide solid phases in acidic hot spring environments. However, the

genome of MK1 encodes for three [NiFe]-hydrogenases that are putatively involved in H_2 oxidation [*i.e.*, groups 1h, 2d, and 2e (Greening et al 2016)]. Group 1h and 2e enzymes are proposed to function in the oxidation of H_2 during aerobic respiration (Schafer et al 2013, Søndergaard et al 2016) whereas group 2d enzymes have been suggested to provide additional reductant for carbon fixation (Brugna-Guiral et al 2003). The previous characterization of strain MK1 apparently did not examine whether H_2 could support or augment autotrophic growth with O_2 or Fe^{3+} as oxidants (Kozubal et al 2008, Kozubal et al 2011). Nonetheless, evidence outlined above strongly suggests that this strain can metabolize H_2 and possibly couple this with reduction of Fe^{3+}. Indeed, serial transfer of the H_2 oxidizing, Fe^{3+} reducing autotrophic culture from RSE1 (sequencing indicates it to be *Metallosphaera* sp. MK1) into autotrophic growth medium with Fe^{2+} as electron donor and O_2 as electron acceptor resulted in growth (data not shown). This suggests that this strain has adapted to use H_2 or Fe^{2+} as electron donors and O_2 or Fe^{3+} as electron acceptors; choice of redox couples is likely dependent on the availability of O_2 considering that its use releases the most amount of energy regardless of electron donor (Shock et al 2010). That a strain such as MK1 can both oxidize and reduce Fe would not be unprecedented among members of the Sulfolobales, since these activities have been previously demonstrated in a strain (MK5) isolated from acidic hot springs in YNP (Kozubal et al 2012).

Conclusions

As summarized in Fig. 3.7, variations in the availability of oxidants influence the abundance and composition of hydrogenotrophic populations. This was found to be particularly true for hydrogenotrophs capable of reducing Fe^{3+}, which results from near surface O_2 dependent oxidation of Fe^{2+} that is enriched in springs that are supplied by vapor phase input relative to those that are not supplied by this input (Nordstrom et al 2009). This observation, when combined with data indicating that the predominant hydrogenotrophs in RSE and RSW differ at the domain level taxonomic rank, is consistent with phylogenetic data that indicates a prominent role for oxidant availability in the diversification of hydrogentrophs at the level of the [NiFe]-hydrogenase enzymes that catalyze H_2 oxidation (Boyd et al 2014, Greening et al 2016). Since surface processes (e.g., atmospheric gas infusion of O_2) are likely to impact springs similarly if spring temperatures are the same, the geochemical composition of hot spring fluids and the precipitated minerals that form in outflow channels of hot springs are ultimately the result of differences in subsurface processes, namely the subsurface phase separation of fluids and interactions of these fluids with bedrock, and mixing of those fluids with meteoric waters. Importantly, oxidants such as Fe(III) and $S°/SO_4^{2-}$ would not be present in these springs and would therefore be unavailable to biology without O_2 to oxidize subsurface sourced Fe(II) and sulfide, respectively (Colman et al 2017, Kozubal et al 2012). We therefore conclude that this study demonstrates the importance of subsurface geological processes and their interaction with surface processes in shaping the ecology, physiology, and evolution of hydrogenotrophic organisms in hydrothermal ecosystems by controlling complex hot spring geochemistry and oxidant availability.

In addition to variation in oxidants used to support H_2 metabolism, the RSE microbial community is dominated by Archaea whereas the RSW community is dominated by Bacteria. The combination of acidic pH and elevated temperature in RSE, which tends to select against Bacteria (Colman et al 2017, Valentine 2007), is likely to explain this marked difference in community composition. The predominant processes that control the geochemistry and therefore the pH of hot springs take place in the subsurface, indicating that such processes had and continue to have a profound influence on the distribution and, by extension, the diversification of microbial life at the highest taxonomic ranks.

Acknowledgements

The authors declare that they have no conflicts of interest. This work was supported by a NASA Exobiology and Evolutionary Biology grant (NNX13AI11G) to T.M.H. and E.S.B., a NASA Early Career Fellowship program grant (NNX17AH10G) to E.S.B., and NSF grant EAR-1529963 to E.L.S. The NASA Astrobiology Institute is supported by grant NNA15BB02A (to T.M.H., E.L.S. and E.S.B.). M.R.L. acknowledges support from the NASA Earth and Space Science Fellowship Program (NNX16AP51H). We are grateful to Natasha Zolotova, Kirt Robinson, and Grayson Boyer for their help in DIC determinations. We thank Christie Hendrix and Stacey Gunther at YNP for research permitting.

Table

Table 3.1. Approximate number of autotrophic cells (per gram dry mass sediment) capable of coupling H_2 oxidation to the reduction of specified oxidants (electron acceptor) as measured by dilution series cultivation experiments.

Electron Acceptor	**RSE1**	**RSE2**	**RSE3**	**RSE4**	**RSE5**	**RSW1**	**RSW3**
O_2	2.6×10^1	1.5×10^6	4.6×10^5	2.5×10^6	2.3×10^7	4.6×10^3	4.4×10^3
NO_3^-	ND	ND	ND	ND	ND	4.6×10^2	4.4×10^5
Fe^{3+}	2.6×10^3	1.5×10^4	4.6×10^6	2.5×10^4	2.3×10^4	ND	ND
SO_4^{2-}	2.6×10^2	ND	ND	ND	ND	4.6×10^3	4.4×10^4
S°	2.6×10^1	1.5×10^2	ND	ND	ND	4.6×10^5	4.4×10^2
Max.	2.6×10^3	1.5×10^6	4.6×10^6	2.5×10^6	2.3×10^7	4.6×10^5	4.4×10^5

Abbreviations: ND, not detected. Max., maximum number of H_2 oxidizing cells.

Figures

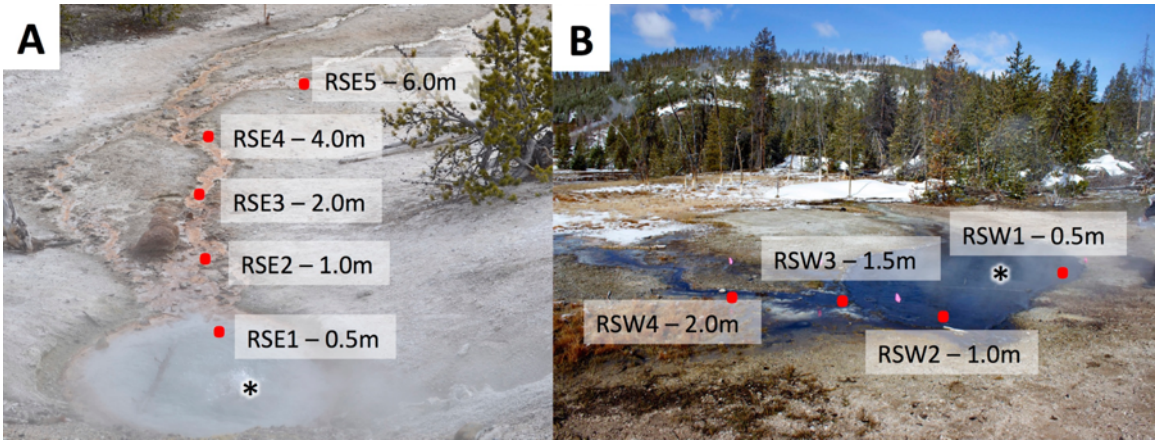


Fig. 3.1. Location of sampling sites (•) along the outflow channels at Roadside East (RSE) (A) and Roadside West (RSW) hot springs (B). The estimated distance of each sampling site from the source of each spring (*) is also indicated.

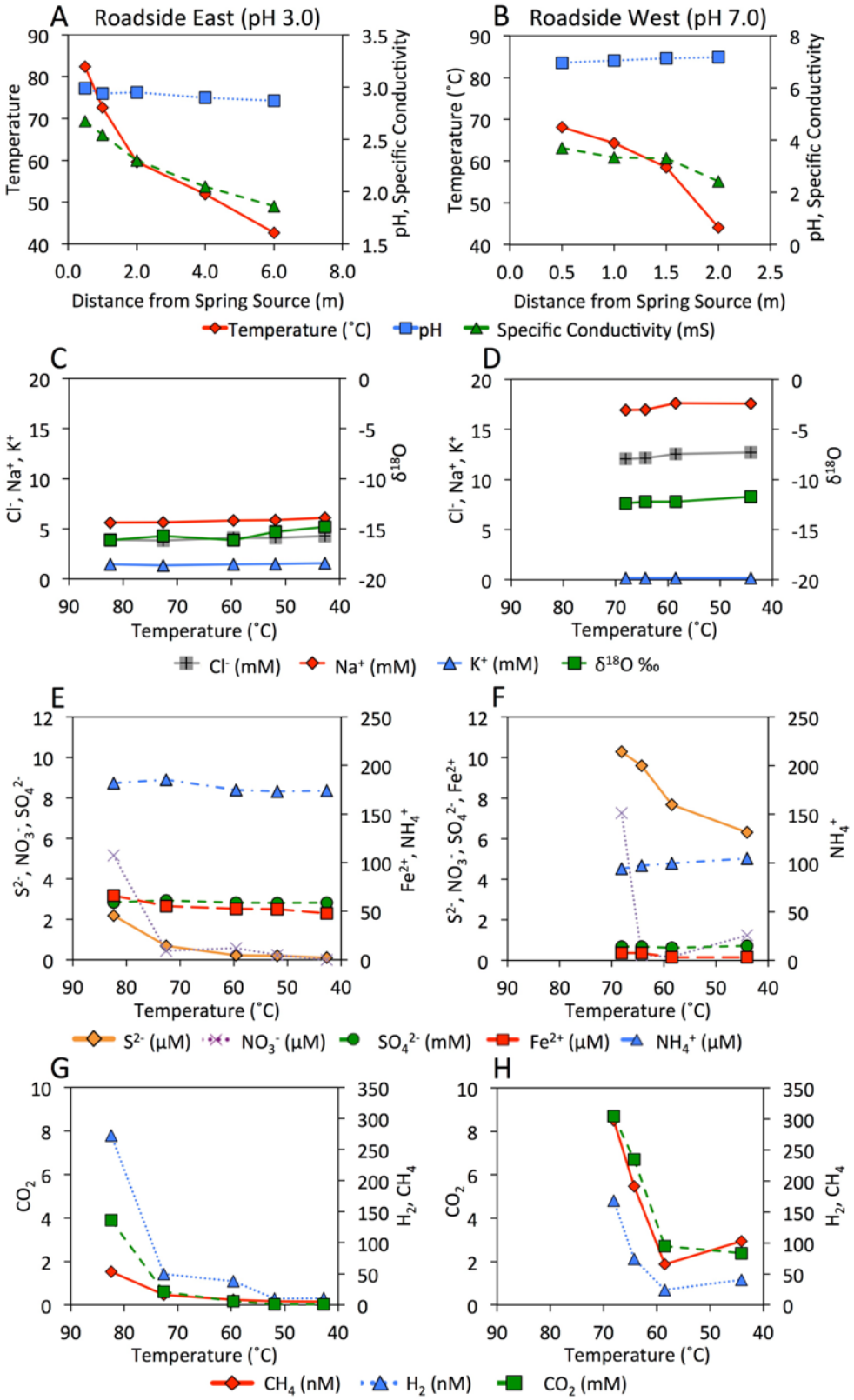

A
Roadside East (pH 3.0)
Temperature
pH, Specific Conductivity
Distance from Spring Source (m)
B
Roadside West (pH 7.0)
Temperature (˚C)
pH, Specific Conductivity
Distance from Spring Source (m)
Temperature (˚C)
pH
Specific Conductivity (mS)
C
D
Cl⁻, Na⁺, K⁺
δ18O
Temperature (˚C)
Cl⁻ (mM)
Na⁺ (mM)
K⁺ (mM)
δ18O ‰
E
F
S2-, NO3-, SO4 2-
Fe2+, NH4+
Temperature (˚C)
S2- (µM)
NO3- (µM)
SO4 2- (mM)
Fe2+ (µM)
NH4+ (µM)
G
H
CO2
H2, CH4
Temperature (˚C)
CH4 (nM)
H2 (nM)
CO2 (mM)

Fig. 3.2. Water temperature, pH, and specific conductance at sampling locations along the outflow channels of RSE (A) and RSW (B). Concentrations of chloride (Cl^-), sodium (Na^+), potassium (K^+), and the isotopic composition of oxygen in water ($\delta^{18}O$) as measured along the outflow channels in RSE (C) and RSW (D). The increase in the concentration of Cl^- was used to normalize the concentrations of the following constituents to account for evaporative concentration: total sulfide (S^{2-}), ferrous iron (Fe^{2+}), nitrate (NO_3^-), sulfate (SO_4^{2-}), ammonium (NH_4^+), and chloride (Cl^-) as measured along the outflow channels in RSE (E) and RSW (F). Concentrations of dissolved methane (CH_4), carbon dioxide (CO_2), and hydrogen (H_2) as measured along the outflow channels in RSE (G) and RSW (H).

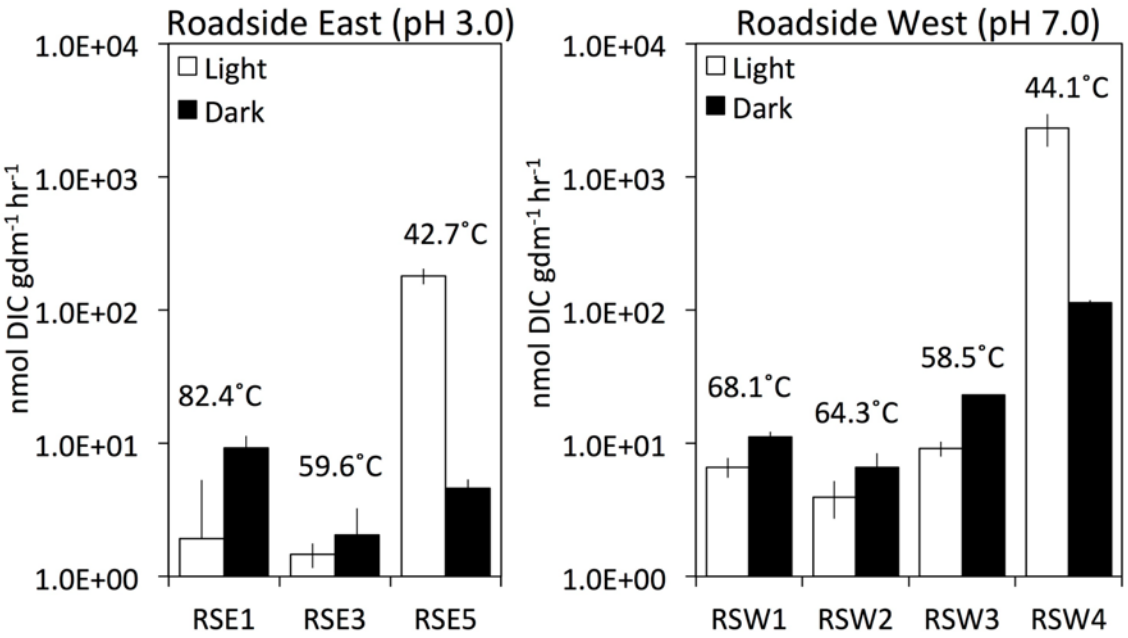


Fig. 3.3. Dissolved inorganic carbon (DIC) uptake in sediment-associated microbial communities sampled along the outflow channels of select RSE and RSW sampling locations. The temperature of each site is indicated.

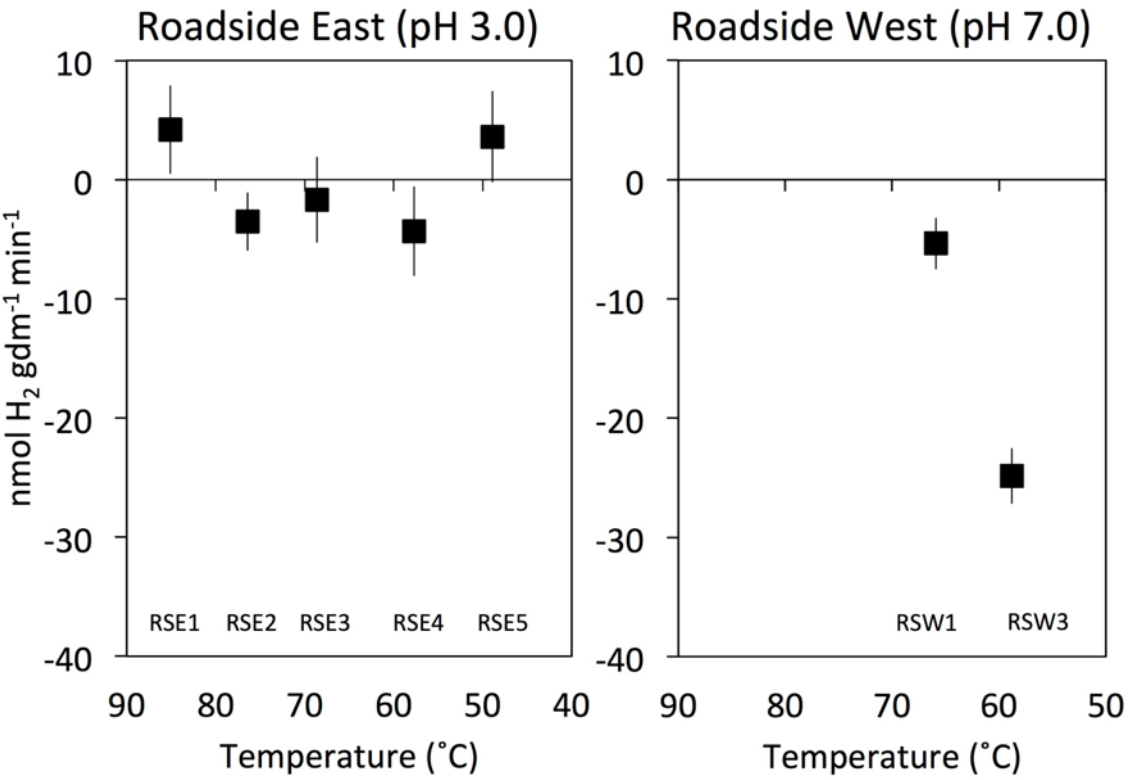


Fig. 3.4. Rates of net H_2 transformation in sediment-associated microbial communities in selected sites along the RSE and RSW hot spring outflow channels. Positive numbers indicate net rates of H_2 production whereas negative numbers indicate net rates of H_2 consumption. The values of H_2 transformation represent the differences between triplicate biotic and abiotic controls.

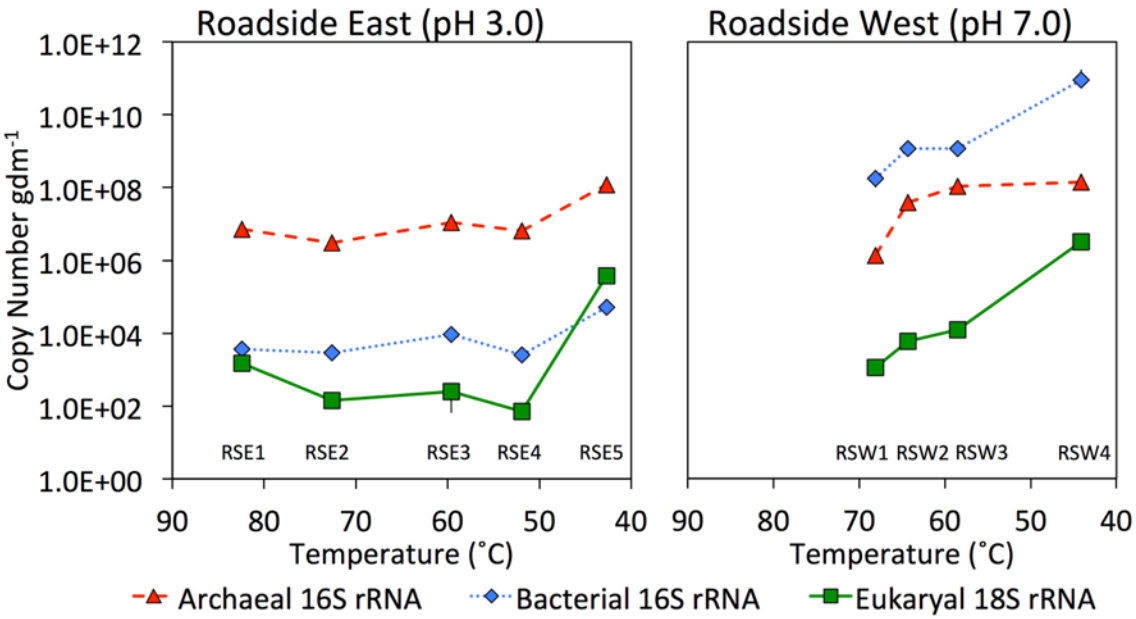


Fig. 3.5. Abundances of 16S (Archaea, Bacteria) and 18S (Eukarya) rRNA gene transcripts associated with sediments sampled along the outflow channels of RSE and RSW hot springs as determined using qPCR on cDNA. Results are presented as the mean of triplicate qPCR assays; error bars represent the standard error of triplicate assays. Many error bars fall within the width of the symbols.

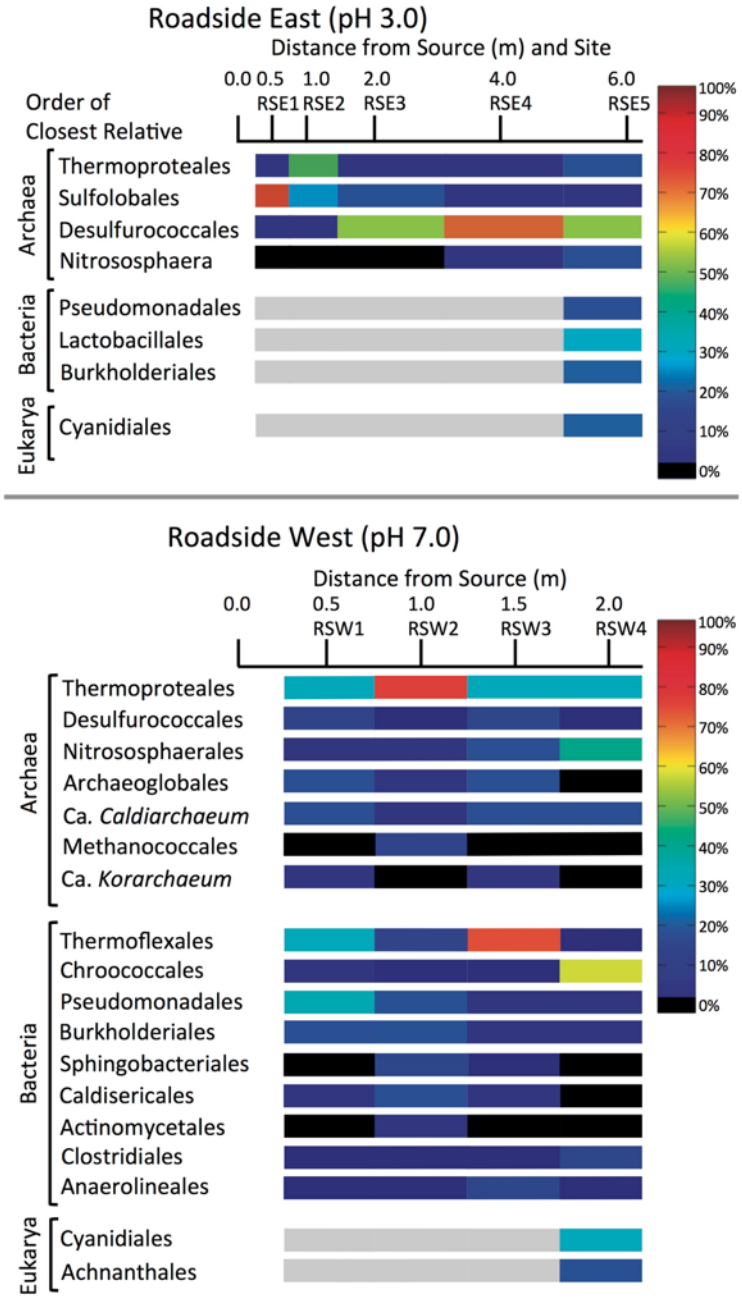


Fig. 3.6. Relative abundances of 16S (Archaea and Bacteria) and 18S (Eukarya) rRNA gene transcript OTUs (as represented by their closest cultivated relatives) as a function of distance along the outflow channels of RSE and RSW. OTUs are color coded according to the relative abundance of sequences within a taxonomic domain as denoted by the bar chart to the right of each panel. Representative OTUs were binned at the order level with orders that represent >5% of total sequences in a sampling location included. Grey bars represent no amplification and thus no sequences for that domain at a given sampling location. Relative abundances (%) for OTUs in each order within a taxonomic domain are also reported in Table 3.10/S9.

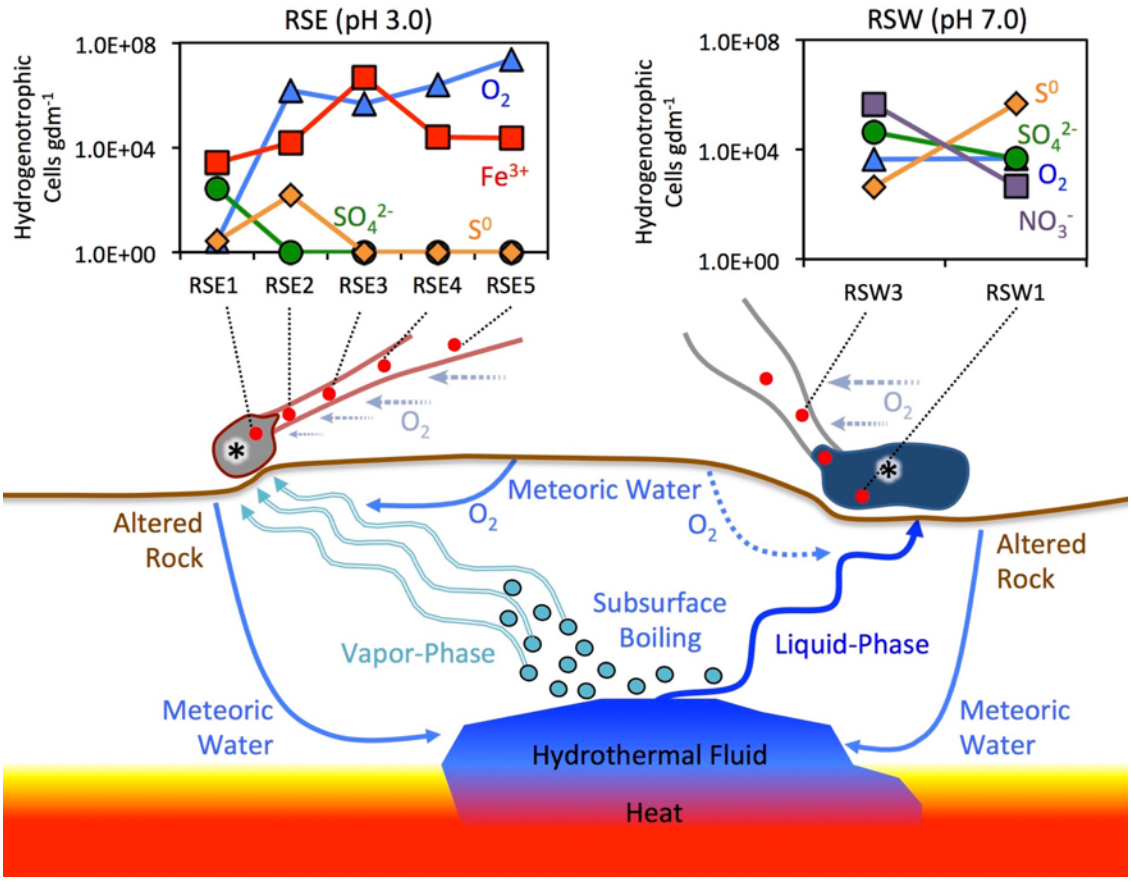


Fig. 3.7. Schematic summarizing the influence of subsurface and surface processes on oxidant availability and use by hydrogenotrophic and autotrophic populations (graphs above cross section). Subsurface hydrothermal fluids recharged by meteoric water can boil during their ascent to the surface and can separate into a vapor phase and a liquid phase. In the case of the paired springs studied here, the vapor phase is thought to interact with meteoric, precipitation-derived fluid forming acidic springs such as RSE (pH 3.0). The liquid phase may also interact, albeit to a lesser extent, with meteoric water and may have slight vapor phase input due to incomplete phase separation resulting in circumneutral springs such as RSW spring (pH 7.0). The acidic pH of RSE is likely the result of microbially catalyzed near-surface oxidation of sulfur that is enriched in the vapor phase by O_2 delivered by meteoric fluid (Colman et al., 2017). The acidic fluid can promote the release of metals from altered host rock, such as Fe^{2+} in RSE. Biological oxidation of this Fe^{2+} can form Fe^{3+} (Kozubal et al., 2012) that can support hydrogenotrophic activity. SO_4^{2-} in RSW is likely due to disproportionation of volcanic sulfur dioxide to HS^-/H_2S and SO_4^{2-} deep in the Yellowstone hydrothermal system prior to phase separation (Nordstrom et al., 2009). Concentrations of SO_4^{2-} in this deep reservoir are estimated to be 0.72 to 0.94 mM, similar to those measured in RSW (range from 0.67 to 0.72 mM). The source of NO_3^- in both springs is hypothesized to be from mixing with near surface groundwater, as has been observed in other geothermal fields (Yeatts, 2006). Increased infusion of atmospheric O_2 as fluids cool along the outflow channels of both springs can support hydrogenotrophic activity directly and indirectly

through reactions involving production of S°, NO_3^-, and Fe^{3+}. Note that sites in the RSE schematic progress from RSE1 to RSE5 to the right in the RSE schematic, while sites in the RSW schematic progress from RSW1 to RSW3 to the left.

Supplemental Tables

Table 3.2/S1. Geochemical parameters measured in waters from sampling locations in April 2014. The conservative, biologically-inert tracers Cl^-, Na^+, and K^+ are presented here, along with water isotope values. The concentration of Cl^- was used to normalize geochemical data presented in Table 3.3/S2 to account for evaporative concentration of chemical constituents in waters along the outflow channels of RSE and RSW. Data correspond with those presented in Fig. 3.2C and Fig. 3.2D.

Site	**Temp. (°C)**	**pH**	**Cl^- (µM)**	**Na^+ (µM)**	**K^+ (µM)**	**δD**	**$\delta^{18}O$**
RSE1	82.4	3.0	3.92	5.62	1.44	-140.4	-16.1
RSE2	72.6	2.9	3.82	5.63	1.34	-139.5	-15.7
RSE3	59.6	3.0	4.09	5.83	1.45	-139.0	-16.1
RSE4	51.9	2.9	4.11	5.87	1.49	-138.1	-15.3
RSE5	42.7	2.9	4.27	6.08	1.56	-138.1	-14.8
RSW1	68.1	7.0	12.05	16.92	0.17	-131.8	-12.4
RSW2	64.3	7.0	12.14	16.96	0.18	-131.8	-12.2
RSW3	58.5	7.1	12.57	17.60	0.16	-131.3	-12.2
RSW4	44.1	7.2	12.69	17.59	0.18	-130.2	-11.7

Table 3.3/S2. Geochemical parameters measured in waters from sampling locations in April 2014, with S^{2-}(T), NO_3^-, Fe^{2+}, and NH_4^+ in sites RSE2, RSE3, RSE4, RSE5, RSW2, RSW3, and RSW4 all normalized to exclude any changes in concentrations due to evaporative concentration (normalized using changes in the conservative, biologically inert tracer Cl^-, values described in Table 3.2/S1). Data correspond with those presented in Fig. 3.2, Table 3.2/S1, and Table 3.4/S3.

Site	Temp. (°C)	pH	S^{2-}(T) (µM)	NO_3^- (µM)	SO_4^{2-} (mM)	Fe^{2+} (µM)	NH_4^+ (µM)
RSE1	82.4	3.0	2.2	5.2	2.83	66.3	181.8
RSE2	72.6	2.9	0.7	0.5	2.93	55.4	185.3
RSE3	59.6	3.0	0.2	0.6	2.81	52.6	175.0
RSE4	51.9	2.9	0.2	0.2	2.81	52.3	173.4
RSE5	42.7	2.9	0.1	0.0	2.81	47.7	174.2
RSW1	68.1	7.0	10.3	7.3	0.67	0.4	94.2
RSW2	64.3	7.0	9.6	0.3	0.68	0.4	97.4
RSW3	58.5	7.1	7.7	0.2	0.62	0.0	99.7
RSW4	44.1	7.2	6.3	1.2	0.72	0.0	104.8

Abbreviations: S^{2-} (T), total sulfide including H_2S, HS^-, and certain metal sulfides.
The detection limits for S^{2-} (T) and Fe^{2+} are 0.01 mg L^{-1} and 0.011 mg L^{-1}, respectively.

Table 3.4/S3. Concentrations of dissolved methane (CH_4), carbon dioxide (CO_2), and hydrogen (H_2) as measured along the outflow channels in RSE and RSW. Data correspond with those presented in Fig. 3.2.

Site	CH_4 (nM)	CO_2 (mM)	H_2 (nM)	DIC (mM)
RSE1	53.2	3.90	272.6	0.050
RSE2	16.4	0.60	49.8	0.001
RSE3	8.3	0.16	38.4	BDL
RSE4	6.1	0.04	10.1	BDL
RSE5	5.0	0.03	11.1	BDL
RSW1	297.6	8.70	168.4	0.98
RSW2	191.2	6.70	74.1	1.02
RSW3	65.6	2.72	24.3	0.92
RSW4	102.7	2.39	40.8	0.93

Abbreviations: BDL, below detection limit. The detection limit for DIC is 0.001 mM.

Table 3.5/S4. Minerals detected by XRD in sediments sampled along the outflow channels of RSE and RSW. Detection of a given mineral is indicated by a grey rectangle and the lack of detection (<1 % of total mass) of a given mineral is indicated by a white rectangle.

Site	Quartz	Alkali Feldspar	Tridymite	Alunite	Kaolinite
RSE1					
RSE2					
RSE3					
RSE4					
RSE5					
RSW1					
RSW2					
RSW3					
RSW4					

Table 3.6/S5. Rates of microbial activities measured in the present study. Data correspond to those presented in Figs. 3.3 and 3.4.

Site	nmol DIC gdm^{-1} hr^{-1} (Light)	nmol DIC gdm^{-1} hr^{-1} (Dark)	Δ H_2 concentration (nmol) hr^{-1} (Dark)
RSE1	1.93	9.19	4.20
RSE2	ND	ND	-3.51
RSE3	1.47	2.06	-1.69
RSE4	ND	ND	-4.34
RSE5	180.19	4.61	3.61
RSW1	6.63	11.12	-5.34
RSW2	3.95	6.59	ND
RSW3	9.13	23.08	-25.85
RSW4	2318.05	113.17	ND

Abbreviations: ND, not detected.

Table 3.7/S6. Predicted sequence coverage and inverse Simpson indices of archaeal, bacterial, and eukaryal small subunit (SSU) rRNA gene sequences amplified from sediment communities at RSE and RSW sampling sites.

Archaea	Coverage	Inverse Simpson
RSE1	0.44	8.613
RSE2	0.44	14.56
RSE3	0.40	13.95
RSE4	0.46	11.73
RSE5	0.47	25.59
RSW1	0.46	29.67
RSW2	0.51	5.669
RSW3	0.48	54.43
RSW4	0.53	19.26

Bacteria	Coverage	Inverse Simpson
RSE5	0.86	13.02
RSW1	0.89	4.715
RSW2	0.88	12.19
RSW3	0.91	2.014
RSW4	0.88	3.863

Eukarya	Coverage	Inverse Simpson
RSE5	0.77	2.990
RSW4	0.52	26.44

Table 3.8/S7. Identities of representative 16S rRNA gene sequences in the enrichment transfers of the most dilute, dilution-to-extinction cultivation experiments (see Table 3.1) conducted on RSE sediments. OTUs were binned at the genus level. The percent identity to a cultivar within the specified genus is indicated within parentheses.

Site	O_2	NO_3^-	Fe^{3+}	SO_4^{2-}	S°
RSE1	*Nitrososphaera* (86%) *Desulfurococcus* (93%) *Metallosphaera* (100%)	ND	*Nitrososphaera* (86%) *Desulfurococcus* (93%) *Metallosphaera* (100%)	NA	*Nitrososphaera* (86%) *Desulfurococcus* (93%)
RSE2	*Nitrososphaera* (86%) *Desulfurococcus* (93%)	ND	*Nitrososphaera* (86%) *Desulfurococcus* (93%)	ND	*Nitrososphaera* (86%) *Desulfurococcus* (93%)
RSE3	NA	ND	*Bacillus* (99%)	ND	ND
RSE4	*Methylophilaceae* (91%) *Sulfobacillus* (100%)	ND	*Paracoccus* (100%)	ND	ND
RSE5	NA	ND	*Methylobacterium* (99%)	ND	ND

Abbreviations: ND, the specified activity was not detected at any of the dilutions or cultivation conditions used in
this study. NA, 16S rRNA genes were not amplifiable.

Table 3.9/S8. Identities of representative 16S rRNA gene sequences in the enrichment transfers of the most dilute, dilution-to-extinction cultivation experiments (see Table 3.1) conducted on RSW sediments. OTUs were binned at the genus level. The percent identity to a cultivar within the specified genus is indicated within parentheses.

Site	O_2	NO_3^-	Fe^{3+}	SO_4^{2-}	S°
RSW1	*Thermofilum* (100%) *Fervidicoccus* (99%)	*Methylobacterium* (99%)	ND	*Thermodesulfo-bacterium* (100%)	*Bacillus* (99%)
RSW3	NA	*Sulfurihydrogenibium* (100%) Ca. *Korarchaeum* (97%)	ND	*Caldiserica* (100%) Ca. *Caldiarchaeum* (97%)	*Dictyoglomus* (100%) *Methanothermo-bacter* (97%)

Abbreviations: ND, the specified activity was not detected at any of the dilutions or cultivation conditions used in this study. NA, 16S rRNA genes were not amplifiable.

Table 3.10/S9. Percent relative abundances of 16S (Archaea and Bacteria) and 18S (Eukarya) rRNA gene transcript OTUs (as represented by their closest cultivated relatives) found in sites at RSE and RSW. Representative OTUs were binned at the order level with orders that represent >5% of total sequences in at least one sampling location included. Data corresponds with those presented in Fig. 3.6.

Order	RSE1	RSE2	RSE3	RSE4	RSE5	RSW1	RSW2	RSW3	RSW4
Thermoproteales	1.58	47.78	1.48	0.33	6.55	32.39	78.64	33.30	31.58
Sulfolobales	73.10	24.65	7.31	3.58	0.05	0.00	0.00	0.00	0.05
Desulfurococcales	0.43	0.05	53.03	72.00	56.62	12.95	0.10	8.74	3.54
Nitrososphaerales	0.00	0.00	0.00	0.14	16.67	7.12	2.87	19.06	43.86
Archaeoglobales	0.05	0.00	0.00	0.00	0.00	13.66	0.72	6.55	0.00
Ca. *Caldiarchaeum*	0.00	0.00	0.00	0.00	0.00	5.35	1.29	9.79	11.18
Methanococcales	0.00	0.00	0.00	0.00	0.00	0.00	6.64	0.00	0.00
Ca. *Korarchaeum*	0.00	0.00	0.00	0.00	0.00	0.29	0.00	2.44	0.00
Thermoflexales	NA	NA	NA	NA	0.70	29.07	11.74	75.37	0.90
Chroococcales	NA	NA	NA	NA	0.78	0.37	0.55	0.37	57.36
Pseudomonadales	NA	NA	NA	NA	18.75	34.41	16.51	0.90	0.42
Lactobacillales	NA	NA	NA	NA	28.47	0.22	2.15	0.20	0.36
Burkholderiales	NA	NA	NA	NA	20.94	10.98	13.63	0.75	0.47
Sphingobacteriales	NA	NA	NA	NA	0.07	0.10	8.61	0.13	0.06
Caldisericales	NA	NA	NA	NA	0.06	0.83	7.01	0.22	0.03
Actinomycetales	NA	NA	NA	NA	0.03	0.07	6.53	0.07	0.04
Clostridiales	NA	NA	NA	NA	1.52	0.08	0.12	0.10	6.33
Anaerolineales	NA	NA	NA	NA	0.14	0.77	1.39	5.17	2.17
Cyanidiales	NA	NA	NA	NA	76.59	NA	NA	NA	35.09
Achnanthales	NA	NA	NA	NA	0.00	NA	NA	NA	15.93

Abbreviations: NA, 16S or 18S rRNA genes were not amplifiable.

Supplemental Figures

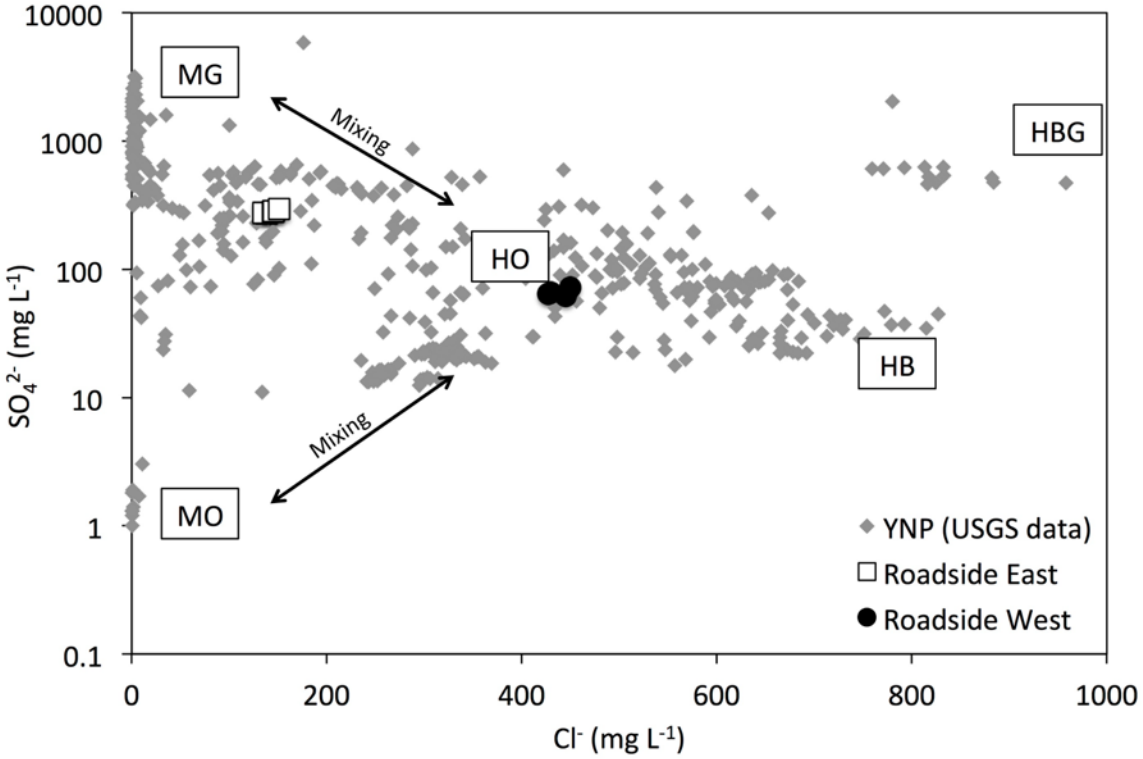


Fig. 3.8/S1. SO_4^{2-} concentrations (mg L^{-1}) in hot spring waters plotted relative to Cl^- concentrations (mg L^{-1}) in the same waters. The nine samples in this study are plotted as white squares (RSE) and black circles (RSW) along with data derived from USGS reports on samples collected throughout YNP during the years 2003-2013 (grey diamonds) (Ball et al., 2006; McCleskey et al., 2014). The classification of hot spring fluid types, as described in Nordstrom et al., 2009, are plotted for reference: hydrothermal water only (HO), hydrothermal water with subsurface boiling (HB), meteoric water only (MO), meteoric water with hot gas discharge (MG), and hydrothermal water with subsurface boiling and hot gas discharge (HBG) (Nordstrom et al., 2009).

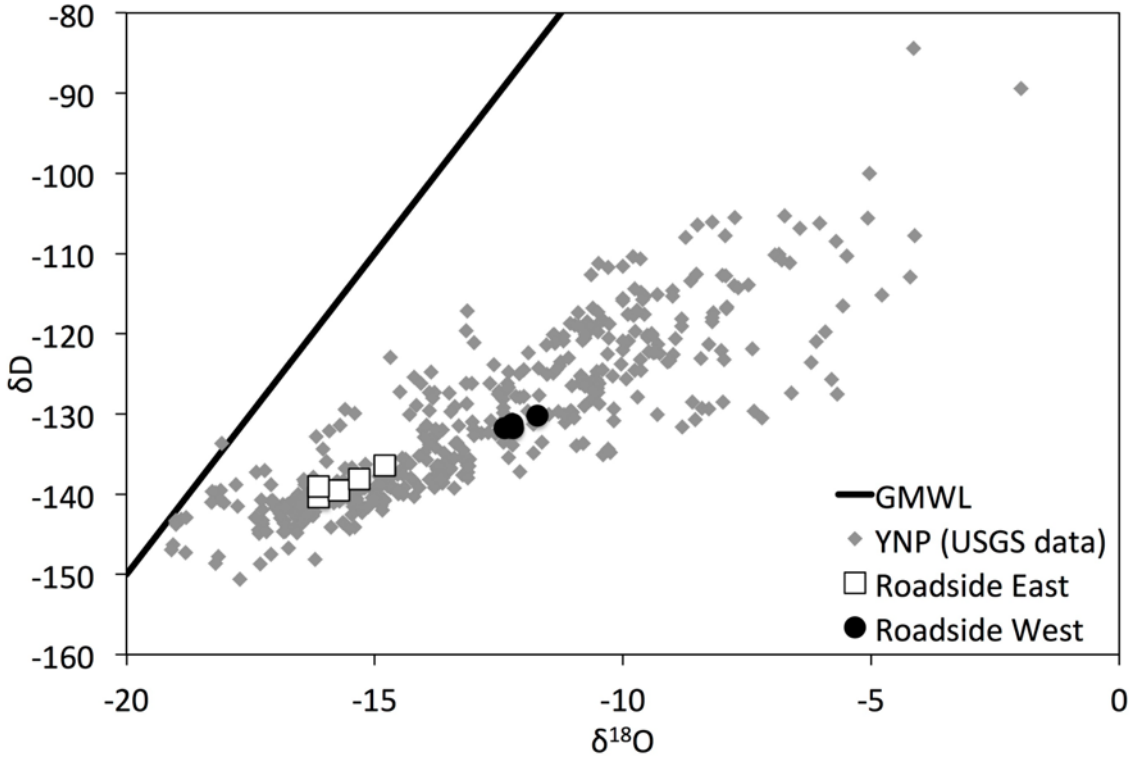


Fig. 3.9/S2. The isotopic composition of hydrogen (δD) in hot spring waters plotted relative to the isotopic composition of oxygen ($\delta^{18}O$) in the same waters. The nine samples in this study are plotted as white squares (RSE) and black circles (RSW) along with data derived from USGS reports on samples throughout YNP as collected from 2003-2013 (grey diamonds) (Ball et al., 2006; McCleskey et al., 2014). The global meteoric water line (GMWL) is plotted for reference (Nordstrom et al., 2009).

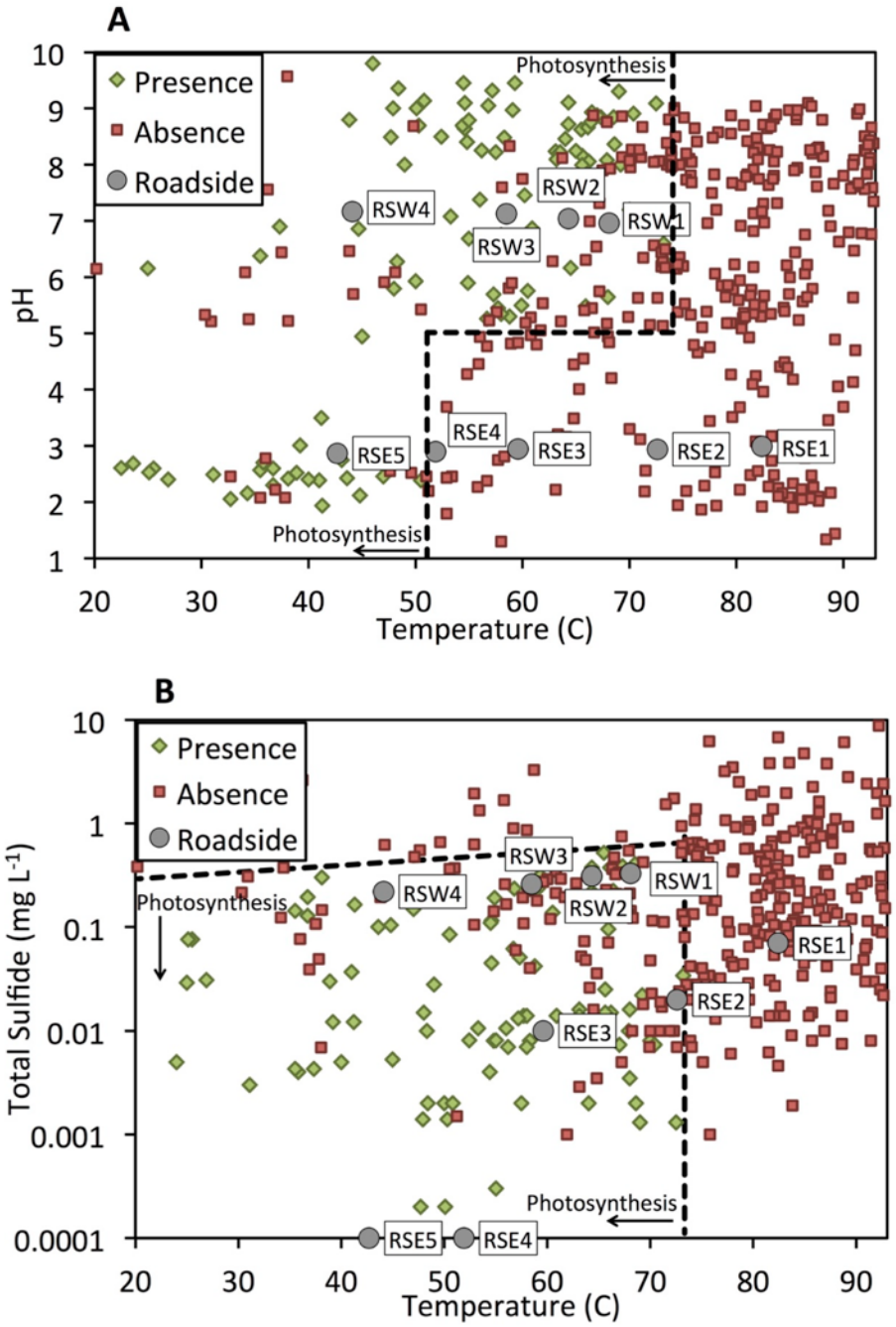


Fig. 3.10/S3. YNP hot spring samples (442 total) plotted as a function of spring water pH and temperature (A) or total sulfide concentration (mg L^{-1}) and temperature (B). Evidence for photosynthesis (presence of phototrophic pigments or genes involved in chlorophyll biosynthesis) is depicted by green diamonds while red squares depict sites where evidence for photosynthesis was not detected. Sampling locations in RSE and RSW hot springs are mapped on this 2-dimensional landscape and are denoted by filled grey circles and labels. Sites RSE5 and RSW4 had evidence of photosynthesis while all other sites in this study did not. Figure adapted from Boyd et al., 2012.

References

Amenabar MJ, Urschel MR, Boyd ES (2015). Metabolic and taxonomic diversification in continental magmatic hydrothermal systems. In: Corien Bakermans DG (ed). *Microbial Evolution under Extreme Conditions.* pp 57-96.

Amenabar MJ, Colman DR, Poudel S, Roden EE, Boyd ES (2018). Electron acceptor availability alters carbon and energy metabolism in a thermoacidophile. *Environmental Microbiology*. **20**: 2523-2537.

Amend JP, Shock EL (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiology Reviews* **25:** 175-243.

Anderson I, Rodriguez J, Susanti D, Porat I, Reich C, Ulrich LE *et al* (2008). Genome sequence of *Thermofilum pendens* reveals an exceptional loss of biosynthetic pathways without genome reduction. *Journal of Bacteriology* **190:** 2957-2965.

Anderson IJ, Sun H, Lapidus A, Copeland A, Glavina Del Rio T, Tice H *et al* (2009). Complete genome sequence of *Staphylothermus marinus* Stetter and Fiala 1986 type strain F1. *Standards in Genomic Sciences* **1:** 183-188.

Ball JW, McCleskey RB, Nordstrom DK, Holloway JM (2006). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, 2003-2005. US Geological Survey Open File Report 2006-1339: Reston, Virginia.

Ball JW, McCleskey RB, Nordstrom DK (2010). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, 2006-2008. US Geological Survey Open File Report 2010-1192: Reston, Virginia.

Bandyopaghyay A, Elvitigala T, Welsh E, Stockel J, Liberton M, Min H *et al* (2011). Novel metabolic attributes of the genus *Cyanothece*, comprising a group of unicellular nitrogen-fixing Cyanobacteria. *mBio* **2:** e00214-00211.

Bernstein HC, Beam JP, Kozubal M, Carlson RP, Inskeep WP (2013). *In situ* analysis of oxygen consumption and diffusive transport in high-temperature acidic iron-oxide microbial mats. *Environmental Microbiology* **15:** 2360-2370.

Boyd ES, Jackson RA, Encarnacion G, Zahn JA, Beard T, Leavitt WD *et al* (2007). Isolation, characterization, and ecology of sulfur-respiring *Crenarchaea* inhabiting acid-sulfate-chloride-containing geothermal springs in Yellowstone National Park. *Applied and Environmental Microbiology* **73:** 6669-6677.

Boyd ES, Leavitt WD, Geesey GG (2009). CO_2 uptake and fixation by a thermoacidophilic microbial community attached to precipitated sulfur in a geothermal spring. *Applied and Environmental Microbiology* **75:** 4289-4296.

Boyd ES, Hamilton TL, Spear JR, Lavin M, Peters JW (2010). [FeFe]-hydrogenase in Yellowstone National Park: evidence for dispersal limitation and phylogenetic niche conservatism. *The ISME Journal* **4:** 1485-1495.

Boyd ES, Lange RK, Mitchell AC, Havig JR, Hamilton TL, Lafreniere MJ *et al* (2011). Diversity, abundance, and potential activity of nitrifying and nitrate-reducing microbial assemblages in a subglacial ecosystem. *Applied and Environmental Microbiology* **77:** 4778-4787.

Boyd ES, Fecteau KM, Havig JR, Shock EL, Peters JW (2012). Modeling the habitat range of phototrophs in Yellowstone National Park: toward the development of a comprehensive fitness landscape. *Frontiers in Microbiology* **3:** 1-11.

Boyd ES, Schut GJ, Adams MWW, Peters JW (2014). Hydrogen metabolism and the evolution of biological respiration. *Microbe* **9:** 361-367.

Brock TD (1967). Life at high temperatures. *Science* **158:** 1012-1019.

Brugna-Guiral M, Tron P, Nitschke W, Stetter KO, Burlat B, Guigliarelli B *et al* (2003). [NiFe] hydrogenases from the hyperthermophilic bacterium *Aquifex aeolicus*: properties, function, and phylogenetics. *Extremophiles* **7:** 145-157.

Campos VL, Valenzuela C, Yarza P, Kampfer P, Vidal R, Zaror C *et al* (2010). *Pseudomonas arsenicoxydans* sp nov., an arsenite-oxidizing strain isolated from the Atacama desert. *Systematic and Applied Microbiology* **33:** 193-197.

Chapelle FH, Vroblesky DA, Woodward JC, Lovley DR (1997). Practical considerations for measuring hydrogen concentrations in groundwater. *Environmental Science and Technology* **31:** 2873-2877.

Colman DR, Poudel S, Hamilton TL, Havig JR, Selensky MJ, Shock EL *et al* (2018). Geobiological feedbacks and the evolution of thermoacidophiles. *The ISME Journal* **12:** 225-236.

Cox A, Shock EL, Havig JR (2011). The transition to microbial photosynthesis in hot spring ecosystems. *Chemical Geology* **280:** 344-351.

Craig H, Gordon LI, Horibe Y (1963). Isotopic exchange effects in the evaporation of water. *Journal of Geophysical Research* **68:** 5079-5087.

D'Imperio S, Leht CR, Oduro H, Druschel G, Kuhl M, McDermott TR (2008). Relative importance of H_2 and H_2S as energy sources for primary production in geothermal springs. *Applied and Environmental Microbiology* **74:** 5802-5808.

Dodsworth JA, Gevorkian J, Despujos F, Cole JK, Murugapiran SK, Ming H *et al* (2014). *Thermoflexus hugenholtzii* gen. nov., sp. nov., a thermophilic, microaerophilic,

filamentous bacterium representing a novel class in the Chloroflexi, Thermoflexia classis nov., and description of Thermoflexaceae fam. nov. and Thermoflexales ord. nov. . *International Journal of Systematic and Evolutionary Microbiology* **64:** 2119-2127.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27:** 2194-2200.

Edwards KJ, Goebel BM, Rodgers TM, Schrenk MO, Gihring TM, Cardona MM *et al* (1999). Geomicrobiology of pyrite (FeS_2) dissolution: case study at Iron Mountain, California. *Geomicrobiology Journal* **16:** 155-179.

Edwards KJ, Bond PL, Gihring TM, Banfield JF (2000). An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. *Science* **10:** 1796-1769.

Fishbain S, Dillon JG, Gough HL, Stahl DA (2003). Linkage of high rates of sulfate reduction in Yellowstone hot springs to unique sequence types in the dissimilatory sulfate respiration pathway. *Applied and Environmental Microbiology* **69:** 3663-3667.

Fogo JK, Popowsky M (1949). Spectrophotometric determination of hydrogen sulfide. *Analytical Chemistry* **21:** 732-734.

Fournier RO, White DE, Truesdell AH (1976). Convective heat flow in Yellowstone National Park. *Proceedings of the 2nd United Nations Symposium on the Development and Use of Geothermal Resources, San Francisco*. U.S. Government Printing Office: Washington, D.C. pp 731-739.

Fournier RO (1989). Geochemistry and dynamics of the Yellowstone National Park hydrothermal system. *Annual Review of Earth and Planetary Sciences* **17:** 13-53.

Giggenbach WF (1978). The isotopic composition of waters from the El Tatio geothermal field, Northern Chile. *Geochimica et Cosmochimica Acta* **42:** 979-988.

Glover DT, Hollingshead SK, Briles DE (2008). *Streptococcus pneumoniae* surface protein pcpA elicits protection against lung infecton and fatal sepsis. *Infection and Immunity* **76:** 2767-2776.

Greening C, Biswas A, Carere CR, Jackson CJ, Taylor MC, Stott MB *et al* (2016). Genomic and metagenomic surveys of hydrogenase distribution indicate H_2 is a widely utilised energy source for microbial growth and survival. *The ISME Journal* **10:** 761-777.

Hamilton TL, Vogl K, Bryant DA, Boyd ES, Peters JW (2012). Environmental constraints defining the distribution, composition, and evolution of chlorophototrophs in thermal features of Yellowstone National Park. *Geobiology* **10:** 236-249.

Hamilton TL, Peters JW, Skidmore ML, Boyd ES (2013). Molecular evidence for an active endogenous microbiome beneath glacial ice. *The ISME Journal* **7:** 1402-1412.

Havig JR, Raymond J, Meyer-Dombard DAR, Zolotova N, Shock EL (2011). Merging isotopes and community genomics in a siliceous sinter-depositing hot spring. *Journal of Geophysical Research: Biogeosciences* **116:** 1-15.

Huber H, Huber R, Stetter KO (2006). *The Prokaryotes*, vol. 3. Springer New York: New York.

Huber H, Stetter KO (2015a). Desulfurococcales ord. nov. In: Whitman W (ed). *Bergey's Manual of Systematics of Archaea and Bacteria.* John Wiley & Sons Inc.

Huber H, Stetter KO (2015b). Thermoproteales. In: Whitman W (ed). *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons, Inc.

Hurwitz S, Lowenstern JB (2014). Dynamics of the Yellowstone hydrothermal system. *Reviews of Geophysics* **52:** 375-411.

Inskeep WP, Ackerman GG, Taylor WP, Kozubal M, Korf S, Macur RE (2005). On the energetics of chemolithotrophy in nonequilibrium systems: case studies of geothermal springs in Yellowstone National Park. *Geobiology* **3:** 297-317.

Jay ZJ, Beam JP, Dohnalkova A, Lohmayer R, Bodle B, Planer-Friedrich B *et al* (2015). *Pyrobaculum yellowstonenesis* strain WP30 respires on elemental sulfur and/or arsenate in circumneutral sulfidic geothermal sediments of Yellowsone National Park. *Applied and Environmental Microbiology* **81:** 5907-5916.

Jeanthon C, L'Haridon S, Cueff V, Banta A, Reysenbach A, Prieur D (2002). Thermodesulfobacterium hydrogeniphilum sp. nov., a thermophilic, chemolithoautotrophic, sulfate-reducing bacterium isolated from a deep-sea hydrothermal vent at Guaymas Basin, and emendation of the genus Thermodesulfobacterium. *International Journal of Systematic and Evolutionary Microbiology* **52:** 765-772.

Johnson DB, Hallberg KB (2005). Acid mine drainage remediation options: a review. *Science of the Total Environment* **338:** 3-14.

Jorgensen NOG, Brandt KK, Nybroe O, Hansen M (2009). *Delftia lacustris* sp. nov., a peptidoglycan-degrading bacterium from fresh water, and emended description of *Delftia tsuruhatensis* as a peptidoglycan-degrading bacterium. *International Journal of Systematic and Evolutionary Microbiology* **59:** 2195-2199.

Kanokratana P, Chanapan S, Pootanakit K, Eurwilaichitr L (2004). Diversity and abundance of Bacteria and Archaea in the Bor Khlueng hot spring in Thailand. *Journal of Basic Microbiology* **44:** 430-444.

Kaster A, Moll J, Parey K, Thauer RK (2011). Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea.

Proceedings of the National Academy of Sciences of the United States of America **108:** 2981-2986.

Kozubal M, Macur RE, Korf S, Taylor WP, Ackerman GG, Nagy A *et al* (2008). Isolation and distribution of a novel iron-oxidizing Crenarchaeon from acidic geothermal springs in Yellowstone National Park. *Applied and Environmental Microbiology* **74:** 942-949.

Kozubal M, Dlakic M, Macur RE, Inskeep WP (2011). Terminal oxidase diversity and function in *Metallosphaera yellowstonensis*: gene expression and protein modeling suggest mechanisms of Fe(II) oxidation in the Sulfolobales. *Applied and Environmental Microbiology* **77:** 1844-1853.

Kozubal M, Macur RE, Jay ZJ, Beam JP, Malfatti SA, Tringe SG *et al* (2012). Microbial iron cycling in acidic geothermal springs of Yellowstone National Park: integrating molecular surveys, geochemical processes, and isolation of novel Fe-active microorganisms. *Frontiers in Microbiology* **3:** 1-16.

Larras F, Keck F, Montuelle B, Rimet F, Bouchez A (2014). Linking diatom sensitivity to herbicides to phylogeny: a step forward for biomonitoring? *Environmental Science and Technology* **48:** 1921-1930.

Laska S, Lottspeich F, Kletzin A (2003). Membrane-bound hydrogenase and sulfur reductase of the hyperthermophilic and acidophilic archaeon *Acidianus ambivalens*. *Microbiology* **149:** 2357-2371.

Lowenstern JB, Bergfeld D, Evans WC, Hurwitz S (2012). Generation and evolution of hydrothermal fluids at Yellowstone: Insights from the Heart Lake Geyser Basin. *Geochemistry, Geophysics, Geosystems* **13:** 1-20.

Martin R, Heilig HGHJ, Zoetendal EG, Jimenez E, Fernandez L, Smidt H *et al* (2007). Cultivation-independent assessment of the bacterial diversity of breast milk among healthy women. *Research in Microbiology* **158:** 31-37.

Matsuzaki M, Misumi O, Shin-I T, Maruyama S, Takahara M, Miyagishima S *et al* (2004). Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. *Nature* **428:** 653-657.

McCleskey RB, Chiu RB, Nordstrom DK, Campbell KM, Roth DA, Ball JW *et al* (2014). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, beginning 2009. USGS Water Resources Dataset: doi:10.5066/F7M043FS.

Nordstrom DK, Ball JW, McCleskey RB (2004). Oxidation reactions for reduced Fe, As, and S in thermal outflows of Yellowstone National Park: biotic or abiotic? In: Wanty RB, Seal II RR (eds). *Water-Rock Interaction*. Taylor & Francis Group: London, UK.

Nordstrom DK, Ball JW, McCleskey RB (2005). Ground water to surface water: chemistry of thermal outflows in Yellowstone National Park In: Inskeep WP, McDermott TR (eds). *Geothermal Biology and Geochmistry in Yellowstone National Park*. Thermal Biology Institute: Montana State University, Bozeman, MT. pp 73-94.

Nordstrom DK, McCleskey RB, Ball JW (2009). Sulfur geochemistry of hydrothermal waters in Yellowstone National Park: IV Acid–sulfate waters. *Applied Geochemistry* **24:** 191-207.

Pihl TD, Schicho RN, Kelly RM, Maier RJ (1989). Characterization of hydrogen-uptake activity in the hyperthermophile *Pyrodictium brockii*. *Proceedings of the National Academy of Sciences of the United States of America* **86:** 138-141.

Poret-Peterson AT, Romaniello SJ, Elser JJ, Shock EL, Hartnett HE, Anbar AD (2013). Active methanogenesis and methanotrophy in a terrestrial hydrothermal ecosystem. *Unpublished, NCBI Entry*.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P *et al* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41:** 590-596.

Rye RO, Truesdell AH (1993). The question of recharge to the geysers and hot springs of Yellowstone National Park. US Geological Survey Professional Paper 93-384.

Schafer C, Friedrich B, Lenz O (2013). Novel, oxygen-insensitive group 5 [NiFe]-hydrogenase in *Ralstonia eutropha*. *Applied and Environmental Microbiology* **79:** 5137-5145.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75:** 7537-7541.

Shock EL, Holland M, Meyer-Dombard DA, Amend JP, Osburn GR, Fischer TP (2010). Quantifying inorganic sources of geochemical energy in hydrothermal ecosystems, Yellowstone National Park, USA. *Geochimica et Cosmochimica Acta* **74:** 4005-4043.

Søndergaard D, Pedersen CNS, Greening C (2016). HydDB: A web tool for hydrogenase classification and analysis. *Scientific Reports* **6:** 1-8.

Spear JR, Walker JJ, McCollom TM, Pace NR (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 2555-2560.

Stieglmeier M, Klingl A, Alves RJE, Rittmann SKMR, Melcher M, Leisch N *et al* (2014). *Nitrososphaera viennensis* gen. nov., sp. nov., an aerobic and mesophilic,

ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *International Journal of Systematic and Evolutionary Microbiology* **64:** 2738-2752.

Stookey LL (1970). Ferrozine - a new spectrophotometric reagent for iron. *Analytical Chemistry* **42:** 772-781.

Toshchakov SV, Korzhenkov AA, Samarov NI, Mazunin IO, Mozhey OI, Shmyr IS *et al* (2015). Complete genome sequence of and proposal of *Thermofilum uzonense* sp. nov. a novel hyperthermophilic crenarchaeon and emended description of the genus *Thermofilum*. *Standards in Genomic Sciences* **10**.

Truesdell AH, Fournier RO (1976). Conditions in the deeper parts of hot spring systems of Yellowstone National Park, Wyoming. US Geological Survey Open File Report 76-428: Reston, Virginia.

Truesdell AH, Nathenson M, Rye RO (1977). The effects of subsurface boiling and dilution on the isotopic compositions of Yellowstone thermal waters. *Journal of Geophysical Research* **82:** 3694-3704.

Urschel MR, Kubo MD, Hoehler TM, Peters JW, Boyd ES (2015). Carbon source preference in chemosynthetic hot spring communities. *Applied and Environmental Microbiology* **81:** 3834-3847.

Valentine DL (2007). Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nature Reviews Microbiology* **5:** 316-323.

van Geldern R, Barth JAC (2012). Optimization of instrument setup and post-run corrections for oxygen and hydrogen stable isotope measurements of water by isotope ratio infrared spectroscopy (IRIS). *Limnology and Oceanography: Methods* **10:** 1024-1036.

Vignais PM, Billoud B, Meyer J (2001). Classification and phylogeny of hydrogenases. *FEMS Microbiology Reviews* **25:** 455-501.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73:** 5261-5267.

White DE, Muffler LJP, Truesdell AH (1971). Vapor-dominated hydrothermal systems compared with hot-water systems. *Economic Geology: Bulletin of the Society of Economic Geologists* **66:** 75-97.

Wilson CL, Hinman NW, Cooper WJ, Brown CF (2000). Hydrogen peroxide cycling in surface geothermal waters of Yellowstone National Park. *Environmental Science and Technology* **34:** 2655-2662.

Wirth R, Chertkov O, Held B, Lapidus A, Nolan M, Lucas S *et al* (2011). Complete genome sequence of *Desulfurococcus mucosus* type strain (O7/1t). *Standards in Genomic Sciences* **4:** 173-182.

Yeatts DS (2006). Characteristics of thermal springs and the shallow ground-water system at Hot Springs National Park, Arkansas. U.S. Geological Survey Scientific Investigations Report 2006-5001: Reston, Virginia.

CHAPTER FOUR

CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

In this dissertation, I aimed to understand how subsurface geological processes influence the distribution, identity, and activities of hydrogenotrophs in hydrothermal environments of YNP. First, using a series of geochemical proxies, I determined that H_2 produced via interactions of water with crustal minerals (most likely those that comprise ferrous iron) in the subsurface partitions into the vapor phase during decompressional boiling of ascending hydrothermal fluids. Variable input of vapor into springs yields substantial differences in H_2 concentration, with the highest concentrations in springs located in highly fractured bedrock near the caldera boundary and at high elevation (e.g. greater than 2500 m). These features together are thought to facilitate gas exhalation and migration to the surface where it can then condense with meteoric water. Based on these data, I hypothesized that springs that are sourced with gas resulting from subsurface phase separation, and that are concentrated near the caldera boundary and at high elevation, are more likely to host abundant H_2-dependent populations. Indeed, springs located in the Smokejumper hot spring area, which exhibit H_2 concentrations that are the highest reported in YNP, are located near the caldera boundary, at a topographic high, and host populations that appear to use H_2 as a reductant. Similarly, microbial populations in springs in other areas at high elevation (i.e. Washburn and Hot Spring Basin), as well as other area of YNP that are sourced with vapor phase gases and that share similar geological features, also appear to be able to use H_2 in their energy

metabolism at increased levels when compared with microbial populations in other springs, as evinced by enrichment of hydrogenase proteins predicted to be capable of H_2 oxidation in reconstructed genomes. The presence and expression of genes that encode oxidative or bidirectional hydrogenases in dominant populations in one of these springs (SJ3) suggests these populations are adapted to use H_2 as an electron donor. The same is expected to be true for populations that inhabit springs that are similarly sourced with vapor phase gases that have mixed with near surface meteoric water (e.g., Washburn and HSB), since they too exhibit high concentration of H_2. Reducing equivalents from H_2 oxidation are likely to support primary production in SJ3 as indicated by cultivation assays and by expression of genes that encode key enzymes involved in CO_2 fixation pathways.

Overall, I suggest that lithogenic H_2 sourced from hydrothermal water-rock interactions in the subsurface of YNP is concentrated in certain geothermal areas by the geological process of phase separation combined with geological and topographic features that facilitate gas migration and exhalation. Springs that are sourced by vapor phase gases have higher concentrations H_2 that, in turn, select for populations capable of using H_2 as a reductant during the assembly of resident microbial communities. It is likely that the links defined between the geological processes that generate and source springs with H_2 and the distribution of H_2-dependent microbial populations in those springs extend to other hydrothermal systems, both today and in the geological past, where similar geological processes are thought to take place.

After determining how subsurface geological processes drive variation in H_2 concentration across YNP hot springs and how these, in turn, influence the distribution of

hydrogenotrophs in hot springs, I then focused on determining how variations in the distribution of oxidants in YNP hot springs selects for different hydrogenotrophs and autotrophs, and how these oxidants impact the rates of activities. Given the apparent role of temporal variation in oxidant availability through geological time and the role of this variation in driving the diversification of [NiFe]-hydrogenases, as outlined previously (Boyd et al 2014, Schut et al 2016), I hypothesized that spatial variation in oxidant availability would influence the distribution of hydrogenotrophs, their taxonomic identity, and their activities in hot springs. I examined this hypothesis in a paired set of springs, RSW and RSE, that are distinguished by the inferred process of phase separation. RSE is acidic due to input of sulfide-replete vapor phase gases that can then be oxidized to yield SO_4^{2-} and protons. In contrast, RSW is circumneutral due primarily to input of sulfide-deplete liquid phase water with a minor component of vapor phase gases due to incomplete separation (continuous model of phase separation). RSE is enriched in SO_4^{2-} and Fe^{2+}, the latter likely due to leaching of bedrock minerals by acid. The near-surface O_2-dependent oxidation of Fe^{2+}, which itself is thought to be mediated by microbial activity (Kozubal et al 2012), leads to enrichment in soluble and solid phase Fe^{3+}. I found a high abundance of hydrogenotrophs capable of metabolically coupling H_2 oxidation with Fe^{3+} oxidation in RSE; this metabolism was less important in RSW where Fe^{2+} and thus Fe^{3+} was less available. Despite high concentrations of SO_4^{2-}, I did not identify abundant H_2 oxidizing, SO_4^{2-}reducing populations, potentially consistent with constraints imposed by acidity (Johnson and Hallberg 2005). However in RSW, hydrogenotrophic organisms coupled H_2 oxidation with SO_4^{2-}, in addition to NO_3^{2-}, and S^0. These observations, when combined with data indicating that the predominant hydrogenotrophs

in RSE and RSW differ at the domain level taxonomic rank, is consistent with phylogenetic data that indicates a prominent role for oxidant availability in the diversification of hydrogentrophs at the level of the [NiFe]-hydrogenase enzymes that catalyze H_2 oxidation (Boyd et al 2014, Greening et al 2016).

The geochemical composition of hot spring fluids and the precipitated minerals that form in outflow channels of hot springs are ultimately the result of differences in subsurface processes, namely the subsurface phase separation of fluids and interactions of these fluids with bedrock and mixing of those fluids with oxygen-rich meteoric waters (Amenabar and Boyd 2019). Importantly, oxidants such as Fe^{3+} and SO_4^{2-} would not be present in these springs and would therefore be unavailable to biology without O_2 to oxidize subsurface sourced Fe^{2+} and sulfide, respectively (Colman et al 2018, Kozubal et al 2012). Thus, it is unlikely that these oxidants would have been available in significant concentrations for biological use on early Earth prior to the emergence of oxygenic photosynthesis and thus available to hydrogenotrophs. This is consistent with the more-recent emergence of [NiFe]-hydrogenase homologs (group 1) that allow for metabolic coupling of H_2 with these substrates when compared with other more deeply branching [NiFe]-hydrogenases (Boyd et al 2014). In summary, this component of the study demonstrated the importance of subsurface geological processes and their interaction with surface processes in shaping the ecology, physiology, and evolution of hydrogenotrophic organisms in hydrothermal ecosystems through their combined role in controlling complex hot spring geochemistry and oxidant availability.

In addition to variation in oxidants used to support H_2 metabolism, the RSE microbial community was shown to be dominated by Archaea, whereas the RSW

community was shown to be dominated by Bacteria. The combination of acidic pH and elevated temperature in RSE, which tends to select for a specific group of Archaea (Colman et al 2018, Valentine 2007), is likely to explain this marked difference in community composition. The predominant processes that control the geochemistry and therefore the pH of hot springs take place in the subsurface, indicating that such processes had and continue to have a profound influence on the distribution and, by extension, the diversification of microbial life at the highest taxonomic ranks.

Throughout this work, I show that subsurface processes exert influence on the distribution and concentrations of H_2, the structure, composition, and abundances of chemosynthetic microbial communities that can utilize abundant and diverse H_2-dependent metabolisms and hydrogenases in the modern hydrothermal system in YNP. This research indicates that the geosphere and biosphere are connected at the level of H_2 in the YNP hydrothermal system. Thus, through conductance of this work I provide new details on the link between geological processes that generate disequilibrium in substrates that can be used to support microbial activity and productivity, with a focus on H_2 and oxidants as a model.

Future Directions

This research provides new understanding of the geomicrobiology of H_2 in YNP hot springs. However, to gain a more complete understanding of how H_2 supports primary producers in YNP hot springs, future directions could aim to identify taxa involved in H_2-dependent primary production. In my future work, I will utilize techniques such as stable isotope probing to identify populations that fix CO_2 using reductant

supplied by lithogenic H_2. Therefore, I will be able to link CO_2 fixation with H_2 oxidation at a taxonomic level, determine the identities and functions associated with key hydrogenotrophic primary producers, and determine how H_2 directly supports primary producers of chemosynthetic communities in Yellowstone. Future directions for this research could also determine how different oxidants affect rates of biological *in situ* H_2 oxidation, and how the enzymes and pathways of active H_2 cycling organisms in hot springs are utilized in H_2 oxidation.

References

Amenabar MJ, Boyd ES (2019). A review of the mechanisms of mineral-based metabolism in early Earth analog rock-hosted hydrothermal ecosystems. *World Journal of Microbiology and Biotechnology* **35:** 1-18.

Boyd ES, Schut GJ, Adams MWW, Peters JW (2014). Hydrogen metabolism and the evolution of biological respiration. *Microbe* **9:** 361-367.

Colman DR, Poudel S, Hamilton TL, Havig JR, Selensky MJ, Shock EL *et al* (2018). Geobiological feedbacks and the evolution of thermoacidophiles. *The ISME Journal*.

Greening C, Biswas A, Carere CR, Jackson CJ, Taylor MC, Stott MB *et al* (2016). Genomic and metagenomic surveys of hydrogenase distribution indicate H2 is a widely utilised energy source for microbial growth and survival. *The ISME Journal* **10:** 761-777.

Inskeep WP, Jay ZJ, Tringe SG, Herrgard MJ, Rusch DB, Members YMPSCaWG (2013). The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem. *Frontiers in Microbiology* **4:** 1-15.

Johnson DB, Hallberg KB (2005). Acid mine drainage remediation options: a review. *Science of the Total Environment* **338:** 3-14.

Kozubal M, Macur RE, Jay ZJ, Beam JP, Malfatti SA, Tringe SG *et al* (2012). Microbial iron cycling in acidic geothermal springs of Yellowstone National Park: integrating molecular surveys, geochemical processes, and isolation of novel Fe-active microorganisms. *Frontiers in Microbiology* **3**.

Schut GJ, Zadvornyy O, Wu C-H, Peters JW, Boyd ES, Adams MWW (2016). The role of geochemistry and energetics in the evolution of modern respiratory complexes from a proton-reducing ancestor. *Biochimica et Biophysica Acta* **1857:** 958-970.

Valentine DL (2007). Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nature Reviews Microbiology* **5:** 316-323.

CUMULATIVE REFERENCES

Adam N, Perner M (2018). Microbially mediated hydrogen cycling in deep-sea hydrothermal vents. *Frontiers in Microbiology* **9:** 1-17.

Adams MWW, Eccleston E, Howard JB (1989). Iron-sulfur clusters of hydrogenase I and hydrogenase II of *Clostridium pasteurianum*. *Proceedings of the National Academy of Sciences of the United States of America* **86:** 4932-4936.

Adams MWW, Mortenson LE (1984). The physical and catalytic properties of hydrogenase II of *Clostridium pasteurianum*. *The Journal of Biological Chemistry* **259:** 7045-7055.

Allen ET, Day AL (1935). *Hot springs of the Yellowstone national park*. Carnegie Institution of Washington: Washington.

Amenabar MJ, Boyd ES (2019). A review of the mechanisms of mineral-based metabolism in early Earth analog rock-hosted hydrothermal ecosystems. *World Journal of Microbiology and Biotechnology* **35:** 1-18.

Amenabar MJ, Colman DR, Poudel S, Roden EE, Boyd ES (2018). Electron acceptor availability alters carbon and energy metabolism in a thermoacidophile. *Environmental Microbiology*. **20:** 2523-2537.

Amenabar MJ, Urschel MR, Boyd ES (2015). Metabolic and taxonomic diversification in continental magmatic hydrothermal systems. In: Corien Bakermans DG (ed). *Microbial Evolution under Extreme Conditions* pp 57-96.

Amend JP, Shock EL (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiology Reviews* **25:** 175-243.

Anbar AD (2008). Elements and Evolution. *Science* **322:** 1481-1483.

Anderson I, Rodriguez J, Susanti D, Porat I, Reich C, Ulrich LE *et al* (2008). Genome sequence of *Thermofilum pendens* reveals an exceptional loss of biosynthetic pathways without genome reduction. *Journal of Bacteriology* **190:** 2957-2965.

Anderson IJ, Sun H, Lapidus A, Copeland A, Glavina Del Rio T, Tice H *et al* (2009). Complete genome sequence of *Staphylothermus marinus* Stetter and Fiala 1986 type strain F1. *Standards in Genomic Sciences* **1:** 183-188.

Atkins PW (1990). Physical Chemistry. *Physical Chemistry*. Oxford University Press.

Ball JW, McCleskey RB, Nordstrom DK (2010). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, 2006-2008. United States Geological Survey Open-File Report 2010-1192: Reston, Virginia.

Ball JW, McCleskey RB, Nordstrom DK, Holloway JM (2006). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, 2003-2005. United States Geological Survey Open-File Report 2006-1339: Reston, Virginia.

Bandyopaghyay A, Elvitigala T, Welsh E, Stockel J, Liberton M, Min H *et al* (2011). Novel metabolic attributes of the genus *Cyanothece*, comprising a group of unicellular nitrogen-fixing Cyanobacteria. *mBio* **2:** e00214-00211.

Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hugler M *et al* (2010). Autotrophic carbon fixation in archaea. *Nature Reviews Microbiology* **8:** 447-460.

Bergfeld D, Lowenstern JB, Hunt AG, Pat Shanks III WC, Evans WC (2014). Gas and Isotope Chemistry of Thermal Features in Yellowstone National Park, Wyoming. *USGS Scientific Investigations Report* **2011-5012**.

Bernstein HC, Beam JP, Kozubal M, Carlson RP, Inskeep WP (2013). *In situ* analysis of oxygen consumption and diffusive transport in high-temperature acidic iron-oxide microbial mats. *Environmental Microbiology* **15:** 2360-2370.

Boga HI, Brune A (2003). Hydrogen-dependent oxygen reduction by homoacetogenic bacteria isolated from termite guts. *Applied and Environmental Microbiology* **69:** 779-786.

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30:** 2114-2120.

Bothe H, Tennigkeit J, Eisbrenner G, Yates MG (1977). The hydrogenase-nitrogenase relationship in the blue-green alga *Anabaena cylindrica*. *Planta* **133:** 237-242.

Boyd ES, Anbar AD, Miller SL, Hamilton TL, Lavin M, Peters JW (2011). A late methanogen origin for molybdenum-dependent nitrogenase. *Geobiology* **9:** 221-232.

Boyd ES, Fecteau KM, Havig JR, Shock EL, Peters JW (2012). Modeling the habitat range of phototrophs in yellowstone national park: toward the development of a comprehensive fitness landscape. *Frontiers in Microbiology* **3:** 221.

Boyd ES, Hamilton TL, Spear JR, Lavin M, Peters JW (2010). [FeFe]-hydrogenase in Yellowstone National Park: evidence for dispersal limitation and phylogenetic niche conservatism. *The ISME Journal* **4:** 1485-1495.

Boyd ES, Jackson RA, Encarnacion G, Zahn JA, Beard T, Leavitt WD *et al* (2007). Isolation, characterization, and ecology of sulfur-respiring *Crenarchaea* inhabiting acid-sulfate-chloride-containing geothermal springs in Yellowstone National Park. *Applied and Environmental Microbiology* **73:** 6669-6677.

Boyd ES, Lange RK, Mitchell AC, Havig JR, Hamilton TL, Lafreniere MJ *et al* (2011). Diversity, abundance, and potential activity of nitrifying and nitrate-reducing microbial assemblages in a subglacial ecosystem. *Applied and Environmental Microbiology* **77:** 4778-4787.

Boyd ES, Leavitt WD, Geesey GG (2009). CO_2 uptake and fixation by a thermoacidophilic microbial community attached to precipitated sulfur in a geothermal spring. *Applied and Environmental Microbiology* **75:** 4289-4296.

Boyd ES, Peters JW (2013). New insights into the evolutionary history of biological nitrogen fixation. *Frontiers in Microbiology* **4:** 1-12.

Boyd ES, Schut GJ, Adams MWW, Peters JW (2014). Hydrogen metabolism and the evolution of biological respiration. *Microbe* **9:** 361-367.

Bradley JA, Daille LK, Trivedi CB, Bojanowski CL, Stamps BW, Stevenson BS *et al* (2017). Carbonate-rich dendrolitic cones: insights into a modern analog for incipient microbialite formation, Little Hot Creek, Long Valley Caldera, California. *NPJ Biofilms* **32**.

Brock TD (1967). Life at high temperatures. *Science* **158:** 1012-1019.

Brugna-Guiral M, Tron P, Nitschke W, Stetter KO, Burlat B, Guigliarelli B *et al* (2003). [NiFe] hydrogenases from the hyperthermophilic bacterium *Aquifex aeolicus*: properties, function, and phylogenetics. *Extremophiles* **7:** 145-157.

Buckel W, Thauer RK (2013). Energy conservation via electron bifurcating ferredoxin reduction and proton/Na^+ translocating ferredoxin oxidation. *Biochimica et Biophysica Acta* **1827:** 94-113.

Campos VL, Valenzuela C, Yarza P, Kampfer P, Vidal R, Zaror C *et al* (2010). *Pseudomonas arsenicoxydans* sp nov., an arsenite-oxidizing strain isolated from the Atacama desert. *Systematic and Applied Microbiology* **33:** 193-197.

Canfield DE, Farquhar J (2009). Animal evolution, bioturbation, and the sulfate concentrations of the oceans. *Proceedings of the National Academy of Science of the United States of America* **106:** 8123-8127.

Canovas III PA (2016). Energy transfer between the geosphere and biosphere, Arizona State University.

Catling DC, Zahnle KJ, McKay CP (2001). Biogenic methane, hydrogen escape, and the irreversible oxidation of early Earth. *Science* **293:** 839-843.

Chapelle FH, Vroblesky DA, Woodward JC, Lovley DR (1997). Practical considerations for measuring hydrogen concentrations in groundwater. *Environmental Science and Technology* **31:** 2873-2877.

Christiansen RL (2001). The Quarternary and Pliocene Yellowstone Plateau Volcanic Field of Wyoming, Idaho, and Montana. USGS Professional Paper 729-G: Denver, CO.

Colman DR, Feyhl-Buska J, Robinson KJ, Fecteau KM, Xu H, Shock EL *et al* (2016). Ecological differentiation in planktonic and sediment-associated chemotrophic microbial populations in Yellowstone hot springs. *FEMS* **92:** 1-13.

Colman DR, Lindsay MR, Boyd ES (2019). Mixing of meteoric and geothermal fluids supports hyperdiverse chemosynthetic hydrothermal communities. *Nature Communications*. **10:** 1-13.

Colman DR, Poudel S, Hamilton TL, Havig JR, Selensky MJ, Shock EL *et al* (2018). Geobiological feedbacks and the evolution of thermoacidophiles. *The ISME Journal* **12:** 225-236.

Cox A, Shock EL, Havig JR (2011). The transition to microbial photosynthesis in hot spring ecosystems. *Chemical Geology* **280:** 344-351.

Craig H, Gordon LI, Horibe Y (1963). Isotopic exchange effects in the evaporation of water. *Journal of Geophysical Research* **68:** 5079-5087.

D'Imperio S, Leht CR, Oduro H, Druschel G, Kuhl M, McDermott TR (2008). Relative importance of H_2 and H_2S as energy sources for primary production in geothermal springs. *Applied and Environmental Microbiology* **74:** 5802-5808.

de Poorter LM, Geerts WJ, Keltjens JT (2005). Hydrogen concentrations in methane-forming cells probed by the ratios of reduced and oxidized coenzyme F_{420}. *Microbiology* **151:** 1697-1705.

Dick GJ, Andersson AF, Baker BJ, Simmons SF, Thomas BC, Yelton AP *et al* (2009). Community-wide analysis of microbial genome sequence signatures. *Genome Biology* **10:** 1-16.

Djokic T, Van Kranendonk MJ, Campbell KM, Walter MR, Ward CR (2016). Earliest signs of life on land preserved in ca. 3.5 Ga hot spring deposits. *Nature Communications* **8:** 15263.

Dodsworth JA, Gevorkian J, Despujos F, Cole JK, Murugapiran SK, Ming H *et al* (2014). *Thermoflexus hugenholtzii* gen. nov., sp. nov., a thermophilic, microaerophilic, filamentous bacterium representing a novel class in the Chloroflexi, Thermoflexia classis nov., and description of Thermoflexaceae fam. nov. and Thermoflexales ord. nov. . *International Journal of Systematic and Evolutionary Microbiology* **64:** 2119-2127.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27:** 2194-2200.

Edwards KJ, Bond PL, Gihring TM, Banfield JF (2000). An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. *Science* **10:** 1796-1769.

Edwards KJ, Goebel BM, Rodgers TM, Schrenk MO, Gihring TM, Cardona MM *et al* (1999). Geomicrobiology of pyrite (FeS2) dissolution: case study at Iron Mountain, California. *Geomicrobiology Journal* **16:** 155-179.

Elderfield H, Schultz A (1996). Mid-ocean ridge hydrothermal fluxes and the chemical composition of the ocean. *Annual Review of Earth and Planetary Science* **24:** 191-224.

Fishbain S, Dillon JG, Gough HL, Stahl DA (2003). Linkage of high rates of sulfate reduction in Yellowstone hot springs to unique sequence types in the dissimilatory sulfate respiration pathway. *Applied and Environmental Microbiology* **69:** 3663-3667.

Fogo JK, Popowsky M (1949). Spectrophotometric determination of hydrogen sulfide. *Analytical Chemistry* **21:** 732-734.

Fournier RO (1989). Geochemistry and dynamics of the Yellowstone National Park hydrothermal system. *Annual Review of Earth and Planetary Science* **17:** 13-53.

Fournier RO (1989). Geochemistry and dynamics of the Yellowstone National Park hydrothermal system. *Annual Review of Earth and Planetary Sciences* **17:** 13-53.

Fournier RO, White DE, Truesdell AH (1976). Convective heat flow in Yellowstone National Park. *Proceedings of the U N Symposium on Development and Use of Geothermal Resources, San Francisco* **1:** 731-739.

Fournier RO, White DE, Truesdell AH (1976). Convective heat flow in Yellowstone National Park. *Proceedings of the 2nd United Nations Symposium on the Development and Use of Geothermal Resources*. U.S. Government Printing Office: Washington, D.C. pp 731-739.

Garvin J, Buick R, Anbar AD, Arnold GL, Kaufman AJ (2009). Isotopic evidence for an aerobic nitrogen cycle in the latest Archaean. *Science* **323:** 1045-1048.

Giggenbach WF (1978). The isotopic composition of waters from the El Tatio geothermal field, Northern Chile. *Geochimica et Cosmochimica Acta* **42:** 979-988.

Giggenbach WF, Poreda RJ (1993). Helium isotopic and chemical composition of gases from volcanic-hydrothermal systems in the Phillipines. *Geothermics* **22:** 369-380.

Giggenbach WF, Sano Y, Wakita H (1993). Isotopic composition of helium, and CO_2 and CH_4 contents in gases produced along the New Zealand part of a convergent plate boundary. *Geochimica et Cosmochimica Acta* **57:** 3427-3455.

Glover DT, Hollingshead SK, Briles DE (2008). *Streptococcus pneumoniae* surface protein pcpA elicits protection against lung infecton and fatal sepsis. *Infection and Immunity* **76:** 2767-2776.

Godfrey LV, Falkowski PG (2009). The cycling and redox state of nitrogen in the Archaean ocean. *Nature Geoscience* **2:** 725-729.

Gouy M, Guindon S, Gascuel O (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27:** 221-224.

Greening C, Biswas A, Carere CR, Jackson CJ, Taylor MC, Stott MB *et al* (2016). Genomic and metagenomic surveys of hydrogenase distribution indicate H_2 is a widely utilised energy source for microbial growth and survival. *The ISME Journal* **10:** 761-777.

Hamilton TL, Peters JW, Skidmore ML, Boyd ES (2013). Molecular evidence for an active endogenous microbiome beneath glacial ice. *The ISME journal* **7:** 1402-1412.

Hamilton TL, Vogl K, Bryant DA, Boyd ES, Peters JW (2012). Environmental constraints defining the distribution, composition, and evolution of chlorophototrophs in thermal features of Yellowstone National Park. *Geobiology* **10:** 236-249.

Havig JR, Raymond J, Meyer-Dombard DAR, Zolotova N, Shock EL (2011). Merging isotopes and community genomics in a siliceous sinter-depositing hot spring. *Journal of Geophysical Research: Biogeosciences* **116:** 1-15.

Hedderich R (2004). Energy-converting [NiFe] hydrogenases from archaea and extremophiles: ancestors of complex I. *Journal of Bioenergetics and Biomembranes* **36:** 65-75.

Hoehler TM (2005). Biogeochemistry of dihydrogen. In: Sigel A, Sigel H, Sigel RKO (eds). *Biogeochemical Cycles of Elements*. Marcel Dekker: New York. pp 9-48.

Hoehler TM, Albert DB, Alperin MJ, Bebout BM, Martens CS, Des Marais DJ (2002). Comparative ecology of H_2 cycling in sedimentary and phototrophic ecosystems. *Antonie van Leeuwenhoek* **81:** 575-585.

Huber H, Huber R, Stetter KO (2006). *The Prokaryotes*, vol. 3. Springer New York: New York.

Huber H, Stetter KO (2015a). Desulfurococcales ord. nov. In: Whitman W (ed). *Bergey's Manual of Systematics of Archaea and Bacteria.* John Wiley & Sons Inc.

Huber H, Stetter KO (2015b). Thermoproteales. In: Whitman W (ed). *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons, Inc.

Hugenholtz P, Pitulle C, Hershberger KL, Pace NR (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *Journal of Bacteriology* **180:** 366-376.

Hunten DM (1973). The escape of light gases from planetary atmospheres. *Journal of the Atmospheric Sciences* **30:** 1481-1494.

Hurwitz S, Lowenstern JB (2014). Dynamics of the Yellowstone hydrothermal system. *Reviews of Geophysics* **52:** 375-411.

Inskeep WP, Ackerman GG, Taylor WP, Kozubal M, Korf S, Macur RE (2005). On the energetics of chemolithotrophy in nonequilibrium systems: case studies of geothermal springs in Yellowstone National Park. *Geobiology* **3:** 297-317.

Inskeep WP, Jay ZJ, Tringe SG, Herrgard MJ, Rusch DB, YNP Metagenome Project Steering Committee and Working Group Members (2013). The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem. *Frontiers in Microbiology* **4:** 1-15.

Jay ZJ, Beam JP, Dohnalkova A, Lohmayer R, Bodle B, Planer-Friedrich B *et al* (2015). *Pyrobaculum yellowstonenesis* strain WP30 respires on elemental sulfur and/or arsenate in circumneutral sulfidic geothermal sediments of Yellowsone National Park. *Applied and Environmental Microbiology* **81:** 5907-5916.

Jeanthon C, L'Haridon S, Cueff V, Banta A, Reysenbach A, Prieur D (2002). Thermodesulfobacterium hydrogeniphilum sp. nov., a thermophilic, chemolithoautotrophic, sulfate-reducing bacterium isolated from a deep-sea hydrothermal vent at Guaymas Basin, and emendation of the genus Thermodesulfobacterium. *International Journal of Systematic and Evolutionary Microbiology* **52:** 765-772.

Jochimsen B, Peinemann-Simon S, Völker H, Stüben D, Botz R, Stoffers P *et al* (1997). *Stetteria hydrogenophila* gen. nov. and sp. nov., a novel mixotrophic sulfur-dependent crenarchaeote isolated from Milos, Greece. *Extremophiles* **1:** 67-73.

Johnson DB, Hallberg KB (2005). Acid mine drainage remediation options: a review. *Science of the Total Environment* **338:** 3-14.

Jorgensen NOG, Brandt KK, Nybroe O, Hansen M (2009). *Delftia lacustris* sp. nov., a peptidoglycan-degrading bacterium from fresh water, and emended description of *Delftia tsuruhatensis* as a peptidoglycan-degrading bacterium. *International Journal of Systematic and Evolutionary Microbiology* **59:** 2195-2199.

Kang DD, Froula J, Egan R, Wang Z (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3:** 1-15.

Kanokratana P, Chanapan S, Pootanakit K, Eurwilaichitr L (2004). Diversity and abundance of Bacteria and Archaea in the Bor Khlueng hot spring in Thailand. *Journal of Basic Microbiology* **44:** 430-444.

Kaster A, Moll J, Parey K, Thauer RK (2011). Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea. *Proceedings of the National Academy of Science of the United States of America* **108:** 2981-2986.

Kelley DS, Karson JA, Fruh-Green GL, Yoerger DR, Shank TM, Butterfield DA *et al* (2005). A sepentinite-hosted ecosystem: the lost city hydrothermal field. *Science* **307:** 1428-1434.

Kita I, Matsuo S, Wakita H (1982). H_2 generation by reaction between H_2O and crushed rock: an experimental study on H_2 degassing from the active fault zone. *Journal of Geophysical Research* **87:** 10789-10795.

Kozubal M, Dlakic M, Macur RE, Inskeep WP (2011). Terminal oxidase diversity and function in *Metallosphaera yellowstonensis*: gene expression and protein modeling suggest mechanisms of Fe(II) oxidation in the Sulfolobales. *Applied and Environmental Microbiology* **77:** 1844-1853.

Kozubal M, Macur RE, Jay ZJ, Beam JP, Malfatti SA, Tringe SG *et al* (2012). Microbial iron cycling in acidic geothermal springs of Yellowstone National Park: integrating molecular surveys, geochemical processes, and isolation of novel Fe-active microorganisms. *Frontiers in Microbiology* **3:** 1-16.

Kozubal M, Macur RE, Korf S, Taylor WP, Ackerman GG, Nagy A *et al* (2008). Isolation and distribution of a novel iron-oxidizing Crenarchaeon from acidic geothermal springs in Yellowstone National Park. *Applied and Environmental Microbiology* **74:** 942-949.

Kral TA, Brink KM, Miller SL, McKay CP (1998). Hydrogen consumption by methanogens on the early Earth. *Origins of life and evolution of the biosphere* **28:** 311-319.

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9:** 357-359.

Larras F, Keck F, Montuelle B, Rimet F, Bouchez A (2014). Linking diatom sensitivity to herbicides to phylogeny: a step forward for biomonitoring? *Environmental Science and Technology* **48:** 1921-1930.

Laska S, Lottspeich F, Kletzin A (2003). Membrane-bound hydrogenase and sulfur reductase of the hyperthermophilic and acidophilic archaeon *Acidianus ambivalens*. *Microbiology* **149:** 2357-2371.

Lenz O, Friedrich B (1998). A novel multicomponent regulatory system mediates H_2 sensing in *Alcaligenes eutrophus*. *Proceedings of the National Academy of Science of the United States of America* **95:** 12474-12479.

Li D, Liu CM, Luo R, Sadakane K, Lam TW (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Brujin graph. *Bioinformatics* **31:** 1674-1676.

Lie TJ, Costa KC, Lupa B, Korpole S, Whitman WB, Leigh JA (2012). Essential anaplerotic role for the energy-converting hydrogenase Eha in hydrogenotrophic methanogenesis. *Proceedings of the National Academy of Science of the United States of America* **109:** 15473-15478.

Lin LH, Hall JR, Lippmann-Pipke J, Ward JA, Lollar BS, DeFlaun M *et al* (2005). Radiolytic H_2 in continental crust: nuclear power for deep subsurface microbial communities. *Geochemistry, Geophysics, Geosystems* **6**.

Lindsay MR, Amenabar MJ, Fecteau KM, Debes RV, Fernandes MC, Urschel MR *et al* (2018). Subsurface Processes Influence Oxidant Availability and Chemoautotrophic Hydrogen Metabolism in Yellowstone Hot Springs. *Geobiology* **16:** 674-692.

Lineweaver CH, Schwartzman D (2004). Cosmic Thermobiology. In: Seckbach J (ed). *Origins. Cellular Origin, Life in Extreme Habitates and Astrobiology*. Springer: Dordrecht.

Lovley DR (1985). Minimum threshold for hydrogen metabolism in methanogenic bacteria. *Applied and Environmental Microbiology* **49:** 1530-1531.

Lowenstern JB, Bergfeld D, Evans WC, Hunt AG (2015). Origins of geothermal gases at Yellowstone. *Journal of Volcanology and Geothermal Research* **302:** 87-101.

Lowenstern JB, Bergfeld D, Evans WC, Hurwitz S (2012). Generation and evolution of hydrothermal fluids at Yellowstone: Insights from the Heart Lake Geyser Basin. *Geochemistry, Geophysics, Geosystems* **13:** 1-20.

Lowenstern JB, Evans WC, Bergfeld D, Hunt AG (2014). Prodigious degassing of a billion years of accumulated radiogenic helium at Yellowstone. *Nature* **506:** 355-358.

Lowenstern JB, Janik CJ (2003). The origins of reservoir liquids and vapors from The Geysers geothermal field, California (USA). In: Simmons SF, Graham I (eds). *Volcanic, geothermal and ore-forming fluids; rulers and witnesses of processes within the Earth*. Society of Economic Geologists Special Publication. pp 181-195.

Luther GW 3rd, Findlay AJ, Macdonald DJ, Owings SM, Hanson TE, Beinart RA, Girguis PR (2011). Thermodynamics and kinetics of sulfide oxidation by oxygen: a look

at inorganically controlled reactions and biologically mediated processes in the environment. *Frontiers in Microbiology* **2**: 1-9.

Lyons TW, Reinhard CT, Planavsky NJ (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **506:** 307-315.

Markowitz VM, Chen IM, Chu K, Szeto E, Palaniappan K, Grechkin Y *et al* (2012). IMG/M: the integrated metagenome data management and comparative analysis system. *Nucleic Acids Research* **40:** D123-129.

Marrink SJ, Berendsen HJC (1996). Permeation process of small molecules across lipid membranes studied by molecular dynamic simulations. *Journal of Physical Chemistry* **100:** 16729-16738.

Martin R, Heilig HGHJ, Zoetendal EG, Jimenez E, Fernandez L, Smidt H *et al* (2007). Cultivation-independent assessment of the bacterial diversity of breast milk among healthy women. *Research in Microbiology* **158:** 31-37.

Martin W, Müller M (1998). The hydrogen hypothesis for the first eukaryote. *Nature* **392:** 37-41.

Martin WF (2012). Hydrogen, metals, bifurcating electrons, and proton gradients: The early evolution of biological energy conservation. *FEBS Letters* **586:** 485-493.

Masukawa H, Mochimaru M, Sakurai H (2002). Disruption of the uptake hydrogenase gene, but not of the bidirectional hydrogenase gene, leads to enhanced photobiological hydrogen production by the nitrogen-fixing cyanobacterium *Anabaena* sp. PCC 7120. *Applied Microbiology and Biotechnology* **58:** 618-624.

Matsuzaki M, Misumi O, Shin-I T, Maruyama S, Takahara M, Miyagishima S *et al* (2004). Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. *Nature* **428:** 653-657.

McCleskey RB, Chiu RB, Nordstrom DK, Campbell KM, Roth DA, Ball JW *et al* (2014). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, beginning 2009. US Geological Survey Data Report.

McCleskey RB, Chiu RB, Nordstrom DK, Holloway JM (2014). Water-chemistry data for selected springs, geysers, and streams in Yellowstone National Park, Wyoming, beginning 2009. USGS Water Resources Dataset: doi:10.5066/F7M043FS.

McCollom TM, Shock EL (1997). Geochemical constrains on chemolithoautotrophic metabolism by microorganisms in seafloor hydrothermal systems. *Geochimica et Cosmochimica Acta* **61:** 4375-4391.

Meyer J (2007). [FeFe] hydrogenases and their evolution: a genomic perspective. *Cellular and Molecular Life Sciences* **64:** 1063-1084.

Mikheenko A, Saveliev V, Gurevich A (2015). MetaQUAST: evaluation of metagenome assemblies. *Bioinformatics* **32:** 1088-1090.

Moore EK, Jelen BI, Giovannelli D, Raanan H, Falkowski PG (2017). Metal availability and the expanding network of microbial metabolisms in the Archaean eon. *Nature Geoscience* **10:** 629-636.

Mori K, Yamaguchi K, Sakiyama Y, Urabe T, Suzuki K (2009). *Caldisericum exile* gen. nov., sp. nov., an anaerobic, thermophilic, filamentous bacterium of a novel bacteriual phylum, *Caldiserica* phyl. nov. originally called the candidate phylum OP5, and description of *Caldisericaceae* fam. nov., *Caldisericales* ord. nov. and *Caldisericia* classis nov. *International Journal of Systemic and Evolutionary Microbiology* **59:** 2894-2898.

Mulder DW, Boyd ES, Sarma R, Lange RK, Endrizzi JA, Broderick JB *et al* (2010). Stepwise [FeFe]-hydrogenase H-cluster assembly revealed in the structure of HydA^DEFG. *Nature* **465:** 248-251.

Nakagawa S, Shtaih Z, Banta A, Beveridge TJ, Sako Y, Reysenbach A-L (2005). *Sulfurihydrogenibium yellowstonense* sp. nov., an extremely thermophilic, facultatively heterotrophic, sulfur-oxidizing bacterium from Yellowstone National Park, and emended descriptions of the genus *Sulfurihydrogenibium, Sulfurihydrogenibium subterraneum* and *Sulfurihydrogenibium azorense*. *International Journal of Systemic and Evolutionary Microbiology* **55:** 2263-2268.

Neal C, Stanger G (1983). Hydrogen generation from mantle source rocks in Oman. *Earth and Planetary Science Letters* **66:** 315-320.

Nicolet Y, Piras C, Legrand P, Hatchikian CE, Fontecilla-Camps JC (1999). *Desulfovibrio desulfuricans* iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure* **7:** 13-23.

Nordstrom DK, Ball JW, McCleskey RB (2004). Oxidation reactions for reduced Fe, As, and S in thermal outflows of Yellowstone National Park: biotic or abiotic? In: Wanty RB, Seal II RR (eds). *Water-Rock Interaction*. Taylor & Francis Group: London, UK.

Nordstrom DK, Ball JW, McCleskey RB (2005). Ground water to surface water: chemistry of thermal outflows in Yellowstone National Park In: Inskeep WP, McDermott (eds). *Geothermal Biology and Geochemistry in Yellowstone National Park*. Thermal Biology Institute: Montana State University, Bozeman, MT. pp 73-94.

Nordstrom DK, McCleskey RB, Ball JW (2009). Sulfur geochemistry of hydrothermal waters in Yellowstone National Park: IV Acid–sulfate waters. *Applied Geochemistry* **24:** 191-207.

Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A *et al* (2013). Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *Journal of Computational Biology* **20:** 714-737.

Pace NR (1991). Origin of life - facing up to the physical setting. *Cell* **65:** 531-533.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research* **25:** 1043-1055.

Peters JW, Lanzilotta WN, Lemon BJ, Seefeldt LC (1998). X-ray crystal structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 angstrom resolution. *Science* **282:** 1853-1858.

Peters JW, Schut GJ, Boyd ES, Mulder DW, Shepard ES, Broderick JB *et al* (2015). [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. *Biochimica et Biophysica Acta - Molecular Cell Research* **1853:** 1350-1369.

Pihl TD, Schicho RN, Kelly RM, Maier RJ (1989). Characterization of hydrogen-uptake activity in the hyperthermophile *Pyrodictium brockii*. *Proceedings of the National Academy of Sciences of the United States of America* **86:** 138-141.

Porat I, Kim E, Hendrickson EL, Xia Q, Zhang Y, Wang T *et al* (2006). Disruption of the operon encoding Ehb hydrogenase limits anabolic CO_2 assimilation in the archaeon *Methanococcus maripaludis*. *Journal of Bacteriology* **188:** 1373-1380.

Poret-Peterson AT, Romaniello SJ, Elser JJ, Shock EL, Hartnett HE, Anbar AD (2013). Active methanogenesis and methanotrophy in a terrestrial hydrothermal ecosystem. *Unpublished, NCBI Entry*.

Poudel S, Tokmina-Lukaszewska M, Colman DR, Refai M, Schut GJ, King PW *et al* (2016). Unification of [FeFe]-hydrogenases into three structural and functional groups. *Biochimica et Biophysica Acta* **1860:** 1910-1921.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P *et al* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41:** 590-596.

Reysenbach A-L, Liu Y, Banta AB, Beveridge TJ, Kirshtein JS, Schouten S *et al* (2006). A ubiquitous thermoacidophilic archaeon from deep-sea hydrothermal vents. *Nature* **442:** 444-447.

Rye RO, Truesdell AH (1993). The question of recharge to the geysers and hot springs of Yellowstone National Park. US Geological Survey Professional Paper 93-384.

Schafer C, Friedrich B, Lenz O (2013). Novel, oxygen-insensitive group 5 [NiFe]-hydrogenase in *Ralstonia eutropha*. *Applied and Environmental Microbiology* **79:** 5137-5145.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75:** 7537-7541.

Schut GJ, Boyd ES, Peters JW, Adams MWW (2013). The modular respiratory complexes involved in hydrogen and sulfur metabolism by heterotrophic hyperthermophilic archaea and their evolutionary implications. *FEMS Microbiology Reviews* **37:** 182-203.

Schut GJ, Zadvornyy O, Wu C-H, Peters JW, Boyd ES, Adams MWW (2016). The role of geochemistry and energetics in the evolution of modern respiratory complexes from a proton-reducing ancestor. *Biochimica et Biophysica Acta* **1857:** 958-970.

Seemann T (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30:** 2068-2069.

Shafaat HS, Rüdiger O, Ogata H, Lubitz W (2013). [NiFe] hydrogenases: A common active site for hydrogen metabolism under diverse conditions. *Biochimica et Biophysica Acta - Bioenergetics* **1827:** 986-1002.

Shock EL, Canovas P (2010). The potential for abiotic organic synthesis and biosynthesis at seafloor hydrothermal systems. *Geofluids* **10:** 161-192.

Shock EL, Holland M, Meyer-Dombard DA, Amend JP, Osburn GR, Fischer TP (2010). Quantifying inorganic sources of geochemical energy in hydrothermal ecosystems, Yellowstone National Park, USA. *Geochimica et Cosmochimica Acta* **74:** 4005-4043.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W *et al* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7:** 1-6.

Søndergaard D, Pedersen CNS, Greening C (2016). HydDB: A web tool for hydrogenase classification and analysis. *Scientific Reports* **6:** 34212.

Spear JR, Walker JJ, McCollom TM, Pace NR (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 2555-2560.

Stamatakis A (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312-1313.

Stevens TO, McKinley JP (2000). Abiotic controls on H_2 production from basalt - water reactions and implications for aquifer biogeochemistry. *Environmental Science and Technology* **34:** 826-831.

Stieglmeier M, Klingl A, Alves RJE, Rittmann SKMR, Melcher M, Leisch N *et al* (2014). *Nitrososphaera viennensis* gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *International Journal of Systematic and Evolutionary Microbiology* **64:** 2738-2752.

Stookey LL (1970). Ferrozine - a new spectrophotometric reagent for iron. *Analytical Chemistry* **42:** 772-781.

Suzuki S, Kakuta M, Ishida T, Akiyama Y (2014). GHOSTX: An improved sequence homology search algorithm using a query suffix array and a database suffix array. *PLOS One* **9:** e103833.

Telling J, Boyd ES, Bone N, Jones EL, Tranter M, MacFarlane JW *et al* (2015). Rock comminution as a source of hydrogen for subglacial ecosystems. *Nature Geoscience* **8:** 851-855.

Thauer RK (1998). Biochemistry of methanogenesis: a tribute to Marjory Stephenson. *Microbiology* **144:** 2377-2406.

Thauer RK, Jungermann K, Decker K (1977). Energy conservation in chemotrophic anaerobic bacteria. *Bacteriological Reviews* **41:** 100-180.

Thauer RK, Kaster AK, Goenrich M, Schick M, Hiromoto T, Shima S (2010). Hydrogenases from methanogenic archaea, nickel, a novel cofactor, and H_2 storage. *Annual Review of Biochemistry* **79:** 507-536.

Therien JB, Artz JH, Poudel S, Hamilton TL, Liu Z, Noone SM *et al* (2017). The physiological functions and structural determinants of catalytic bias in the [FeFe]-hydrogenases CpI and CpII of *Clostridium pasteurianum* strain W5. *Frontiers in Microbiology* **8:** 1-11.

Tian F, Toon OB, Pavlov AA, De Sterck H (2005). A hydrogen-rich early Earth atmosphere. *Science* **308:** 1014-1017.

Toshchakov SV, Korzhenkov AA, Samarov NI, Mazunin IO, Mozhey OI, Shmyr IS *et al* (2015). Complete genome sequence of and proposal of *Thermofilum uzonense* sp. nov. a novel hyperthermophilic crenarchaeon and emended description of the genus *Thermofilum*. *Standards in Genomic Sciences* **10**.

Trembley P, Lovley DR (2012). Role of the NiFe hydrogenase Hya in oxidative stress defense in *Geobacter sulfurreducens*. *Journal of Bacteriology* **194:** 2248-2253.

Truesdell AH, Fournier RO (1976). Conditions in the deeper parts of hot spring systems of Yellowstone National Park, Wyoming. US Geological Survey Open-File Report 76-428.

Truesdell AH, Nathenson M, Rye RO (1977). The effects of subsurface boiling and diultion on the isotopic compositions of Yellowstone thermal waters. *Journal of Geophysical Research* **82:** 3694-3704.

Urschel MR, Kubo MD, Hoehler TM, Peters JW, Boyd ES (2015). Carbon source preference in chemosynthetic hot spring communities. *Applied and Environmental Microbiology* **81:** 3834-3847.

Valentine DL (2007). Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nature Reviews Microbiology* **5:** 316-323.

van Geldern R, Barth JAC (2012). Optimization of instrument setup and post-run corrections for oxygen and hydrogen stable isotope measurements of water by isotope ratio infrared spectroscopy (IRIS). *Limnology and Oceanography: Methods* **10:** 1024-1036.

van Geldern R, Barth JAC (2012). Optimization of instrument setup and post-run corrections for oxygen and hydrogen stable isotope measurements of water by isotope ratio infrared spectroscopy (IRIS). *Limnology and Oceanography: Methods* **10:** 1024-1036.

Vignais PM, Billoud B (2007). Occurence, classification, and biological function of hydrogenases: an overview. *Chemical Reviews* **107:** 4206-4272.

Vignais PM, Billoud B, Meyer J (2001). Classification and phylogeny of hydrogenases. *FEMS Microbiology Reviews* **25:** 455-501.

Vignais PM, Colbeau A (2004). Molecular Biology of Microbial Hydrogenases. *Current Issues in Molecular Biology* **6:** 159-188.

Wächtershäuser G (1990). Evolution of the first metabolic cycles. *Proceedings of the National Academy of Sciences of the United States of America* **87:** 200-204.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73:** 5261-5267.

Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S *et al* (2016). The physiology and habitat of the last universal common ancestor. *Nature Microbiology* **1:** 1-8.

White DE, Muffler LJP, Truesdell AH (1971). Vapor-dominated hydrothermal systems compared with hot-water systems. *Economic Geology: Bulletin of the Society of Economic Geologists* **66:** 75-97.

Wilson CL, Hinman NW, Cooper WJ, Brown CF (2000). Hydrogen peroxide cycling in surface geothermal waters of Yellowstone National Park. *Environmental Science and Technology* **34:** 2655-2662.

Windman T, Zolotova N, Schwander R, Shock EL (2007). Formate as an energy source for microbial metabolism in chemosynthetic zones of hydrothermal ecosystems. *Astrobiology* **7:** 873-890.

Wirth R, Chertkov O, Held B, Lapidus A, Nolan M, Lucas S *et al* (2011). Complete genome sequence of *Desulfurococcus mucosus* type strain (O7/1t). *Standards in Genomic Sciences* **4:** 173-182.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL (2012). Primer-BLAST: a tool to design target specific primers for polymerase chain reaction. *BMC Bioinformatics* **13:** 1-11.

Yeatts DS (2006). Characteristics of thermal springs and the shallow ground-water system at Hot Springs National Park, Arkansas. U.S. Geological Survey Scientific Investigations Report 2006-5001: Reston, Virginia.

APPENDICES

APPENDIX A

MIXING OF METEORIC AND GEOTHERMAL FLUIDS SUPPORTS HYPERDIVERSE CHEMOSYNTHETIC HYDROTHERMAL COMMUNITIES

Manuscript Information

Daniel R. Colman, Melody R. Lindsay, Eric S. Boyd*

*Corresponding Author

Nature Communications

Status of Manuscript:
____ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-reviewed journal
___ Accepted by a peer-reviewed journal
_X_ Published in a peer-reviewed journal

Published February 2019, doi:10.1038/s41467-019-08499-1

Mixing of meteoric and geothermal fluids supports hyperdiverse chemosynthetic hydrothermal communities

Daniel R. Colman[1], Melody R. Lindsay[1], and Eric S. Boyd[1,2*]

[1]*Department of Microbiology and Immunology, Montana State University, Bozeman, MT, 59718*

[2]*NASA Astrobiology Institute, Ames Research Center, Mountain View, CA*

[*]Author of correspondence: Eric S. Boyd (*eboyd@montana.edu*)
Department of Microbiology & Immunology
Montana State University
PO Box 173520
Bozeman, MT 59717
Phone: (406) 994-7046
Fax: (406) 994-4926

Abstract

Little is known of how mixing of meteoric and geothermal fluids supports biodiversity in non-photosynthetic ecosystems. Here, we use metagenomic sequencing to investigate a chemosynthetic microbial community in a hot spring (SJ3) of Yellowstone National Park that exhibits geochemistry consistent with mixing of a reduced volcanic gas-influenced end member with an oxidized near-surface meteoric end member. SJ3 hosts an exceptionally diverse community with representatives from ~50% of known higher-order archaeal and bacterial lineages, including several divergent deep-branching lineages. A comparison of functional potential with other available chemosynthetic community metagenomes reveals similarly high diversity and functional potentials (i.e., incorporation of electron donors supplied by volcanic gases) in springs sourced by mixed fluids. Further, numerous closely-related SJ3 populations harbor differentiated metabolisms that may function to minimize niche overlap, further increasing endemic diversity. We suggest that dynamic mixing of waters generated by subsurface and near-surface geological processes may play a key role in the generation and maintenance of chemosynthetic biodiversity in hydrothermal and other similar environments

Introduction

Mixing of subsurface- and surface-derived fluids defines Earth's habitable (or critical) zone, which represents the outermost layers of the continental crust that are affected by atmospheric, geological, and hydrological processes. The most studied portions of the critical zone are in the near-surface (vadose zone) that are heavily

influenced by photosynthetic productivity. However, the critical zone extends empirically to depths up to ~5 km (Moser et al 2005) and could extend to depths exceeding 12 km (Kieft 2016) where influences from photosynthesis are likely to be minimal and where life is likely constrained by a combination of high temperature and nutrient or energy limitation (Hoehler 2004, LaRowe et al 2017). Numerous studies have documented the presence of microbial communities in subsurface environments (Moser et al 2005, Stetter et al 1993, Stevens and Mckinley 1995, Takai et al 2001); however, little is known of how the confluence of subsurface and surface-associated processes promotes the generation and maintenance of biodiversity in subsurface, chemosynthetic systems. This is due largely to difficulties in accessing these ecosystems and isolating fluid types that source these environments.

Continental hydrothermal systems integrate waters from the deep subsurface and surface to generate extraordinarily diverse geochemical compositions in accessible terrains (Shock et al 2010). At temperatures > ~70°C in circumneutral to alkaline (pH > 6.5) springs or > 54°C in acidic (pH < 3.5) springs, photosynthesis is excluded and microbial life is supported by chemosynthesis (Boyd et al 2012). At first order, extremes in temperature constrain microbial diversity in chemosynthetic hot springs (Colman et al 2018, Miller et al 2009). However, studies conducted in thermal springs have revealed chemosynthetic communities that are exceptionally diverse (Barns et al 1994, Hugenholtz et al 1998, Inskeep et al 2013), while others have revealed communities that comprise as few as one or two populations, despite similar spring temperatures (Inskeep et al 2013).

These discrepancies suggest geochemical controls on the generation and maintenance of chemosynthetic hot spring community biodiversity.

Thermal springs have served as integral platforms for examining the role of geochemical variation in shaping microbial diversity and have provided new insights into novel forms of microbial life and the processes that sustain this life (Inskeep et al 2013, Meyer-Dombard et al 2005, Spear et al 2005). Indeed, some of the earliest environmental studies of uncultivated microbial diversity were conducted in thermal springs like Obsidian Pool, Yellowstone National Park (YNP), Wyoming, leading to the recognition of dozens of new archaeal and bacterial taxonomic divisions from this single spring (Barns et al 1994, Hugenholtz et al 1998). Likewise, recent environmental genomic surveys from YNP hot spring communities have yielded new insights into the genomic and physiological diversity of uncultivated lineages including those thought to represent a missing link to proto-Eukaryotes (Zaremba-Niedzwiedzka et al 2017), those that serve as analogs to cosmopolitan subsurface taxa (Baker et al 2016), and have provided insights into some of the smallest known microbial symbionts (Wurch et al 2016). Many of these newly described lineages contribute to our understanding of the physiological limits of life on Earth (Colman et al 2018, Valentine 2007) and have served as analogs for understanding the capacity for life in the deep subsurface of Earth (Colman et al 2017). Yet, little investigation has been conducted to discern how the generation and maintenance of such diversity is related to the geological processes that influence geochemical variation in hydrothermal systems. In other words, considerable research has attempted to identify what types of diversity are present in hydrothermal systems, but

little attention has been directed towards understanding how the diversity arises and how it is differentially promoted across spring types.

The geochemistry of hot spring waters is controlled by interaction between subsurface and surface geologic processes (Fournier 1989). Paramount among these is separation of waters into a gas-poor, circumneutral liquid phase that is enriched in non-volatile solutes (e.g., Cl^-) and a gas-rich vapor phase that is enriched in reduced gases such as H_2S, CH_4, and H_2 (Bergfeld D. et al 2014, Fournier 1989, Nordstrom et al 2009). Gas-rich vapor can condense and mix with O_2-rich meteoric waters as it ascends to the surface, promoting the oxidation of H_2S and the subsequent generation of acidic springs (Nordstrom et al 2009, Nordstrom et al 2005). Broadly these two geothermal fluids result in two geothermal end member water types with circumneutral (pH ~6.5-7.5) and acidic (pH ~2.5-3.5) pH (Nordstrom et al 2009, Nordstrom et al 2005). These two geothermal water types, along with precipitation-derived or meteoric waters, represent three end member water compositions that broadly define hot spring geochemistry (Nordstrom et al 2009). The microbial communities associated with these end member compositions have been characterized extensively (Amenabar et al 2015). However, dynamic mixing of these end members can generate spring waters that are intermediate in composition between the end member waters described above (i.e., pH ~4.0 ~ 6.0). Far less is known about the microbial communities inhabiting springs sourced by mixed end member waters.

Here, we describe a hyperdiverse chemosynthetic thermal spring community from YNP termed 'Smokejumper 3' (SJ3; pH=5.4, T=61.9°C), which hosts an extensive diversity of archaeal and bacterial lineages. The geochemical composition of SJ3 waters represent an extreme example of mixing of a reduced vapor phase-influenced water enriched in volcanic gases with oxidized near-surface meteoric water. Many of the lineages in SJ3 have not previously been characterized via genomic analyses, nor have several identified functionalities (e.g., methanogenesis) been previously identified in the lineages that are present within the SJ3 community. Comparison of the SJ3 metagenome with other available YNP chemosynthetic metagenomes indicates that spring waters with geochemical compositions underpinned by similar geological processes also host exceptionally diverse microbial communities that are functionally similar to that of SJ3. The results are discussed in the context of how dynamic spatial and temporal mixing of water compositions likely support and maintain biodiverse chemosynthetic communities in hydrothermal settings.

Results

Geologic and geochemical context of Smokejumper geyser basin.

The Smokejumper geyser basin (SJGB) is located in Southwestern Yellowstone National Park (YNP), Wyoming, at a topographic high in the Summit Lake Rhyolite (lava) Flow (Fig. 1). The SJGB is also located at high elevation, and near the continental hydrological divide of the Americas and on the edge of the Yellowstone caldera boundary (Fig. 1), where extensive faulting and fracturing of volcanic bedrock is thought to promote the release of gases that were exsolved during subsurface phase separation

(Christiansen 2001). SJ3 is a gas-rich, hydrothermal spring located within the SJGB (N 44°24'57.42"; W -110°57'20.76"). SJ3 waters exhibit low conductivity (0.42 mS, Supplementary Figure 1) and a pH of 5.4 (Supplementary Figure 2), characteristics that have been suggested to result from mixing of vapor phase gas-influenced waters with meteoric waters (Amenabar et al 2015, Boyd et al 2010). Moreover, the very low chloride concentrations measured in SJGB springs (~1.0 – 1.1 mg L^{-1}) (Bergfeld et al 2014) suggest minimal to no input of waters from the deep hydrothermal reservoir (Fournier 1989, Nordstrom et al 2009), while SO_4^{2-} concentrations are elevated (~54-70 mg L^{-1}; Supplementary Figure 3) relative to meteoric waters (Bergfeld et al. 2014), likely due to the oxidation of vapor phase-sourced sulfide (Supplementary Figure 3**)**. Taken together, the low conductivity, slightly acidic pH, low Cl^-, and moderate levels of SO_4^{2-} suggests low residence time of meteoric waters. Phase separated volcanic gas-influenced geothermal fluids are often enriched in H_2, CO, and short chain alkanes due to interaction between geothermal fluids and ferrous iron containing minerals in bedrock during ascension of the gases and waters to the surface (Bergfeld et al 2014). Consistent with this observation, dissolved H_2 and CH_4 concentrations in springs of the Smokejumper geyser basin are among the highest concentrations reported in hot spring waters in YNP (Bergfeld et al. 2014).

<u>High taxonomic and functional diversity of the SJ3 community.</u>

Assembly of contigs from the SJ3 metagenome (~61 Gbp of reads, 278 Mbp assembled length, Supplementary Table 1) and subsequent binning of genomes into metagenome assembled genomes (MAGs) based on tetranucleotide frequencies and

coverage profiles resulted in 108 draft MAGs. Of these 108 MAGs, 82 were estimated to be > 50% complete and exhibited < 7% contamination, meeting currently accepted criteria as medium to high quality draft MAGs (Bowers et al 2017) (Fig. 2**,** Supplementary Data 9). Populations closely related to the chemolithoautotrophic $H_2S/S_2O_3^-/S^0$ oxidizing *Sulfurihydrogenibium yellowstonense* (phylum: Aquificae) and chemoheterotrophic $S^0/S_2O_3^-/SO_3^{2-}$ reducing *Caldisericum exile*(Mori et al 2009) (phylum: Caldiserica; formerly candidate division OP5) dominate the community (~15% relative abundance each; Supplementary Data 9), which is consistent with the prevalence of *Sulfurihydrogenibium* populations in weakly acidic YNP hot springs (Mitchell 2009). The remainder of the genomes comprised diverse low abundance populations with 10 and 23 archaeal order/phylum-level and bacterial phylum-level groups present, respectively, and with multiple representatives within many of the groups (Fig. 3, Supplementary Data 9).

Populations within the SJ3 community represented ~50% of the known higher-order lineages within both Archaea (order/class level) and Bacteria (phylum level) (Fig. 3). This level of diversity is particularly high for chemosynthetic hydrothermal system sediment/water communities that have previously been shown to typically comprise only a few lineages (Colman 2015, Colman et al 2016, Inskeep et al 2013). Further, SJ3 genomes contributed 8% of the branch length to both archaeal and bacterial phylogenetic trees inclusive of all the major lineages for each domain. While the total branch length and taxonomic breadth within a phylogeny is dependent on *i*) the taxa that are included in the analysis, *ii*) lineage representation in available databases, and *iii*) lineage definitions

that can be controversial (Hug et al 2016, Parks et al 2018), these analyses nevertheless suggest that this hot spring community is hyperdiverse at the taxonomic and phylogenetic levels when compared to previously analyzed hot spring communities.

A number of the lineages observed in SJ3 are underrepresented in other hot springs (Boyd et al 2013, Colman 2015, Colman et al 2016, Inskeep et al 2013, Mitchell 2009, Xie et al 2014), including many that have only rarely been observed in YNP hot springs such as the *Candidatus* (*Ca.*) 'Acetothermia', Archaeoglobales, *Ca.* 'Bathyarchaeota', Deep sea Hydrothermal Vent Euryarchaeota group 2 (DHVE2), DPANN-related lineages, *Ca.* 'Hadesarchaea', Thaumarchaeota, and *Ca.* 'Verstraetarchaeota', among others. Phylogenomic reconstructions indicate that several SJ3 MAGs represent the currently deepest-branching genomic representatives of *Ca.* 'Aigarchaeota', Korarchaeota, Thaumarchaeota, *Ca.* 'Bathyarchaeota', *Ca.* 'Verstraetarchaeota', and the DHVE2 (Supplementary Data 1), suggesting that they may provide new insight into the divergence of these archaeal groups. Indeed, several of these deep-branching lineages belong to monophyletic clades that are phylogenetically distant from other characterized members of their respective lineages (e.g., as in the Thaumarchaeota, Korarchaeota, and *Ca.* 'Aigarchaeota'), and thus may comprise lineages distinct from those indicated in Fig. 3.

<u>End member fluid mixing and chemosynthetic diversity.</u>

To investigate whether the functional coding potential within the SJ3 metagenome was high among chemosynthetic YNP hot spring communities in general, protein family

diversity was compared among SJ3 and publicly available YNP chemosynthetic spring community metagenomes. Rarefaction analysis and random subsampling of protein families in chemosynthetic communities from SJ3 and other publicly available YNP hot spring communities (n=48; Supplementary Data 10) revealed that the SJ3 community comprises one of the highest levels of protein diversity among metagenomically characterized chemosynthetic YNP spring communities (Fig. 4). This result held with increasing subsampling depth (Fig. 4) indicating that the deep sequencing of SJ3 (298,127 proteins in the analysis) did not inflate protein family diversity relative to less-thoroughly characterized YNP metagenomes (e.g., those produced with Sanger sequencing). Indeed, several metagenomes generated from less total sequence when compared to SJ3 yielded similarly high levels of protein diversity even when compared at the level of only 10,000 subsampled proteins (Fig. 5a; Supplementary Data 10). Two of these metagenomes, 'Obsidian Pool Prime', and 'Washburn Spring', were produced with Sanger sequencing (Inskeep et al 2013) and comprised > 2 orders of magnitude less assembled sequence data than SJ3, suggesting that sequencing effort, per se, did not preclude comparisons of overall protein diversity when selecting random subsets of rarefied proteins from each metagenome.

A significant linear relationship (adjusted $R^2 = 0.33$, $p < 0.001$, n=41; Fig. 5) was observed when comparing protein coding gene diversity (after subsampling at 10,000 random proteins) against the ratio of SO_4^{2-}/Cl^- of waters in springs where these data were available (Supplementary Data 10). Spring waters with elevated SO_4^{2-}/Cl^- ratios, such as SJ3, are interpreted to reflect mixing of vapor phase (contributing elevated SO_4^{2-} from

H_2S oxidation) with near surface meteoric waters that lack significant Cl^- input from the deeply-sourced hydrothermal reservoir (Nordstrom et al 2009). Conversely, spring waters with low SO_4^{2-}/Cl^- ratios are indicative of less input of vapor phase and are interpreted to reflect input from the deep hydrothermal reservoir (Markusson and Stefansson 2011, Nordstrom et al 2009). The large amount of variation unaccounted for in the linear model can be attributed to the role of other variables and processes that are unaccounted for in this model that may influence the functional diversity of spring communities. This includes temperature and the temporal dynamics of the spring, both of which have been suggested to influence the biodiversity of thermal springs (Cole et al 2013, Colman et al 2016, Costa et al 2009, Miller et al 2009). Regardless, the prevalence of higher protein coding gene diversity in communities inhabiting springs that are sourced by mixing of end member fluids indicates that this process promotes functional diversity in thermal springs, and likely also contributes to the elevated taxonomic diversity observed in these springs (Barns et al 1994, Hugenholtz et al 1998). Consistent with this interpretation, several of the communities that encoded higher protein coding gene diversity than SJ3, and which had published geochemical information on spring waters (i.e., 'Washburn Spring', 'Obsidian Pool Prime', and Obsidian Pool (*sensu stricto*)), all shared similar attributes with SJ3 including hosting waters with similar pH, high SO_4^{2-}, and low Cl^- concentrations (Fig. 5**,** Supplementary Data 10). This observation suggests that the geologic processes that define the geochemical composition of these spring types (i.e., end member water mixing) is potentially also involved in supporting and maintaining the high levels of taxonomic and functional biodiversity in these springs.

While the role of geothermal fluid mixing in supporting community diversity has not been extensively investigated in continental hydrothermal systems, several investigations of marine hydrothermal vent systems suggest that fluid mixing influences the diversity of vent communities (Huber et al 2002, Huber et al 2003, Nakagawa et al 2005, Schrenk et al 2003). Indeed, mixing of high temperature hydrothermal fluids (>200^{o}C) with cold seawater promotes chemical disequilibrium in redox reactions that can support microbial metabolism (McCollom and Shock 1997), likely leading to increased niche space capable of supporting higher diversity. Although there is considerable evidence that mixing of hydrothermal fluids with seawater influences community diversity, to the best of our knowledge, an explicit comparison of marine hydrothermal vent community diversity levels with chemical proxies that can be used to deduce the extent of fluid mixing has not been conducted. Nevertheless, recent metagenomic analyses have revealed high levels of metabolic and taxonomic diversity in marine sediments within environments hosting fluids that are reflective of considerable hydrothermal fluid/seawater mixing at the Juan de Fuca Ridge Flank (Jungbluth et al 2017), as well as in the Guaymas Basin (Dombrowski et al 2017). Consequently, it is likely that mixing of end member fluids leads to similar ecological outcomes in both marine and continental hydrothermal systems.

Subsurface-sourced gases support SJ3 community metabolism.

To compare SJ3 functional potential to other chemosynthetic YNP metagenomes and investigate the links between geothermal fluid mixing and SJ3 metabolic diversity, the metabolic protein-coding differences between SJ3 and a subset (n=14) of available

YNP metagenomes were compared. The encoded protein complements involved in energy metabolism in SJ3 populations were like those from 'Obsidian Pool Prime' (OPP_17) and 'Washburn Spring' (WS_18). Together, these communities formed a distinct cluster in a PCoA ordination of inferred protein composition dissimilarities (Fig. 6a). The encoded proteins that separated these three springs (particularly SJ3) from other springs analyzed included [NiFe]-hydrogenases (F_{420} hydrogenase, Frh; NADP-coupled hydrogenase, Hyh; bidirectional hydrogenase, Hox), carbon monoxide dehydrogenase (Coo), carbonic anhydrase (Cah), formate dehydrogenase (Fdh), and those associated with methanogenesis/anaerobic alkane oxidation (methyl coenzyme-M reductases, Mcr; tetrahydromethanopterin S-methyltransferase, Mtr; formylmethanofuran dehydrogenase, Fwd) (Fig. 6b**,** Supplementary Data 3). In addition, genes encoding homologs of acetate kinases (Ack) and acetate synthases (Acs) putatively involved in acetogenesis (Ragsdale and Pierce 2008) were also among the more enriched protein coding genes in SJ3.

Genes encoding putative Mcr proteins were identified in SJ3 MAGs associated with the *Ca.* 'Verstraetarchaeota' (SJ3.Bin27 and SJ3.Bin82.1.1) and Archaeoglobales (SJ3.Bin34). Mcr protein homologs have been recently identified in MAGs of putatively methanogenic (or alkanotrophic) lineages outside of the canonical euryarchaeotal methanogen groups, including representatives of the *Ca.* 'Bathyarchaeota' and *Ca.* 'Verstraetarchaeota' (Evans et al 2015, Vanwonterghem et al 2016). Unlike previous analyses of *Ca.* ‘Bathyarchaeota’ from coal bed methane wells (Evans et al 2015), no evidence for methanogenesis or methanotrophy was identified in the five phylogenetically distinct (Supplementary Data 1) *Ca.* ‘Bathyarchaeota’ MAGs that were

present in the SJ3 community. Regardless, the enrichment of genes coding for proteins putatively involved in the metabolism of substrates that are abundant in volcanic gases (e.g., H_2, CH_4 or other alkanes, CO), and that are particularly abundant in SJ3 (Bergfeld et al. 2014), is consistent with the SJ3 community being adapted to take advantage of these substrates as electron donors and/or carbon sources. Many of the genes coding for the above-mentioned functionalities were also enriched in the 'Obsidian Pool Prime' and 'Washburn Spring' metagenomes (Fig. 6b, Supplementary Data 3), which are likely to be sourced by vapor phase-influenced waters (and/or subsurface-sourced gases) mixing with meteoric waters, based on similar geochemical profiles as SJ3 (Inskeep et al 2013, Shock et al 2005) (Fig. 5, Supplementary Data 10).

Metabolic differentiation also promotes high diversity.

The detection of closely related genome bins within the diverse SJ3 community prompted investigation of putative ecological mechanisms that may contribute to the maintenance of such high levels of diversity in mixed fluid systems and whether such diversity could be related to the process of end member fluid mixing. Specifically, we sought to determine whether metabolic differentiation of taxa (i.e., the minimization of niche overlap) belonging to the same higher-order taxonomic groups allows for population coexistence. Comparison of archaeal and bacterial populations belonging to the same lineages (and exhibiting > 75% completeness) were conducted for ten archaeal taxonomic groups (*Ca.* 'Aigarchaeota', Archaeoglobales, *Ca.* 'Bathyarchaeota', Desulfurococcales, Korarchaeota, Thaumarchaeota, Thermoplasmata, Thermoproteales, and *Ca.* 'Verstraetarchaeota'), and seven bacterial groups (Acidobacteria, *Ca.*

‘Aminicenantes’, Deltaproteobacteria, Dictyoglomi, Spirochaetes, Thermodesulfobacteria, and Dictyoglomi-like unclassified Bacteria) (Supplementary Data 9). Comparisons of G+C% content, coverage profiles, and phylogenetic distances among MAGs indicated that they represent distinct populations and were not artefacts of the genome binning approaches used here (Supplementary Data 4**,** Supplementary Data 5).

We first hypothesized that closely related MAGs that are differentiated due to the loss or acquisition of traits should exhibit a greater level of difference in the composition of metabolism-related protein coding genes relative to genetic processing protein coding genes. To test this hypothesis, we first compared differentiation in encoded proteins among genomes that were annotated within the KEGG category of 'metabolism' (which primarily comprises proteins involved in energy conservation) versus differentiation of encoded proteins among genomes involved in 'genetic information processing'. A greater extent of differentiation of metabolism-related proteins relative to genetic processing proteins was observed in all but one (Dictyoglomi SJ3.Bin31.1.2 vs. SJ3.Bin76.1.7) of the 49 pairwise comparisons, suggesting that metabolic differentiation was ubiquitous among SJ3 populations that belonged to the same taxonomic lineage (Fig. 7a).

Further comparison of the functional potentials among MAGs within the same taxonomic group indicates that closely related taxa are differentiated at the metabolic level, potentially allowing for their co-existence and the maintenance of high levels of diversity in SJ3 (Fig. 7). For example, a single Archaeoglobales-related MAG

(SJ3.Bin34; 98.7% estimated completeness) with the putative capacity for methanogenesis, methanotrophy, or higher chain alkane oxidation (Supplementary Figure 4, Supplementary Data 6**)**, co-occurred with other Archaeoglobales (SJ3.Bin61; 89.5% estimated completeness) that exhibited metabolic potential that was more typical of other Archaeoglobales (i.e., sulfate/sulfite reduction/thiosulfate reduction (Garrity and Holt 2001); Fig. 7c; Supplementary Data 6). The SJ3.Bin34 and SJ3.Bin61 MAGs are closely-related phylogenetically, and encode proteomes that are 72% $\pm$ 13% identical at the amino acid level, suggesting that they likely belong to the same family- or genus-level clade (Konstantinidis and Tiedje 2005). SJ3.Bin34 forms an outgroup to SJ3.Bin61 and another unpublished MAG from a Chinese hot spring (*Archaeoglobus* sp. JZ bin_24; IMG genome ID: 2721755890). However, genes coding for proteins involved in methanogenesis or anaerobic alkane oxidation (e.g., McrAGCDE), in addition to numerous other methanogenesis-related accessory proteins, are only present in the SJ3.Bin34 MAG (Fig. 7c; Supplementary Figures 4 and 5, Supplementary Data 6). Conversely, genes coding for key proteins involved in sulfate/sulfite/thiosulfate reduction (AprAB, DsrAB) are present in the SJ3.Bin61 MAG, but not that of SJ3.Bin34.

Intriguingly, the SJ3.Bin34 McrA homolog forms a basal lineage to *Ca.* 'Verstraetarchaeota' McrA homologs identified here and elsewhere (Vanwonterghem et al 2016) (Supplementary Figure 6). Given that the SJ3.Bin34/SJ3.Bin61/JZ bin_24 clade is nested within other well-characterized sulfate/sulfite/thiosulfate-reducing Archaeoglobales (Fig. 7c), the genetic capacity for putative methane or alkane metabolism was either horizontally acquired in the SJ3.Bin34 lineage from a yet to be

identified source, or otherwise retained as an ancestral trait that was lost in other Archaeoglobales. Methylotrophic methanogenesis has been hypothesized in the *Ca.* 'Verstraetarchaeota'(Vanwonterghem et al 2016), although the lack of cultivation information precludes definitive assignment of this functionality. However, the phylogenetic association of the Archaeoglobales/ *Ca.* 'Verstraetarchaeota' McrA homologs with McrA homologs from hydrogenotrophic methanogens such as *Methanopyrus*, Methanobacteriales, and Methanococcales (Garcia et al 2000) (Supplementary Figure 6**)**, leads us to cautiously speculate that the SJ3 Archaeoglobales population represented by the SJ3.Bin34 MAG may be involved in hydrogenotrophic methanogenesis. This assertion is consistent with the lack of protein homologs necessary for methylotrophic methanogenesis (Supplementary Data 6), and the presence of protein homologs necessary for hydrogenotrophic methanogenesis (Supplementary Figure 4**)**. Nevertheless, here and in other comparisons, it is possible that the lack of genome completeness could lead to the apparent absence of some functional gene homologs. However, this is unlikely for the SJ3.Bin34 MAG owing to its high estimated completeness (~99%) and other lines of evidence suggesting the potential for hydrogenotrophic methanogenesis. The capacity to produce small amounts of CH_4 *in vitro* has been previously documented for *Archaeoglobus fulgidus* (Achenbach-Richter et al 1987), despite that Mcr protein coding genes have not been previously observed in the genome of *A. fulgidus* or other published Archaeoglobales isolates or genomes. Thus, while other Archaeoglobales may be able to generate minor amounts of CH_4 *in vitro*, it must necessarily be via a mechanism unlike that of the energy conservation pathway in methanogens and potentially also the population represented by SJ3.Bin34. Regardless,

this apparent transfer or retention of ancestral Mcr homologs in the Archaeoglobales lineage has likely allowed it to inhabit an available niche whose dimensions are defined by variable input of volcanically sourced gases such as H_2, CO_2, CH_4 or short-chain alkanes.

Potential metabolic differentiation was also observed in *Ca.* 'Verstraetarchaeota' MAGs (n=4) wherein genes encoding McrABG were identified in later diverging *Ca.* 'Verstraetarchaeota' MAGs (SJ3.Bin27 and SJ3.Bin82.1.1; Supplementary Figure 7) but were absent in earlier-diverging lineages represented by the MAGs SJ3.Bin21 and SJ3.Bin40 (Supplementary Data 7). The only published evidence of functional potential in the *Ca.* 'Verstraetarchaeota' derives from MAGs recovered from anaerobic digesters (Vanwonterghem et al 2016), wherein these populations were hypothesized to be capable of methylotrophic methanogenesis due to the presence of protein coding genes necessary to utilize methanol (MtaA; [methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase), methylamines (MtbA; [methyl-Co(III) methylamine-specific corrinoid protein]:coenzyme M methyltransferase, MtmB; methylamine-corrinoid protein Co-methyltransferase, MtbC; dimethylamine corrinoid protein), or methanethiols (MtsA; methylated-thiol-coenzyme M methyltransferase). Methylotrophy-associated homologs were present in SJ3.Bin27 (MtaA, MttC; trimethylamine corrinoid protein, MtmB) and SJ3.Bin82.1.1 (MtaA, MttC, MtmB, MtbC) (Supplementary Data 7), suggesting a similar metabolic potential as those described previously. In contrast, none of the above homologs were identified in SJ3.Bin21 or SJ3.Bin40, or other closely related, earlier-diverging *Ca.* 'Verstraetarchaeota' (Fig. 7e).

This suggests relatively recent metabolic differentiation of the SJ3 *Ca.* 'Verstraetarchaeota' populations, potentially allowing them to capitalize on available CH_4, short-chain alkanes, or methylated compounds in SJ3 (Fig. 7e). Moreover, the capacity for carbon monoxide (CooS) or formate (FdhF) utilization was variously distributed within SJ3 *Ca.* 'Verstraetarchaeota' population MAGs, in addition to the metabolism of hydrogen through a variety of [NiFe]-hydrogenase isoforms encoded in these genomes (Peters et al 2015), further indicating widely disparate physiological potential among the SJ3 *Ca.* 'Verstraetarchaeota' populations (Fig. 7e; Supplementary Data 7).

In addition to the two lineages discussed above, three phylogenetically distinct, deeply-branching Thaumarchaeota-related MAGs were recovered from SJ3 with > 75% estimated completeness (and two others > 50% complete). None of these MAGs encoded proteins allowing for the oxidation of ammonium (Supplementary Data 8), a trait that has thus far typified many Thaumarchaeota cultivars and genomes (Brochier-Armanet et al 2012). Rather, these deeper-branching SJ3 thaumarchaeotes (SJ3.Bin56, SJ3.Bin70) harbored the functional capacity to consume or produce H_2 via the activity of [NiFe]-hydrogenase isoforms predicted to be involved in reversible H_2 transformation (e.g., group 3d and 4 enzymes; Fig. 7g; Supplementary Data 8) (Peters et al 2015). Conversely, genes coding for carbon monoxide dehydrogenase (CoxL) and thiosulfate sulfurtransferase (Tst) suggests the ability to use CO and $S_2O_3^-$ as electron donors, respectively, in the population represented by the SJ3.Bin66 MAG (and other later-diverging thaumarchaeotes). Genes encoding these proteins were absent in the deeper-

branching thaumarchaeote MAGs. This observation coupled with the presence of NADH dehydrogenase (Nuo) subunit homologs in later evolving thaumarchaeotes that were absent in the earlier-evolving genomes suggests an overall shift in mechanisms allowing for respiration in these lineages (Schut et al 2016). This may represent a transition from anaerobic metabolisms of earlier-diverging Thaumarchaeota to aerobic metabolisms within later-evolving Thaumarchaeota. The presence of genes encoding proteins in the earlier-evolving thaumarchaeote MAGs that are involved in methanogenesis including HdrABC (heterodisulfide reductase) and MttB (trimethylamine methyltransferase) subunits (which are involved in methylotrophic methanogenesis in the Methanomassiliicoccales, and potentially in the *Ca.* 'Verstraetarchaeota' (Poulsen et al 2013, Vanwonterghem et al 2016)) further highlight these evolutionary shifts. However, the lack of other necessary protein complements for methanogenesis (e.g. Mcr, Mtr, Fwd) in these two thaumarchaeote MAGs (inclusive of SJ3.Bin56) suggests that they are unlikely to be involved in methylotrophic methanogenesis (Supplementary Data 8).

These observations, taken together, suggest that phylogenetically similar populations within SJ3 are metabolically differentiated in a manner that is likely related to selective pressure to diversify into new niches made available by the distinctive environment of SJ3 and other similar hydrothermal environments. Thus, metabolic differentiation and subsequent habitation of environmental niches created by dynamic mixing of volcanic gas-enriched hydrothermal fluids and meteoric waters may also help explain the high level of diversity that is present in SJ3 relative to other chemosynthetic hot spring communities. The data presented here suggests that much of this

differentiation can be linked to diversification to take advantage of lithogenically-sourced compounds in SJ3 waters made available by variable mixing of end member geothermal fluid types that characterize hot spring environments.

Discussion

Hydrothermal springs with pH values between 4.0-6.0 are understudied microbial ecosystems relative to other hydrothermal spring types. The pH of these springs reflects underlying geothermal fluid mixing processes wherein volcanic gas phase-influenced end member waters mix with near-surface sourced meteoric waters and/or deeply-sourced hydrothermal fluids, resulting in weakly acidic spring waters. We suggest such conditions promote and maintain extensive disequilibria in oxidation-reduction reactions that support distinctively high chemosynthetic microbial biodiversity relative to other springs that reflect minimal or no mixing of fluid types. Although little studied in continental hydrothermal systems, considerable evidence exists for the influential role of hydrothermal and seawater mixing on deep-sea vent microbial diversity. Chemosynthetic communities within intermediate pH springs across YNP are similar in microbial taxonomic and functional compositions suggesting that such water mixing regimes generally selects for and sustains unique and exceptionally diverse microbial chemosynthetic communities. Several lineages of microorganisms with inferred metabolisms not observed previously in those lineages suggests that such conditions might also even promote the emergence of new biodiversity. This metabolic differentiation functions to minimize niche overlap while maximizing the use of available nutrients (particularly gases), further promoting increased biodiversity in these settings.

Taken together, these results provide new understanding of how geologic settings lead to fluid mixing dynamics that then sustains high chemosynthetic community biodiversity and may also lead to the generation of new chemosynthetic biodiversity. Moreover, these data provide important context for understanding the role of subsurface and near-surface water mixing in promoting chemosynthetic biodiversity in geothermal systems, and other similar environments, including the subsurface of Earth today and in pre-photosynthetic ecosystems of early Earth.

Methods

Sampling, metagenomic sequencing, and data processing.

Sediment samples for community analyses were collected from SJ3 spring on July 22, 2014. Spring temperature and pH were measured with a portable pH meter and a temperature-compensated probe (WTW 3100; WTW, Weilheim, Germany). Water conductivity was measured using a temperature-compensated probe (YSI EC300; YSI Inc. Yellow Springs, Ohio). Triplicate sediment samples (~250 mg) were collected sterilely, immediately frozen on dry ice, and transported back to the laboratory on July 22, 2014. DNA was extracted from triplicate sediment samples (~250 mg each) using the Fast DNA Spin Kit for Soil (MP Biomedicals, Irvine, CA) following the manufacturer's instructions as conducted previously (Hamilton et al 2013). Equal volumes of triplicate DNA extracts were then pooled for further analyses.

Whole community shotgun metagenomic sequencing was conducted on total genomic DNA from SJ3 sediments (~5 ng total). Paired-end sequencing (2 x 250 bp) was

conducted at the Genomics Core Facility at the University of Wisconsin-Madison on the Illumina Hiseq 2500 Rapid platform. Fragmented DNA was prepared using the Illumina Nextera DNA library preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer's protocols. Reads were quality trimmed and cleaved of Illumina sequencing adapters using Trimmomatic v.0.35(Bolger et al 2014) and the following parameters: LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, and MINLEN:36. MEGAHIT v.1.1.1(Li et al 2015) was used to assemble the quality-filtered reads into contigs using a range of k-mer values (21, 29, 39, 59, 79, 99, 119, and 141) and default parameters. MetaQUAST v.3.2(Mikheenko et al 2016) was then used to assess the quality of the assemblies and a final assembly was chosen using a k-mer size of 141. Raw reads were aligned and mapped to the final contigs using Bowtie2 (Langmead and Salzberg 2012).

Contigs (> 2.5 kbp) from the highest quality assembly (k-mer size=141) were binned into draft metagenome-assembled-genomes using unsupervised binning in MetaBAT v.0.26.3 (Kang et al 2015) based on tetranucleotide frequency distribution patterns and differential sequence coverage profiles with the "verysensitive" parameter settings. Draft genome bins were then assessed for quality, contamination, and completeness using CheckM v.1.0.5 (Parks et al 2015). Manual curation of bins was conducted using several methods to improve the quality of bins that appeared to represent multiple populations based on marker gene 'contamination' estimates. First, 'outlier' contigs that were defined as outside of 95% of the distributions for tetranucleotide word frequency distance, G+C content, or coding density of each genome bin were removed

using CheckM. Second, genome bins that clearly consisted of multiple populations were re-binned using MetaBAT specifying increasingly more stringent sensitivity parameters, until 'contamination' in each bin was minimized. Contigs in each bin were also surveyed for obvious coverage or G+C% value deviations from the majority of the bin's contigs. Only medium-high quality draft genome bins are included in the analyses presented here (> 50% complete, < 7% contamination).

Gene prediction and annotations were then conducted using Prodigal v.2.6.3 (Hyatt et al 2010) as implemented in Prokka v.1.11 (Seemann 2014) using the default parameters, or, as implemented in CheckM using the default parameters. To compare the robustness of genome bin assignments, additional binning was performed using the unsupervised binning program CONCOCT v.0.4.1 (Alneberg et al 2014), and the quality of the CONCOCT bins were assessed as described above. The MetaBat- and CONCOCT-produced bins were compared against one another to assess the reliability of genome bin differentiation within and between the two methods. In some cases, the completeness of genome bins that were estimated to be abundant (> 1.0% estimated relative abundance) exhibited low completeness, as is common in deep metagenomic sequencing of environmental genomic DNA. Improvement in the quality of these bins was attempted by recruiting contigs to genomes that were publicly available or were available from other in-house metagenomes and shared high marker protein identity to the low-completeness bins. Contig recruitment was conducted with the MG Wrapser tool (https://github.com/dunfieldlab/mg_wrapser) using an 80-90% identity level and 50% alignment length, followed by extraction of quality-filtered reads mapped to the contigs

using the Multi-metagenome package (Albertsen et al 2013). Extracted reads were then reassembled using the Spades v.3.10.0 assembler (Nurk et al 2017), which produced higher quality assemblies than MEGAHIT for individual populations. K-mer sizes of 21,33,55,77,99,127 and the --meta and --only-assembler options were used for assembly. Contigs > 1,000 bp were used for these bins, as the reference-based extraction procedure allowed for higher-quality genome assemblies. Assessment of assembly quality and population homogeneity was conducted with MetaQUAST and CheckM, and final genome bins were curated as described above using RefineM (Parks et al 2018) to identify and filter outlier contigs from bins and to also employ k-means clustering to separate bins with multiple populations on the basis of various genomic properties (e.g., coverage profile, tetranucleotide frequencies). Relative abundances were calculated for individual populations following the methods of Hu *et al.* 2016 (Hu et al 2016), wherein relativized coverage was first calculated for each bin by dividing total mapped reads by total length. This measurement was then normalized to estimated genome size by multiplying relativized coverage by estimated genome size. Finally, the relative abundance for each population was determined as the fraction represented by each population of the total summed coverage values.

Phylogenetic analyses.

Curated genome bins were surveyed for the presence of 104 archaeal-specific or 31 bacterial-specific single copy phylogenetic marker genes with Amphora2 (Wu and Scott 2012). Publicly available references from genome databases (IMG and NCBI) were downloaded based on homology searches of RpoB proteins from each bin (or ribosomal

protein sequences if RpoB was not present) against each database. Only references with > 50% estimated completeness were included in the final trees. Each of the proteins were aligned individually using Clustal Omega v.1.2.0 (Sievers et al 2011), and the protein alignments were concatenated into a super matrix for each domain. The concatenated alignments for Archaea (n=713; 55,322 informative amino acid positions) and Bacteria (n=497; 15,290 informative amino acid positions) were subjected to Maximum Likelihood phylogenetic analysis in RAxML v.8.2.9 (Stamatakis 2014) specifying an LG protein substitution model and a Gamma shape distribution. The robustness of each clade's monophyly was assessed with 100 rapid algorithm bootstraps. Taxonomic clades were annotated based on monophyly and previously published designations. Clades without current genomic representation were identified based on monophyletic groupings and branch-lengths equal-to or greater than those used to define commonly accepted division-level designations. To assess the total branch length associated with genomes from SJ3, the branch lengths of all taxa in the final trees were calculated using the 'distRoot' function in the 'adephylo' package for R with default settings (Jombart and Dray 2008).

Protein family diversity and comparison to other metagenomes.

To assess the diversity of protein diversity and functionalities in SJ3 relative to other YNP metagenomes, proteins > 50 amino acids in length from SJ3 and 47 other YNP metagenomes publicly available in the IMG database as of 10/01/2017 (Supplementary Data 10) were clustered into protein family 'bins' using CD-HIT v.4.6.5 (Fu et al 2012) with the following parameter settings: -g 1 -G 0 -aS 0.8 -d 0.

Communities were identified as 'chemosynthetic' based on previously published information for the specific metagenomes or on the springs that were analyzed. The entire dataset was first clustered at the 90% homology level (k=5), followed by clustering at the 60% homology level (k=4), and a final clustering step at the 30% homology level (-ce 1e-6). Abundance-weighted protein bin counts were then calculated within each metagenome, and this dataset was used in a rarefaction analysis (i.e., random selection of increasingly larger individual proteins from each sample, and concomitant identification of total protein bin diversity) using the mothur software package (Schloss et al 2009). Metagenomes that appeared to be derived from non-natural samples (i.e., enrichment cultures or synthetically derived communities) and/or potentially photosynthetic communities based on published or other available data were excluded from the analyses. To compare protein coding gene diversity against environmental factors, relevant geochemical parameters (pH, temperature, SO_4^{2-}, Cl^-) were sought for each spring (Supplementary Data 10), based on publications referring either to the specific metagenome, or to other geochemical analyses for the spring. Additional information about the source of metadata for each spring is provided in Supplementary Data 10.

The functional content of SJ3 was then compared against 14 other well-characterized YNP metagenomes to identify proteins that differentiated SJ3 from other YNP springs (Supplementary Data 10). Total proteins from each of the 15 metagenomes were first annotated against the KEGG functional database (Kanehisa et al 2017) using the KEGG Automatic Annotation Server (Moriya et al 2007). An abundance-weighted table of KEGG orthology (KO) assignments was then computed as described above. A

subset of KOs were extracted (n=887) that encompassed proteins in the 'Energy Metabolism' subcategory of metabolism. The abundances of each KO were normalized to total metagenome KO size to account for unequal sequencing efforts between SJ3 and the much smaller, previously characterized YNP metagenomes. A Bray-Curtis distance matrix was then produced from the KO matrix using the veg.dist function in the R v.3.4.3 (https://www.r-project.org) vegan package v.2.4-6 (https://cran.r-project.org/web/packages/vegan/), which was then used in a Principal Coordinates Analysis (PCO) using the labdsv package v.1.8.0 (https://cran.r-project.org/web/packages/labdsv/) for R. The enrichment of specific KO-annotated proteins in each metagenome was assessed by first normalizing the KO abundances for each metagenome based on the total number of annotated KOs within each metagenome. The normalized data were then expressed as enrichment values where the mean for the whole dataset was subtracted from the value for each metagenome, and then normalized by the mean of the whole dataset (e.g.: [KO_Y normalized abundance for Sample Y – total mean KO_X]/ total mean KO_X). The resultant value indicates the enrichment of individual KOs within each metagenome, relative to all 15 YNP metagenomes in the dataset.

Functional comparisons among taxa.

We attempted to further identify the ecological mechanisms underlying the prevalence of numerous distinct populations of the same higher-order taxonomic lineages. Intra-lineage comparisons were first identified by assessing genomes (> 75% estimated completeness only) that were present within the same monophyletic clades determined in the phylogenomic analyses. The number of shared and unshared

'Metabolism' category KOs were identified and compared against the number of shared and unshared 'Genetic Information Processing' category KOs to determine the magnitude of metabolism-related protein differences in comparison to differences of proteins involved in genetic processing. Differences in genetic processing proteins would be expected to be more conservative among closely-related taxa and was thus used as a baseline. The relative numbers of protein differences for both categories were then identified across all intra-lineage comparisons (n=49 pairwise comparisons). To provide additional resolution for some pairwise comparisons, pairwise amino acid identity values (Konstantinidis and Tiedje 2005) were calculated between genomes using the enveomics software package (Rodriguez and Konstantinidis 2016).

Data Availability.

Assembled contigs have been uploaded to the Integrated Microbial Genomes (IMG) database under the genome ID 3300029625. The source data for Figs. 4 and 5 and Supplementary Figure 3 are provided as a source data file.

Acknowledgements

This work was supported by a Montana Space Grant Consortium grant to D.R.C. and E.S.B and a grant from the National Science Foundation to E.S.B. and D.R.C. (EAR-1820658). MRL acknowledges support from the NASA Earth and Space Science Fellowship program (NNX16AP51H). The NASA Astrobiology Institute is supported by grant numbers NNA15BB02A (to E.S.B.). We thank Christie Hendrix, Stacey Gunther, and Annie Carlson at YNP for research permitting.

Author Contributions

DRC, MRL, and ESB were responsible for generating and analyzing the metagenomics data. DRC and ESB wrote the manuscript with contributions from MRL.

Competing Interests

The authors declare no competing interests.

Materials & Correspondence

Correspondence and requests for materials should be addressed to E.S.B. (eboyd@montana.edu).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature. DOI: https://doi.org/10.1038/s41467-019-08499-1.

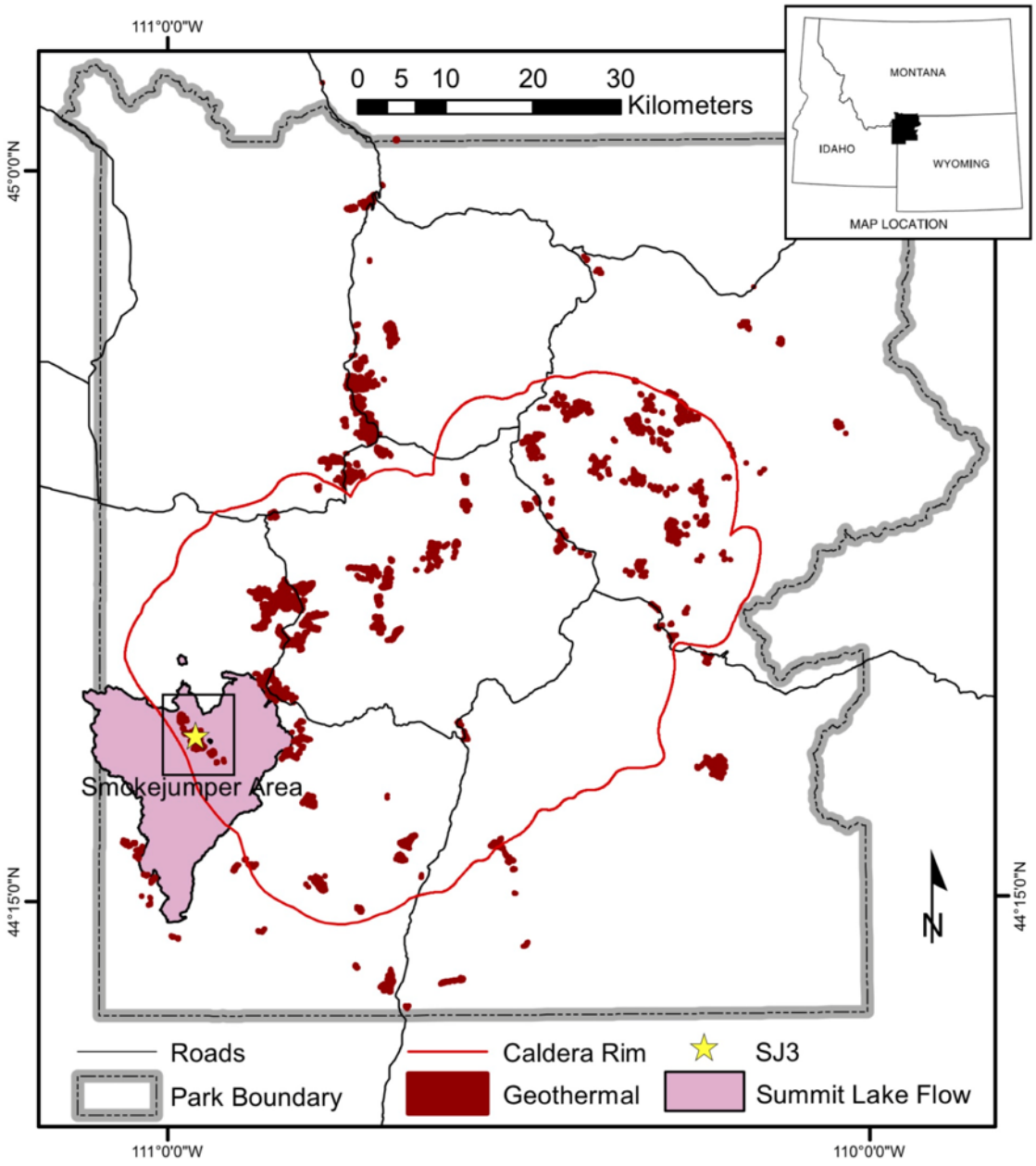


Figure 1. Map of Yellowstone National Park (YNP) showing the location of the Smokejumper Geyser Basin (SJGB) and 'Smokejumper 3' spring (SJ3). The caldera rim, road reference layers, and Summit Lake Rhyolite Flow reference is the same as in Christiansen (2001).

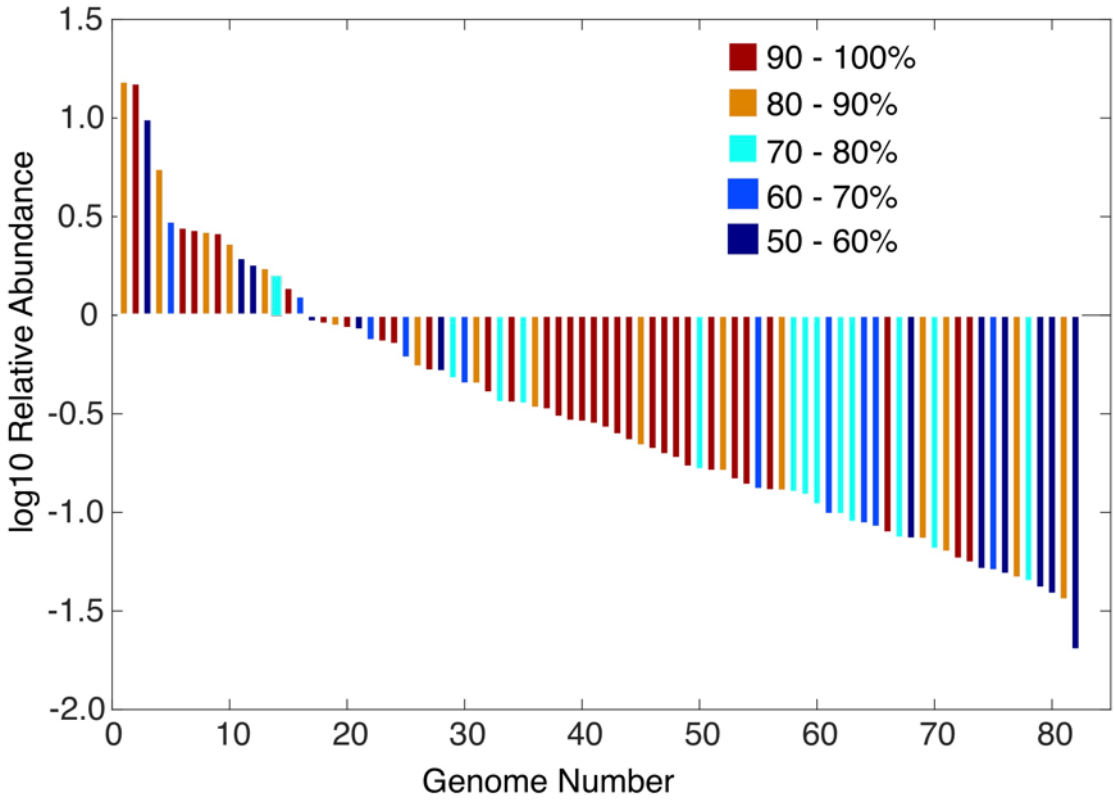


Figure 2. Rank-abundance plot of SJ3 reconstructed population level bins. Each vertical bar represents a reconstructed genome bin that has an estimated completeness > 50% (n=82). Genome bins are arranged by $\log_{10}$ transformed relative abundance (as a percentage) in decreasing order, as determined by read mapping. The scale on the upper right corresponds to the estimated percent completeness of each genome bin.


a
Bathyarchaeota (5)
Korarchaeota (2)
Aigarchaeota (3)
Thaumarchaeota (5)
Verstraetarchaeota (4)
Asgard Group
Geoarchaeota
MSBL1
Thermoproteales (5)
Hadesarchaea (1)
Sulfolobales (1)
Methanosarcinales
Desulfurococcales (13)
0.1
Methanocellales
DPANN (2)
Methanomicrobia
Theionarchaea
Halobacteria
Thermococci
Archaeoglobales (2)
Methanococci
DHVE2 (1)
Methanobacteria
Marine Group II
Methanomasiliicoccales
Thermoplasmatales (1)
b
Thermotogae (1)
Firmicutes (2)
Acetothermia (2)
Atribacteria/OP9
Chloroflexi (1)
Deinococcus-Thermus (1)
Synergistetes
Fervidibacteria(1)
Mollicutes
Armatimonadetes/OP10 (1)
Fusobacteria
Cyanobacteria
Unclassified (2)
Dictyoglomi (2)
CPR (1)
Thermodesulfobium (1)
Caldiserica/OP5 (1)
Spirochaetes (2)
Bacteroidetes
Aquificae (2)
Calescamantes/EM19
Balneolaeota
Ca. Firestonebacteria (1)
Chlorobi
0.1
Elusimicrobia (1)
Kryptonia (1)
Alpha/Beta/Gamma/ Epsilon-Proteobacteria (1)
Ignavibacteriae
Calditrichaeota
Chrysiogenetes/ Deferribacteres
Marinimicrobia
Deltaproteobacteria/ Thermodesulfobacteria (6)
Zixibacteria/ Latescibacteria
Nitrospirae
Nitrospinae
Gemmatimonadetes
Tectomicrobia
NC10
Fibrobacteres
WOR-3 (1)
KSB-1
Aminicenantes/OP8 (2)
Lentisphaerae
Acidobacteria (3)
Verrucomicrobia (1)
Ca. Poribacteria/NKB19
Chlamydiae
Omnitrophica/OP3
Planctomycetes

Figure 3. Maximum likelihood phylogenetic reconstruction of archaeal and bacterial population level MAGs from SJ3. Trees representing the major **a)** archaeal order- (or class/phylum-) level clades and the **b)** bacterial phylum-level clades were constructed using up to 104 or 31 phylogenetic marker genes, respectively (> 50% in any genome). Clades are collapsed as triangles with the taxonomic designation provided to the right. The numbers of MAGs that were identified in a specified clade are indicated in parentheses. Clades encompassing MAGs from 'SJ3' that were > 50% complete are shown in blue, those with MAGs > 75% are shown in red, and those representing multiple MAGs that ranged in completeness from 50 to > 75% are shown in purple. The scale bar shows the expected number of substitutions per site. The phylogenetic trees without collapsed clades showing individual MAG placement are provided as Supplementary Data 1 (Archaea) and Supplementary Data 2 (Bacteria).

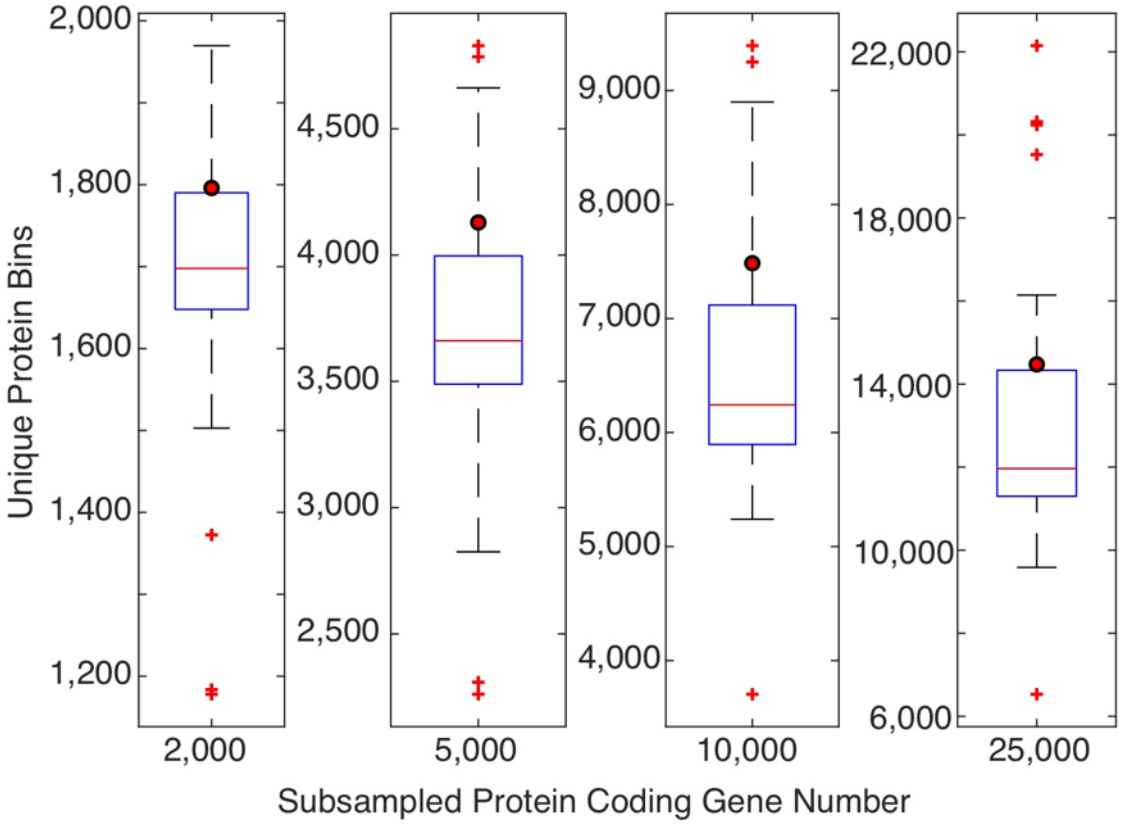


Figure 4. Comparison of protein family diversity in SJ3 relative to other chemosynthetic YNP hot spring metagenomes. Boxplots show the distribution of unique protein bins after subsampling 48 chemosynthetic YNP metagenomes at multiple depths at the levels shown below each box plot. Note that the Y-axis scales differ between panels. The values for SJ3 are shown as red circles. Red lines indicate the median for each distribution, while the boxes represent the interquartile range between the 25th and 75th percentiles. Whiskers show the full range of the data, not considering outliers and outliers are shown as red crosses. Source data are provided within the source data file accompanying this manuscript.

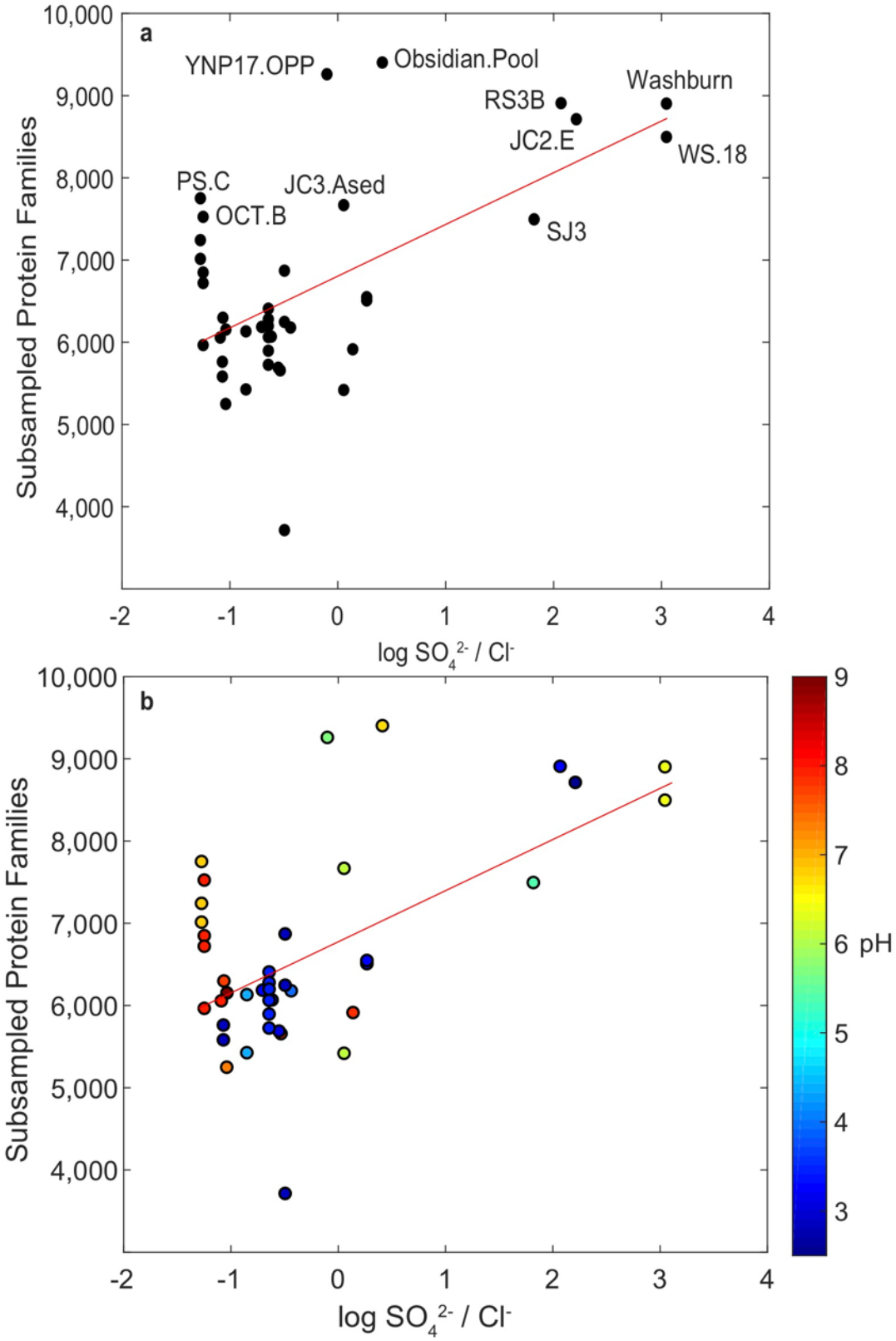

a
b
YNP17.OPP
Obsidian.Pool
Washburn
RS3B
JC2.E
WS.18
PS.C
JC3.Ased
OCT.B
SJ3
Subsampled Protein Families
log SO4 2- / Cl-
pH
10,000
9,000
8,000
7,000
6,000
5,000
4,000
-2
-1
0
1
2
3
4
9
8
7
6
5
4
3

Figure 5. Comparison of subsampled protein family diversity in chemosynthetic YNP hot spring metagenomes and inferred spring water origins. Protein subsampling was conducted on 10,000 randomly chosen protein homologs from each metagenome. The log transformed SO_4^{2-}/Cl^- ratio was calculated from publicly available data for each spring (Supplementary Data 10). Those metagenomes with higher protein family diversity than SJ3 after subsampling are indicated in panel **a**, with IDs corresponding to those provided in Supplementary Data 10. Red line shows linear regression model (adjusted R^2=0.33, $P < 0.001$, n=41). Source data are provided within the source data file accompanying this manuscript.

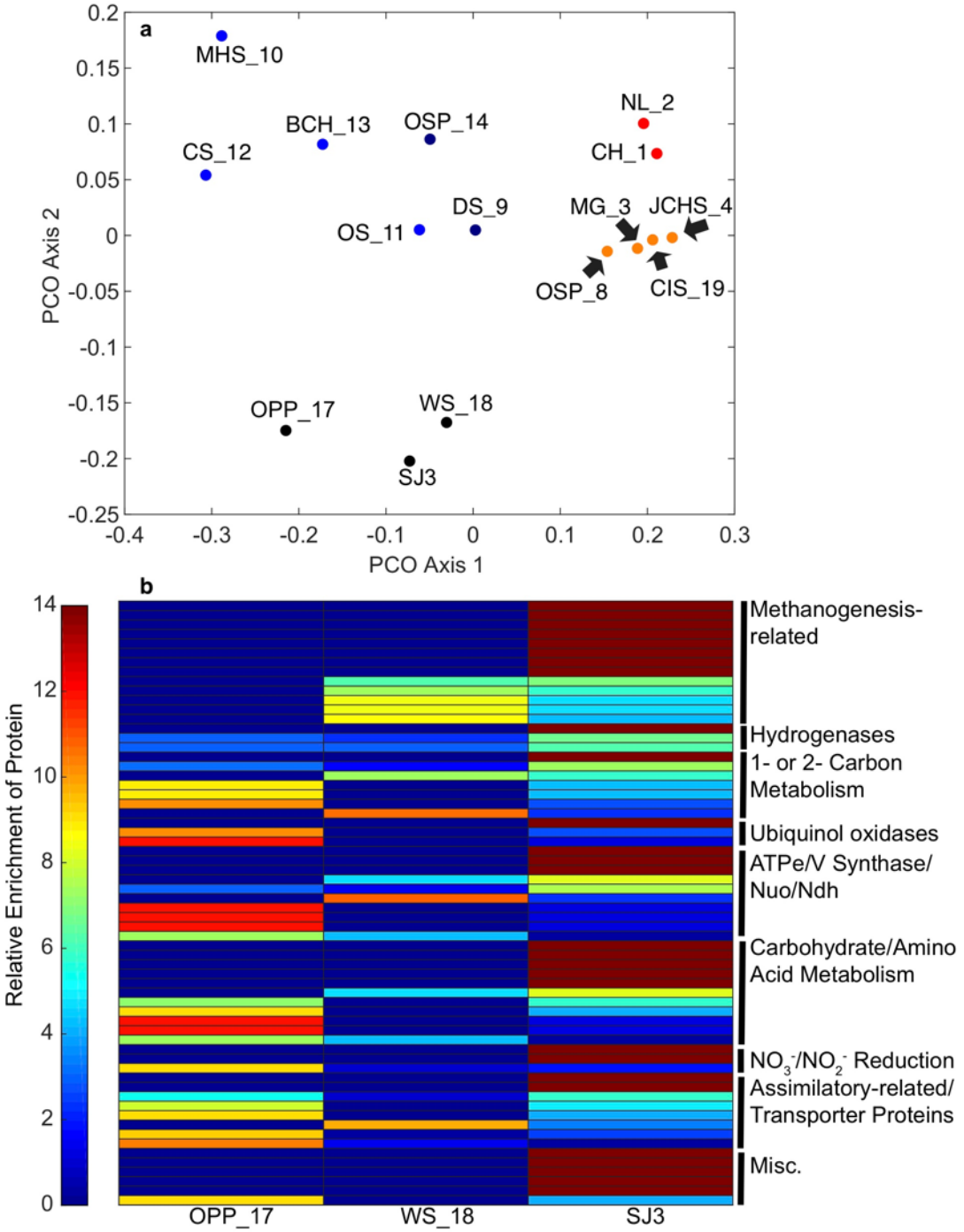


Figure 6. Similarity in proteins putatively involved in energy metabolism among SJ3 and 14 other published YNP metagenomes. a) Principal coordinates analysis (PCO) shows the similarity in proteins that are putatively involved in energy metabolism pathways among metagenomes, where each point represents a metagenome. Points are colored based on taxonomic and geochemical differences: light blue, high pH Aquificales-dominated communities (pH 6.5-7.9); dark blue, low pH Aquificales-

dominated communities (pH 3.1-3.5); red, low pH Sulfolobales-dominated communities (pH 2.6-4); orange, low-mid pH Crenarchaeaota-dominated communities (pH 3.4-6.4); black, mid pH Crenarchaeota/Aquificales/uncultured bacterial division dominated communities (pH 5.4-6.4). **b)** Heatmap showing enrichment of specific proteins putatively involved in the energy metabolism in the SJ3 community and two other similar chemotrophic YNP thermal spring communities: OPP_17 and WS_18. The heatmap represents only the proteins that were exclusive to the SJ3, OPP_17, and WS_18 communities, and where at least one homolog was present in SJ3. Relative enrichment was calculated after normalizing annotated proteins with the total number of proteins for each metagenome to account for unequal sequencing efforts. The scale bar on the left indicates relative enrichment where a value ≥ 14 indicates it was exclusively found in one metagenome, relative to the other 14 considered, and a value of ≤ 0 indicates that the protein was not detected in the metagenome. Proteins are grouped into broader functional categories shown on the right. All protein enrichment values that are depicted here are shown in further detail in Supplementary Data 3.

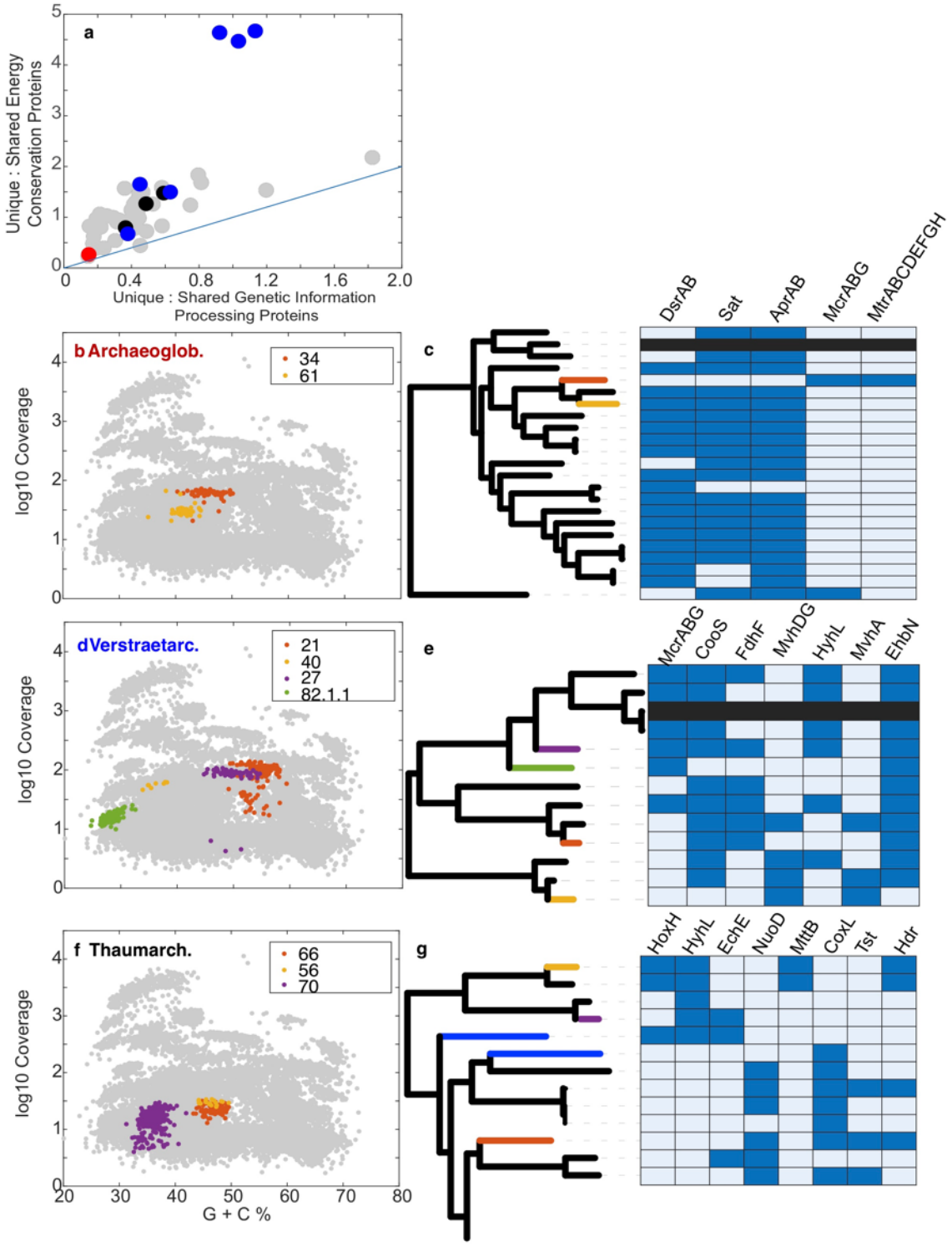

a
Unique : Shared Energy Conservation Proteins
Unique : Shared Genetic Information Processing Proteins
b Archaeoglob.
34
61
log10 Coverage
c
DsrAB
Sat
AprAB
McrABG
MtrABCDEFGH
dVerstraetarc.
21
40
27
82.1.1
e
McrABG
CooS
FdhF
MvhDG
HyhL
MvhA
EhbN
f Thaumarch.
66
56
70
G + C %
g
HoxH
HyhL
EchE
NuoD
MtlB
CoxL
Tst
Hdr

Figure 7. Differentiation of energy metabolism proteins within intra-taxon genomic comparisons. a) Ratio of unique/shared energy conservation-related proteins and the ratio of unique/shared genetic processing proteins for lineage comparisons. Each point represents a within-lineage pairwise comparison between genomes (n=49). Ratios were calculated based on the number of unshared KEGG-annotated proteins relative to the number of shared KEGG-annotated proteins for the 'Metabolism' and 'Genetic Information Processing' categories. Line shows a 1:1 relationship. Differentiation in genetic information processing proteins in closely-related taxa should be low, due to close phylogenetic relationships. Thus, little differentiation in metabolism-related proteins, relative to genetic processing proteins would indicate a lack of metabolic differentiation, whereas the converse would indicate greater metabolic differentiation. The Archaeoglobales, *Ca.* 'Verstraetarchaeota', and Thaumarchaeota comparisons are highlighted in red, black, and blue, respectively. Scatter plots **(b,d,f)** show the distribution of contigs based on G+C% and average sequencing coverage for genomic bins from three lineages with multiple representatives in SJ3: Archaeoglobales, *Ca.* 'Verstraetarchaeota', and Thaumarchaeota. Phylogenies and heatmaps of key differential functional genes **(c,e,g)** are shown for each of the three comparisons, respectively. Phylogenies are subsets from the full phylogenetic analysis represented in Fig. 3a, and branches are colored according to the color schemes for each lineage comparison on the left. Blue branches represent additional genomic bins from SJ3 that were estimated to be < 75% complete, and thus not included in these comparisons. Only the subset of the deepest-branches of the Thaumarchaeota are shown in **g**, while all genomes in our database are shown for the Archaeogloblaes and *Ca.* 'Verstraetarchaeota' in **c** and **e**, respectively. Protein abbreviations are as follows: DsrAB (dissimilatory sulfite reductase), Sat (sulfate adenylyltransferase), AprAB (adenosine 5'-phosphosulfate reducatase), McrABG (methyl coenzyme-M reductase), Mtr (N5-methyltetrahydromethanopterin methyltransferase), HoxH ([NiFe]-hydrogenase group 3d large subunit, LSU), HyhL ([NiFe]-hydrogenase group 3b LSU), EchE ([NiFe]-hydrogenase group 4 energy conserving hydrogenase LSU), NuoD (NADH dehydrogenase), MttB (trimethylamine--corrinoid protein co-methyltransferase), CoxL (aerobic carbon monoxide dehydrogenase LSU), Tst (thiosulfate/3-mercaptopyruvate sulfurtransferase), Hdr (heterodisulfide reductase subunits ABC), CooS (anaerobic carbon monoxide dehydrogenase catalytic subunit), FdhF (formate dehydrogenase-H), MvhADG ([NiFe]-hydrogenase group 3c methyl viologen reducing hydrogenase subunits DG), EhbN (energy converting hydrogenase B subunit N).

References

Achenbach-Richter L, Stetter KO, Woese CR (1987). A possible biochemical missing link among archaebacteria. *Nature* **327:** 348-349.

Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH (2013). Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nature Biotechnology* **31:** 533-+.

Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ *et al* (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods* **11:** 1144-1146.

Amenabar MJ, Urschel MR, Boyd ES (2015). Metabolic and taxonomic diversification in continental magmatic hydrothermal systems. In: Bakermans C (ed). *Microbial Evolution Under Extreme Conditions*. De Gruyter: Berlin, Germany.

Baker BJ, Saw JH, Lind AE, Lazar CS, Hinrichs KU, Teske AP *et al* (2016). Genomic inference of the metabolism of cosmopolitan subsurface Archaea, Hadesarchaea. *Nature Microbiology* **1:** 16002.

Barns SM, Fundyga RE, Jeffries MW, Pace NR (1994). Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proceedings of the National Academy of Sciences of the United States of America* **91:** 1609-1613.

Bergfeld D., Lowenstern JB, Hunt AG, III SWCP, Evans WC (2014). Gas and isotope chemistry of thermal features in Yellowstone National Park, Wyoming (ver 1.1 September 2014). In: Report USGSSI (ed).

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30:** 2114-2120.

Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK *et al* (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology* **35:** 725-731.

Boyd ES, Hamilton TL, Spear JR, Lavin M, Peters JW (2010). [FeFe]-hydrogenase in Yellowstone National Park: evidence for dispersal limitation and phylogenetic niche conservatism. *The ISME Journal* **4:** 1485-1495.

Boyd ES, Fecteau KM, Havig JR, Shock EL, Peters JW (2012). Modeling the habitat range of phototrophs in Yellowstone National Park: toward the development of a comprehensive fitness landscape. *Frontiers in Microbiology* **3**.

Boyd ES, Hamilton TL, Wang J, He L, Zhang CL (2013). The role of tetraether lipid composition in the adaptation of thermophilic archaea to acidity. *Frontiers in Microbiology* **4:** 62.

Brochier-Armanet C, Gribaldo S, Forterre P (2012). Spotlight on the Thaumarchaeota. *The ISME Journal* **6:** 227-230.

Christiansen RL (2001). The Quaternary and Pliocene Yellowstone Plateau volcanic field of Wyoming, Idaho, and Montana. *USGS Professional Paper*.

Cole JK, Peacock JP, Dodsworth JA, Williams AJ, Thompson DB, Dong HL *et al* (2013). Sediment microbial communities in Great Boiling Spring are controlled by temperature and distinct from water communities. *The ISME Journal* **7:** 718-729.

Colman DR (2015). Diversity of understudied archaeal and bacterial populations of Yellowstone National Park: from genes to genomes. Ph.D. thesis, University of New Mexico, Albuquerque, NM.

Colman DR, Feyhl-Buska J, Robinson KJ, Fecteau KM, Xu H, Shock EL *et al* (2016). Ecological Differentiation in Planktonic and Sediment-Associated Chemotrophic Microbial Populations in Yellowstone Hot Springs. *FEMS Microbiology Ecology*.

Colman DR, Poudel S, Stamps BW, Boyd ES, Spear JR (2017). The deep, hot biosphere: Twenty-five years of retrospection. *Proceedings of the National Academy of Sciences of the United States of America* **114:** 6895-6903.

Colman DR, Poudel S, Hamilton TL, Havig JR, Selensky MJ, Shock EL *et al* (2018). Geobiological feedbacks and the evolution of thermoacidophiles. *The ISME Journal* **12:** 225-236.

Costa KC, Navarro JB, Shock EL, Zhang CLL, Soukup D, Hedlund BP (2009). Microbiology and geochemistry of great boiling and mud hot springs in the United States Great Basin. *Extremophiles* **13:** 447-459.

Dombrowski N, Seitz KW, Teske AP, Baker BJ (2017). Genomic insights into potential interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal sediments. *Microbiome* **5**.

Evans PN, Parks DH, Chadwick GL, Robbins SJ, Orphan VJ, Golding SD *et al* (2015). Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science* **350:** 434-438.

Fournier RO (1989). Geochemistry and Dynamics of the Yellowstone-National-Park Hydrothermal System. *Annual Review of Earth and Planetary Sciences* **17:** 13-53.

Fu L, Niu B, Zhu Z, Wu S, Li W (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28:** 3150-3152.

Garcia JL, Patel BKC, Ollivier B (2000). Taxonomic phylogenetic and ecological diversity of methanogenic Archaea. *Anaerobe* **6:** 205-226.

Garrity GM, Holt JG (2001). Class VI. *Archaeoglobi* class nov.. In: Boone DR, Castenholz RW (eds). *Vol. 1: The Archaea and the Deeply Branching and Phototrophic Bacteria.*, 2nd edn. Springer-Verlag: New York.

Hamilton TL, Peters JW, Skidmore ML, Boyd ES (2013). Molecular evidence for an active endogenous microbiome beneath glacial ice. *The ISME Journal* **7:** 1402-1412.

Hoehler TM (2004). Biological energy requirements as quantitative boundary conditions for life in the subsurface. *Geobiology* **2:** 205-215.

Hu P, Tom L, Singh A, Thomas BC, Baker BJ, Piceno YM *et al* (2016). Genome-Resolved Metagenomic Analysis Reveals Roles for Candidate Phyla and Other Microbial Community Members in Biogeochemical Transformations in Oil Reservoirs. *MBio* **7:** e01669-01615.

Huber JA, Butterfield DA, Baross JA (2002). Temporal changes in archaeal diversity and chemistry in a mid-ocean ridge subseafloor habitat. *Applied and Environmental Microbiology* **68:** 1585-1594.

Huber JA, Butterfield DA, Baross JA (2003). Bacterial diversity in a subseafloor habitat following a deep-sea volcanic eruption. *FEMS Microbiology Ecology* **43:** 393-409.

Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ *et al* (2016). A new view of the tree of life. *Nature Microbiology* **1:** 16048.

Hugenholtz P, Pitulle C, Hershberger KL, Pace NR (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *Journal of Bacteriology* **180:** 366-376.

Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11:** 119.

Inskeep WP, Jay ZJ, Tringe SG, Herrgård MJ, Rusch DB, YNP Metagenome Project Steering Committee *et al* (2013). The YNP Metagenome Project: Environmental Parameters Responsible for Microbial Distribution in the Yellowstone Geothermal Ecosystem. *Frontiers in Microbiology* **4:** 67.

Jombart T, Dray S (2008). adephylo: exploratory analyses for the phylogenetic comparative method. *Bioinformatics* **26:** 1907-1909.

Jungbluth SP, Amend JP, Rappe MS (2017). Metagenome sequencing and 98 microbial genomes from Juan de Fuca Ridge flank subsurface fluids (vol 4, 170037, 2017). *Scientific Data* **4**.

Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research* **45:** D353-D361.

Kang DD, Froula J, Egan R, Wang Z (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3:** e1165.

Kieft TL (2016). Microbiology of the Deep Continental Biosphere. In: Hurst C (eds). *Their World: A Diversity of Microbial Environments. Advances in Environmental Microbiology, Vol. 1.* Springer, Cham.

Konstantinidis KT, Tiedje JM (2005). Towards a genome-based taxonomy for prokaryotes. *Journal of Bacteriology* **187:** 6258-6264.

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9:** 357-359.

LaRowe DE, Burwicz E, Arndt S, Dale AW, Amend JP (2017). Temperature and volume of global marine sediments. *Geology* **45:** 275-278.

Li D, Liu CM, Luo R, Sadakane K, Lam TW (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31:** 1674-1676.

Markusson SH, Stefansson A (2011). Geothermal surface alteration of basalts, Krysuvik Iceland-Alteration mineralogy, water chemistry and the effects of acid supply on the alteration process. *Journal of Volcanology and Geothermal Research* **206:** 46-59.

McCollom TM, Shock EL (1997). Geochemical constraints on chemolithoautotrophic metabolism by microorganisms in seafloor hydrothermal systems. *Geochimica et Cosmochimica Acta* **61:** 4375-4391.

Meyer-Dombard DR, Shock EL, Amend JP (2005). Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. *Geobiology* **3:** 211-227.

Mikheenko A, Saveliev V, Gurevich A (2016). MetaQUAST: evaluation of metagenome assemblies. *Bioinformatics* **32:** 1088-1090.

Miller SR, Strong AL, Jones KL, Ungerer MC (2009). Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline

hot springs in Yellowstone National Park. *Applied and Environmental Microbiology* **75:** 4565-4572.

Mitchell KR (2009). Controls on microbial community structure in thermal environments; exploring Bacterial diversity and the relative influence of geochemistry and geography. Ph.D. thesis, University of New Mexico, Albuquerque.

Mori K, Yamaguchi K, Sakiyama Y, Urabe T, Suzuki K (2009). Caldisericum exile gen. nov., sp. nov., an anaerobic, thermophilic, filamentous bacterium of a novel bacterial phylum, Caldiserica phyl. nov., originally called the candidate phylum OP5, and description of Caldisericaceae fam. nov., Caldisericales ord. nov. and Caldisericia classis nov. *International Journal of Systematic and Evolutionary Microbiology* **59:** 2894-2898.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research* **35:** W182-185.

Moser DP, Gihring TM, Brockman FJ, Fredrickson JK, Balkwill DL, Dollhopf ME *et al* (2005). Desulfotomaculum and Methanobacterium spp. dominate a 4- to 5-kilometer-deep fault. *Applied and Environmental Microbiology* **71:** 8773-8783.

Nakagawa S, Takai K, Inagaki F, Chiba H, Ishibashi J, Kataoka S *et al* (2005). Variability in microbial community and venting chemistry in a sediment-hosted backarc hydrothermal system: Impacts of subseafloor phase-separation. *FEMS Microbiology Ecology* **54:** 141-155.

Nordstrom DK, McCleskey RB, Ball JW (2009). Sulfur geochemistry of hydrothermal waters in Yellowstone National Park: IV Acid-sulfate waters. *Applied Geochemistry* **24:** 191-207.

Nordstrom KD, Ball JW, McCleskey RB (2005). Ground water to surface water: chemistry of thermal outflows in Yellowstone National Park. In: Inskeep WP, McDermott TR (eds). *Geothermal biology and geochemistry in Yellowstone National Park*. Montana State University: Bozeman, MT. pp 73 - 94.

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Research* **27:** 824-834.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research* **25:** 1043-1055.

Parks DH, Rinke C, Chuvochina M, Chaumeil PA, Woodcroft B, Evans PN *et al* (2018). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life (vol 2, pg 1533, 2017). *Nature Microbiology* **3:** 253-253.

Peters JW, Schut GJ, Boyd ES, Mulder DW, Shepard EM, Broderick JB *et al* (2015). [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. *Biochimica et Biophysica Acta* **1853:** 1350-1369.

Poulsen M, Schwab C, Jensen BB, Engberg RM, Spang A, Canibe N *et al* (2013). Methylotrophic methanogenic Thermoplasmata implicated in reduced methane emissions from bovine rumen. *Nature Communications* **4:** 1428.

Ragsdale SW, Pierce E (2008). Acetogenesis and the Wood-Ljungdahl pathway of CO(2) fixation. *Biochimica et Biophysica Acta* **1784:** 1873-1898.

Rodriguez, RL, Konstantinidis, KT (2016). The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. Preprint at https://peerj.com/preprints/1900/.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: Open-source, Platform-Independent, Community-supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology* **75:** 7537-7541.

Schrenk MO, Kelley DS, Delaney JR, Baross JA (2003). Incidence and diversity of microorganisms within the walls of an active deep-sea sulfide chimney. *Applied and Environmental Microbiology* **69:** 3580-3592.

Schut GJ, Zadvornyy O, Wu CH, Peters JW, Boyd ES, Adams MW (2016). The role of geochemistry and energetics in the evolution of modern respiratory complexes from a proton-reducing ancestor. *Biochimica et Biophysica Acta* **1857:** 958-970.

Seemann T (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30:** 2068-2069.

Shock EL, Holland M, Meyer-Dombard DR, Amend JP (2005). Geochemical Sources of Energy for Microbial Metabolism in Hydrothermal Ecosystems: Obsidian Pool, Yellowstone National Park. In: Inskeep WP, McDermott TR (eds). *Geothermal biology and geochemistry in Yellowstone National Park*. Montana State University Bozeman, MT. pp 95-110.

Shock EL, Holland M, Meyer-Dombard D, Amend JP, Osburn GR, Fischer TP (2010). Quantifying inorganic sources of geochemical energy in hydrothermal ecosystems, Yellowstone National Park, USA. *Geochimica et Cosmochimica Acta* **74:** 4005-4043.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W *et al* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7:** 539.

Spear JR, Walker JJ, McCollom TM, Pace NR (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 2555-2560.

Stamatakis A (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312-1313.

Stetter KO, Huber R, Blochl E, Kurr M, Eden RD, Fielder M *et al* (1993). Hyperthermophilic Archaea Are Thriving in Deep North-Sea and Alaskan Oil-Reservoirs. *Nature* **365:** 743-745.

Stevens TO, Mckinley JP (1995). Lithoautotrophic Microbial Ecosystems in Deep Basalt Aquifers. *Science* **270:** 450-454.

Takai K, Moser DP, DeFlaun M, Onstott TC, Fredrickson JK (2001). Archaeal diversity in waters from deep South African gold mines. *Applied and Environmental Microbiology* **67:** 5750-5760.

Valentine DL (2007). Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nature Reviews Microbiology* **5:** 316-323.

Vanwonterghem I, Evans PN, Parks DH, Jensen PD, Woodcroft BJ, Hugenholtz P *et al* (2016). Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nature Microbiology* **1:** 16170.

Wu M, Scott AJ (2012). Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **28:** 1033-1034.

Wurch L, Giannone RJ, Belisle BS, Swift C, Utturkar S, Hettich RL *et al* (2016). Genomics-informed isolation and characterization of a symbiotic Nanoarchaeota system from a terrestrial geothermal environment. *Nature Communications* **7:** 12115.

Xie W, Zhang CL, Wang J, Chen Y, Zhu Y, de la Torre JR *et al* (2014). Distribution of ether lipids and composition of the archaeal community in terrestrial geothermal springs: impact of environmental variables. *Environmental Microbiology*.

Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Backstrom D, Juzokaite L, Vancaester E *et al* (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541:** 353-358.

APPENDIX B

COPIES OF PERMISSIONS TO REPRINT

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Feb 26, 2019

This Agreement between Montana State University -- Melody Lindsay ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4536601320113
License date	Feb 26, 2019
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Geobiology
Licensed Content Title	Subsurface processes influence oxidant availability and chemoautotrophic hydrogen metabolism in Yellowstone hot springs
Licensed Content Author	Melody R. Lindsay, Maximiliano J. Amenabar, Kristopher M. Fecteau, et al
Licensed Content Date	Jul 23, 2018
Licensed Content Volume	16
Licensed Content Issue	6
Licensed Content Pages	19
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Geomicrobiology of Hydrogen in Yellowstone Hot Springs
Expected completion date	Mar 2019
Expected size (number of pages)	
Requestor Location	Montana State University 1916 South Black Avenue BOZEMAN, MT 59715 United States Attn: Montana State University
Publisher Tax ID	EU826007151
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a"Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.

- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order,** is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.

- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner.**For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts,** You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The Creative Commons Attribution Non-Commercial (CC-BY-NC)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The Creative Commons Attribution Non-Commercial-NoDerivs License (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library

http://olabout.wiley.com/WileyCDA/Section/id-410895.html

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

nature COMMUNICATIONS

Article | OPEN | Published: 08 February 2019

Mixing of meteoric and geothermal fluids supports hyperdiverse chemosynthetic hydrothermal communities

Daniel R. Colman, Melody R. Lindsay & Eric S. Boyd

Nature Communications **10**, Article number: 681 (2019) | Download Citation

Rights and permissions

This page is available in the following languages:

Creative Commons License Deed

Attribution 4.0 International (CC BY 4.0)

This is a human-readable summary of (and not a substitute for) the license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material

for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.