SYSTEMS BIOLOGY AND ECOLOGY OF MICROBIAL MAT COMMUNITIES

EDITED BY : Martin G. Klotz, Donald A. Bryant, Jim K. Fredrickson, William P. Inskeep and Michael Kühl

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**

SYSTEMS BIOLOGY AND ECOLOGY OF MICROBIAL MAT COMMUNITIES

Topic Editors:

Martin G. Klotz, Queens College of The City University of New York (CUNY), USA
Donald A. Bryant, The Pennsylvania State University, USA
Jim K. Fredrickson, Pacific Northwest National Laboratory, USA
William P. Inskeep, Montana State University, USA
Michael Kühl, University of Copenhagen, Denmark
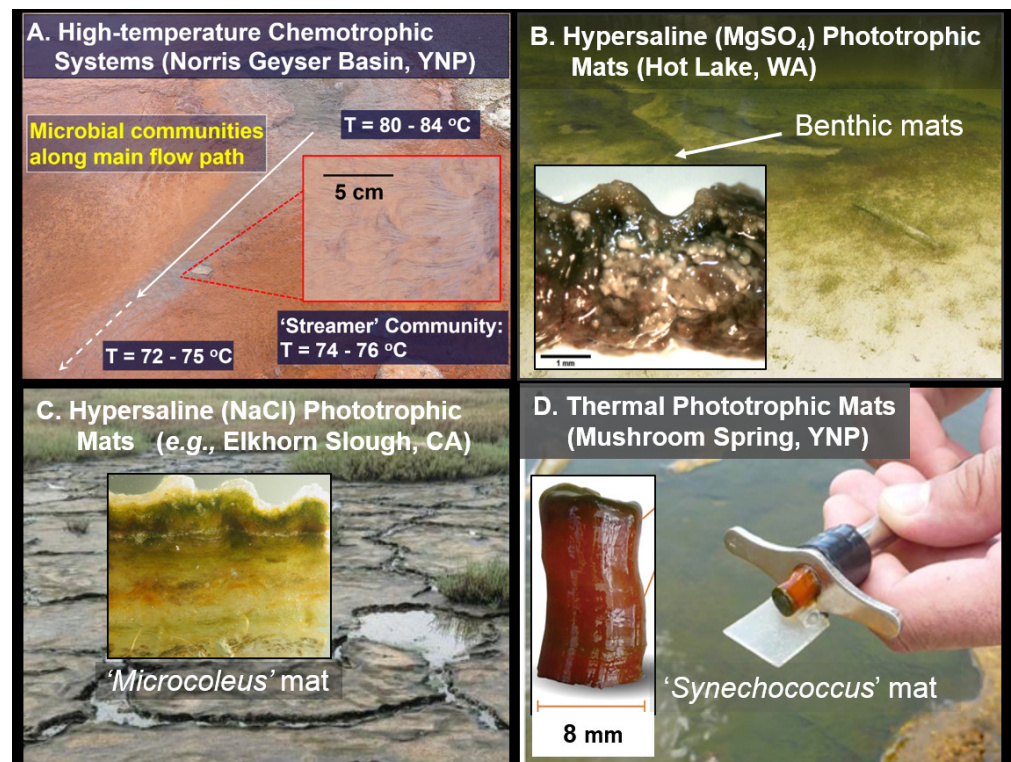
The primary aim of this research topic was to examine principles of systems biology in different types of stratified microbial mats present in extreme environments: high-temperature chemotrophic (A), hypersaline (B, C), and high-temperature phototrophic (D) communities. These environments exhibit contrasting gradients in key environmental variables (e.g., light, temperature, oxygen) and provide tractable systems for studying the tight metabolic coupling that occurs over relevant microbial scales.

Images adapted from:
A. Inskeep WP, Jay ZJ, Herrgard MJ, Kozubal MA, Rusch DB, Tringe SG, Macur RE, Jennings RdeM, Boyd ES, Spear JR and Roberto FF (2013) Phylogenetic and functional analysis of metagenome sequence from high-temperature archaeal habitats demonstrate linkages between metabolic potential and geochemistry. Front. Microbiol. 4:95. doi: 10.3389/fmicb.2013.00095.

B. Lindemann SR, Moran JJ, Stegen JC, Renslow RS, Hutchison JR, Cole JK, Dohnalkova AC, Tremblay J, Singh K, Malfatti SA, Chen F, Tringe SG, Beyenal H and Fredrickson JK (2013) The epsomitic phototrophic microbial mat of Hot Lake, Washington: community structural responses to seasonal cycling. Front. Microbiol. 4:323. doi: 10.3389/fmicb.2013.00323

C. Lawrence Livermore National Laboratory Genomic Sciences SFA, J. Pett-Ridge

D. Kim Y-M, Nowack S, Olsen MT, Becraft ED, Wood JM, Thiel V, Klapper I, Kühl M, Fredrickson JK, Bryant DA, Ward DM and Metz TO (2015) Diel metabolomics analysis of a hot spring chlorophototrophic microbial mat leads to new hypotheses of community member metabolisms. Front. Microbiol. 6:209. doi: 10.3389/fmicb.2015.00209

Microbial mat communities consist of dense populations of microorganisms embedded in exo-polymers and/or biomineralized solid phases, and are often found in mm-cm thick assemblages, which can be stratified due to environmental gradients such as light, oxygen or sulfide. Microbial mat communities are commonly observed under extreme environmental conditions, deriving energy primarily from light and/or reduced chemicals to drive autotrophic fixation of carbon dioxide. Microbial mat ecosystems are regarded as living analogues of primordial systems on Earth, and they often form perennial structures with conspicuous stratifications of microbial populations that can be studied in situ under stable conditions for many years. Consequently, microbial mat communities are ideal natural laboratories and represent excellent model systems for studying microbial community structure and function, microbial dynamics and interactions, and discovery of new microorganisms with novel metabolic pathways potentially useful in future industrial and/or medical applications. Due to their relative simplicity and organization, microbial mat communities are often excellent testing grounds for new technologies in microbiology including micro-sensor analysis, stable isotope methodology and modern genomics. Integrative studies of microbial mat communities that combine modern biogeochemical and molecular biological methods with traditional microbiology, macro-ecological approaches, and community network modeling will provide new and detailed insights regarding the systems biology of microbial mats and the complex interplay among individual populations and their physicochemical environment. These processes ultimately control the biogeochemical cycling of energy and/or nutrients in microbial systems. Similarities in microbial community function across different types of communities from highly disparate environments may provide a deeper basis for understanding microbial community dynamics and the ecological role of specific microbial populations. Approaches and concepts developed in highly-constrained, relatively stable natural communities may also provide insights useful for studying and understanding more complex microbial communities.

Table of Contents

CrossMark

Editorial: Systems Biology and Ecology of Microbial Mat Communities

Martin G. Klotz[1]*, Donald A. Bryant[2,3], Jim K. Fredrickson[4], William P. Inskeep[5]* and Michael Kühl[6]

[1] Department of Biology, Queens College of The City University of New York, New York, NY, USA, [2] Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA, [3] Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA, [4] Pacific Northwest National Laboratory, Richland, WA, USA, [5] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA, [6] Department of Biology, University of Copenhagen, Helsingør, Denmark

The Editorial on the Research Topic

Systems Biology and Ecology of Microbial Mat Communities

The goals of systems biology and microbial ecology are to gain a predictive understanding of how microbial communities function in dynamic systems, inclusive of interactions among community members and spatiotemporal changes in the physicochemical environment. Microorganisms in natural systems experience cycles of environmental change over different periodicities and amplitudes, and these processes are reflected in the composition, genetic repertoire and activity of microbial community members, and community function as a whole. A primary emphasis of this research topic was to focus on reports of tractable microbial communities in extreme environments (e.g., temperature, salinity, light, pH) where a foundation of genome sequence and other molecular (-omic) and geochemical measurements provide evidence of specific functional attributes of individual community members, which are directly linked with spatiotemporal changes in key environmental variables as well as the metabolic dynamics of other community members. Thermal or saline microbial mats are often stratified with respect to key environmental variables (e.g., light, oxygen) and exhibit compositional simplicity and low heterogeneity relative to habitats such as soils and/or natural waters. Consequently, the majority of contributions to this research topic focus on either high-temperature chemotrophic microbial mats, high-temperature phototrophic mats, or hypersaline phototrophic mats in marine or epsomitic systems. The structure and function of high-temperature systems of Yellowstone National Park (YNP) was evaluated using metagenome sequence and geochemical observations across a wide range of environmental conditions, which provided a basis for understanding the distribution of thermophiles in YNP and led to the discovery of several new archaeal and bacterial lineages. Genome sequences of relevant ecotypes have provided a foundation for interrogating more detailed spatiotemporal aspects of thermophilic phototrophic communities, including microsensor analysis of their physical and chemical microenvironment, and provided a rationale for comparison to hypersaline phototrophic mats. The fixation of carbon dioxide as a primary carbon source and the production of key cofactors by autotrophs are important processes, which support diverse heterotrophs across widely different environmental circumstances. Specific metabolic linkages among community members (e.g., production of storage compounds, nitrogen fixation, fermentation, sulfate reduction, hydrogen, and vitamin production) were documented in phototrophic mats, which revealed that these

processes changed in unexpected ways across a diel cycle. The tight metabolic coupling among specific populations that are highly adapted to spatial and/or temporal conditions is a common theme in natural environments, and this was demonstrated in detail using different light-adapted ecotypes of cyanobacteria (*Synechococcus* spp.) present in alkaline siliceous geothermal mats as primary producers. Finally, controlled experiments using pure cultures and/or consortia as "systems" revealed the importance of specific nutrient requirements, and importantly, how the exoproteome of a bacterium is controlled by the level of cellular oxidation. Ultimately, a predictive understanding of the complex network of abiotic and biotic interactions that occur in natural systems will also require detailed knowledge of the regulatory and physicochemical processes that govern gene expression, post-translational modifications, and protein activity. It is our hope that the articles included in this research topic advance a more comprehensive understanding of natural microbial communities, and demonstrate the utility of coupling molecular methods with detailed spatiotemporal measurements and dissection of microbial community function across gradients in key environmental variables such as light, temperature, pH, oxygen, or hydrogen. Integrated approaches across relevant microbial scales will lead to predictive capabilities useful for engineering microbial communities (or consortia) and for understanding how natural systems may respond to changes in key environmental variables (e.g., climate change).

AUTHOR CONTRIBUTIONS

WI drafted the manuscript, DB, JF, MK, and MGK revised the draft and all authors agreed to the final version. The articles in the RT were edited by MGK (10), WI (3), MK (1), and DB (1).

ACKNOWLEDGMENTS

The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem

William P. Inskeep[1,2]*, Zackary J. Jay[1,2], Susannah G. Tringe[3]*, Markus J. Herrgård[4], Douglas B. Rusch[5] and YNP Metagenome Project Steering Committee and Working Group Members[†]

[1] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman MT, USA
[2] Thermal Biology Institute, Montana State University, Bozeman MT, USA
[3] Department of Energy, Joint Genome Institute, Walnut Creek, CA, USA
[4] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm Denmark
[5] Center for Genomics and Bioinformatics, Indiana University, Bloomington, IN, USA

The Yellowstone geothermal complex contains over 10,000 diverse geothermal features that host numerous phylogenetically deeply rooted and poorly understood archaea, bacteria, and viruses. Microbial communities in high-temperature environments are generally less diverse than soil, marine, sediment, or lake habitats and therefore offer a tremendous opportunity for studying the structure and function of different model microbial communities using environmental metagenomics. One of the broader goals of this study was to establish linkages among microbial distribution, metabolic potential, and environmental variables. Twenty geochemically distinct geothermal ecosystems representing a broad spectrum of Yellowstone hot-spring environments were used for metagenomic and geochemical analysis and included approximately equal numbers of: (1) phototrophic mats, (2) "filamentous streamer" communities, and (3) archaeal-dominated sediments. The metagenomes were analyzed using a suite of complementary and integrative bioinformatic tools, including phylogenetic and functional analysis of both individual sequence reads and assemblies of predominant phylotypes. This volume identifies major environmental determinants of a large number of thermophilic microbial lineages, many of which have not been fully described in the literature nor previously cultivated to enable functional and genomic analyses. Moreover, protein family abundance comparisons and in-depth analyses of specific genes and metabolic pathways relevant to these hot-spring environments reveal hallmark signatures of metabolic capabilities that parallel the distribution of phylotypes across specific types of geochemical environments.

Keywords: thermophiles, geochemistry, microbial interactions, microbial mats, functional genomics

INTRODUCTION

The Yellowstone hotspot is responsible for an enormous number ($>14,000$) and diversity of thermal features that cover a wide range in pH (2–10), temperature (40–92°C), and geochemical properties (Fournier, 1989; Rye and Truesdell, 2007). The waters, rocks, and mineral surfaces in these geothermal sites provide an assortment of electron donors such as hydrogen, sulfide, and ferrous iron, as well as electron acceptors (e.g., dissolved oxygen or elemental S) that are vital to the survival of thermophilic microorganisms (Brock, 1978). Remarkably, many sites exhibit relative stability or regular patterns of temporal change, despite the dynamic nature of geothermal activity (Ball et al., 1998, 2002, 2010; McCleskey et al., 2005). Microbial communities in high-temperature environments are often dominated by just several types of microorganisms (phylotypes), and in general are significantly less diverse than lower-temperature habitats. Geothermal systems thus serve as valuable models for establishing linkages between genomic potential and environmental parameters, and for understanding environmental

controls on community structure and function. Moreover, many current thermal environments in YNP harbor modern-day analogs of microbial communities potentially important in ancient Earth, perhaps most notably, cyanobacterial mats that may have analogs dating back to the beginning of oxygenic photosynthesis and the "Great Oxidation Event" (Canfield, 2005; Konhauser, 2006, 2009; Sessions et al., 2009). Consequently, the expansive field site of YNP provides a natural laboratory for studying microbial evolution and adaptation across uniquely protected and preserved thermal environments.

The evolutionary history of all three branches of Life (*Bacteria*, *Archaea*, and *Eucarya*) is intertwined with major geological events that have shaped the planet (Nealson and Ghiorse, 2001; Konhauser, 2006). Relationships between geological events and microbial function are subjects of considerable debate and interest across numerous disciplines, and provide a strong basis for interdisciplinary collaboration. The extant global distribution of microorganisms reveals an astounding diversity of functional

attributes, many of which are specialized to specific physical or chemical constraints. The importance of microbially mediated reactions in the geosciences has been discussed (Newman and Banfield, 2002; Reysenbach and Shock, 2002; Falkowski et al., 2008), and numerous metabolic pathways are directly involved in key global processes such as nutrient cycling (e.g., C, N, S, Fe), gas exchange (CO_2 and CH_4), and trace element cycling (e.g., As, Sb, Hg). However, the microbial complexity typical of temperate soil and aquatic habitats (e.g., Madigan et al., 2012) has contributed to the difficulty in understanding specific linkages between community members and material cycling (Tringe et al., 2005; DeLong et al., 2006; Rusch et al., 2007). Metagenomics of high-temperature communities in YNP provides an unparalled opportunity to study specific microbiological associations across widely different geochemistry and mineralogy under highly controlled, pseudo-steady-state conditions.

Metagenome sequencing of environmental DNA has provided a useful tool for studying microbial community structure and function in numerous environments (e.g., Tyson et al., 2004; Tringe et al., 2005; Baker et al., 2006; DeLong et al., 2006; Rusch et al., 2007; Dick et al., 2009). Prior metagenome sequencing in geothermal systems of YNP has been used to dissect phototrophic mat communities (Klatt et al., 2011), and establish linkages among geochemical processes and microbial populations of chemotrophic communities (Inskeep et al., 2010). These studies provide specific examples of what can be gained from detailed phylogenetic and functional analysis of metagenome sequence. Here we describe, characterize and compare metagenomes and associated metadata collected across 20 different geothermal sites (**Figure 1**), which were selected to cover some of the major types of geochemical and microbiological systems active in YNP. The specific objectives of this manuscript were to (i) discuss rationale for the study design that focused on specific linkages between environmental parameters and microbial community structure, (ii) demonstrate the primary geochemical and geophysical attributes that separate different high-temperature microbial habitats, and (iii) provide an overview of metagenome sequence content and high-level (TIGRFAM) functional analyses that serve to introduce subsequent studies focused on three main geobiological ecosystem types: (1) phototrophic mats; (2) *Aquificales*-rich "filamentous-streamer" communities; and (3) archaeal-dominated sediments. These studies are the outgrowth of one of the major collaborative activities of an NSF Research Coordination Network, established to promote and coordinate research focused on geothermal biology and geochemistry in YNP. Importantly, the research presented here represents the first comprehensive and coordinated non-PCR based survey of microorganisms distributed across a broad spectrum of high-temperature habitats in YNP.

RESULTS AND DISCUSSION
GEOCHEMISTRY AND ENVIRONMENTAL CONTEXT
A geo-referenced geothermal database developed by the Yellowstone Center for Resources (U.S. National Park Service) contains information for numerous geothermal features including pH, temperature, electrical conductivity (EC), and site context (with image cataloging). A version of this database (~8000 entries) is available

FIGURE 1 | Map of Yellowstone National Park (YNP) showing the locations of 20 geothermal sites sampled for metagenome and geochemical analysis (green = phototrophic sites; yellow/red = archaeal sites; blue = Aquificales "streamer" communities; gray = Obsidian Pool Prime; site numbers defined in Table 1). The dashed line represents the boundary of the most recent caldera.
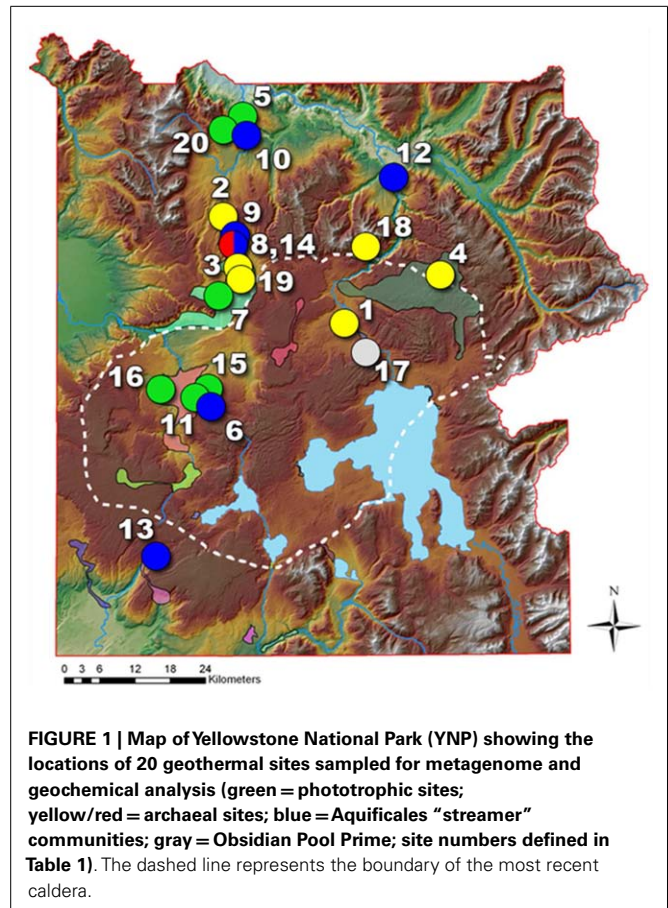
at the RCN YNP website[1], as well as other data on geothermal sites collected by either the U.S. Geological Survey (e.g., Ball et al., 1998, 2002, 2010; McCleskey et al., 2005) or individual researchers. The distribution of pH values across geothermal features in YNP is bimodal with peaks near pH 2.5 and 6.5 (**Figure 2**, based on $n \sim 7700$). While sites chosen for this study (colored bars along x-axis) strategically cover the bimodal pH ranges observed in YNP (**Figure 2**), the microbial communities inhabiting this wide range in pH are also strongly influenced by temperature, oxygen, sulfide, and other physical variables. EC values for a similar number of sites ($n \sim 6450$) show that most geothermal systems in YNP exhibit EC values ranging from 0.1–3 dS/m (or mmho/cm) (**Figure 2**), which corresponds to ionic strengths (I) of ~0.005–0.04 M. Compared to marine ($I \sim 0.7$ M), and other saline environments, the geothermal habitats sampled herein have significantly lower levels of dissolved ions, which would not be expected to impart significant environmental selection toward halophilic (high salt) or alkaliphilic (high-pH) microorganisms.

The current study was designed to cover a broad range of geochemical and temperature conditions, and to represent several of the major types of geothermal systems in YNP (Figure S1 in Supplementary Material). The breadth of habitats sampled capitalizes on a range of several key "system-defining" variables that give rise

[1]http://www.rcn.montana.edu

FIGURE 2 | Distribution of pH (A) and electrical conductivity (EC) (B) values across 7700 and 6450 geothermal features, respectively, cataloged in the Yellowstone Center for Resources (YCR) database (available on line via www.rcn.montana.edu). Vertical bars along pH and EC axes indicate values of the 20 sites described in the current study (Aquificales "streamer" communities = blue; archaeal-dominated sites = red; phototrophic mats = green). Ionic strengths ranging from ∼0.005–0.04 M are indicated with dashed lines (WS_18 has the highest EC in the study of ∼3.2 mmho/cm).

to different niches for thermophilic microorganisms. These key environmental variables and resulting habitat types are shown as a *decision-tree* (**Figure 3**), which is a descriptive tool to visualize the study design (showing replication of similar sites when possible). The nodes in the decision-tree also serve to represent hypotheses of the primary environmental factors that are most important for defining niches occupied by specific phylotypes. The corresponding metadata (geochemistry and physical context) associated with each of these sites are provided as supplemental material (Table S2 in Supplementary Material), but key attributes (T, pH, sulfide, or elemental S, and physical context) of each site are included here to understand the predominant site groupings, as well as the targeted communities and/or populations the study aimed to elucidate (**Table 1**; **Figure 3**).

The separation of sites as a function of pH represents a key decision point in the study design. As the first node, pH allows for an important distinction between the effective temperature limit for colonization by known phototrophs in acidic (pH < 5) versus more neutral (pH 5–9) environments. It is well established that the upper temperature limit for oxygenic phototrophs is ∼70–74°C (Brock, 1978; Madigan et al., 2012), but cyanobacteria do not generally colonize habitats below pH 4.5–5. Consequently, at pH values < 5, a different phototrophic limit is established near 54–56°C, corresponding to the upper temperature limit of members of the Cyanidales (red algae), diatoms, or possible representatives from the mildly acidophilic purple-bacteria (Toplin et al., 2008; Madigan et al., 2012). Consequently, bacterial phototrophic mats are confined to the upper node of the decision-tree, and are not expected as dominant organisms in habitat types below pH 5. Diverse eukaryotic (algal) phototrophic mats exist in YNP (<55°C), but these were not considered in the current study. The phototrophic sites (5, 6, 7, 15, 16, and 20) were chosen to target communities dominated by either oxygenic phototrophs (OP), filamentous anoxygenic phototrophs (FAPs), or anoxygenic phototrophs (AP). As noted, the sample from Obsidian Pool Prime (OPP_17) was obtained from biofilm growth on a large glass plate and represents a different physical context relative to other phototrophic sites.

Both low- and high-pH sites are further separated based on the presence or absence of sulfide and/or elemental S (**Figure 3**). This is an important variable because high sulfide concentrations also imply low oxygen (hypoxia). When sulfide levels are above detection (or if significant elemental sulfur and/or thiosulfate are present), there is generally rapid abiotic consumption of oxygen by reduced S species (Nordstrom et al., 2005, 2009). This does not rule out the possibility that O_2 influx from the atmosphere will influence the resulting microbial community, but systems with plentiful sulfide and/or elemental S will (i) contain little to no detectable dissolved oxygen, and (ii) contain species of S that can serve as potential electron donors or acceptors. Utilization of an O_2 variable in the decision-tree does not offer the same utility; low O_2 does not imply the presence of reduced S species. Some geothermal springs discharge non-sulfidic, hypoxic waters that contain other reduced constituents such as ferrous Fe (e.g., CP_7, OSP_8).

Two additional variables related to physical context are necessary to explain the study design and the corresponding habitat types included for metagenome analysis. Although these variables are included here as a simple *yes* or *no*, they represent important physical determinants of microbial community structure. Gradients in oxygen, light quantity, and quality, and sulfide are substantial as a function of mat depth and influence the distribution of microorganisms (Ward et al., 1998; Ramsing et al., 2000; van der Meer et al., 2005). For instance, sub-surface mat (SSMAT) layers are known to be less oxygenated and generally contain a greater abundance of FAPs relative to surface layers, which show a greater abundance of cyanobacteria (Ward et al., 2006; Klatt et al., 2011). Two of the phototrophic sites sampled in this study were from sub-surface mat positions at Mushroom Spring (MS_15) and Fairy Geyser (FG_16), where mat dissection was conducted to focus on novel FAP communities discovered in prior studies (Ward et al., 1998; Boomer et al., 2002), and because other

FIGURE 3 | The thermophilic microbial communities included in the YNP Metagenome Project were separated into predominant habitat types using a decision-tree with four "nodes" or primary environmental factors: pH, temperature (°C), the presence or absence of dissolved sulfide and/or elemental S, and additional physiographic context (4a: SSMat = sub-surface sample obtained from laminated phototrophic mats; 4b: Flow = primary flow channel). Decision nodes are represented by brown squares, node output by blue rectangles, and decision endpoints by triangles (i.e., predominant habitat types, site numbers). Targeted habitat types: FAP = filamentous anoxygenic phototrophs, OP = oxygenic phototrophs, AP = anoxygenic phototrophs, AQ-*Therm/Sulf/Hydro* = Aquificales "streamer" communities dominated by one of three major Aquificales lineages (i.e., *Thermocrinis*, *Sulfurihydrogenibium*, *Hydrogenobaculum*); S-SED = bottom or suspended sulfur-rich sediments dominated by archaea; Fe-Oxide = Fe-oxyhydroxide microbial mat dominated by archaea. The bulk aqueous characteristics of OPP_17 place this sample closest to sites WC_6 and CP_7 in the decision-tree, however, this sample was collected from biofilm growth on a large glass plate (see text, **Table 1**).

metagenome investigations were focused on the top layers of the MS_15 phototrophic mats (Klatt et al., 2011; Liu et al., 2011). Other phototrophic samples (CP, WC, BLVA) were obtained from surface mats that occur within geothermal outflow channels, and thus flow is an implied physical variable that defines these habitats.

Hydrodynamic conditions (i.e., flow rate, turbulence) are crucial for defining the microenvironmental context of microorganisms in geothermal settings. All Aquificales "streamer" communities (DS_9, MHS_10, CS_12, OSP_14, OS_11, BCH_13) were obtained within the primary flow path of high-velocity (0.1–0.3 ms^{-1}) outflow channels, as well as site OSP_8 (Fe-oxide mat), which is immediately down-gradient of OSP_14. Members of the deeply rooted bacterial Aquificales are known for assuming filamentous morphology in turbulent environments where the discharge of hot reduced waters occurs in the presence of air or oxygenated waters (e.g., marine hydrothermal vents, Reysenbach et al., 2000). Flow and turbulence of discharged geothermal waters encourages the degassing of $H_2S(aq)$ as well as the ingassing of $O_2(g)$ (Inskeep et al., 2005; Nordstrom et al., 2005). Consequently, the high-velocity, stream-channel habitats are considerably different from geothermal pools, which accumulate sediments of various sulfides, S^0, alunite, kaolinite, and polymorphs of SiO_2, depending on specific geochemical conditions such as pH, and levels of H_2S, As, Sb, and Fe (Inskeep et al., 2009). Sulfidic and elemental S-rich sediments (CH_1, NL_2, MG_3, CIS_19,

JCHS_4, WS_18) were sampled from low-flow habitats over a wide pH range (2.5–6.1) at temperatures considerably greater than the phototrophic limit (**Figure 3**). Sulfidic sites less than pH 5–6 can be further separated based on pH, an important variable influencing the distribution of specific crenarchaea (e.g., organisms within the order Sulfolobales increase in abundance with decreasing pH).

The oxidation of aqueous Fe(II) to solid-phase ferric oxyhydroxide is an exergonic reaction under most hydrothermal conditions (Amend and Shock, 2001; Amend et al., 2003; Inskeep et al., 2005), but of course requires dissolved O_2. The in-channel flow environments promote equilibration with atmospheric conditions and encourage O_2 influx. The acidic Fe-oxide mat (0.5–1 cm depth) sampled in Norris Geyser Basin (OSP_8) occurs within a high-velocity channel where oxygenation results in $O_2(aq)$ concentrations of $\sim$ 40–60 μM, which are 20–30% of theoretical saturation at this temperature and pressure. The concentration of Fe(II) does not appear in the decision-tree, because pH and the inference of oxygen (e.g., lack of sulfide) are already included as major determinants (concentrations of total Fe generally increase three orders of magnitude per pH unit decrease). Consequently, it is not the concentration of Fe *per se* that determines whether Fe(II) is oxidized, but whether the hydrodynamic or physical setting promotes oxygenation. For example, the more acidic CH_1 site contains $\sim$5 times more soluble Fe(II) than

Table 1 | Site names, abbreviations, and minimal set of environmental metadata necessary to separate different microbial community types (temperature, pH, dissolved sulfide or the presence of elemental S; Figure 3).

Site	Site Abbrev.	T (°C)	pH	DS (μM)	S⁰	Physical Context[1]	Target Phylum/Other[2]
White Creek	WC_6	48-50	8.2	<1	N	Flow channel, thermal creek	Cyanobacteria (*Mastigocladus* sp.), Chloroflexi
Chocolate Pots	CP_7	52	6.2	<1	N	Flow channel, spring source	Cyanobacteria, Chloroflexi, Rhodococcales
Mushroom Spring	MS_15	60	8.2	<1	N	Flow channel, subsurface mat	Chloroflexi (Red), Cyanobacteria (*Synechococcus* sp.)
Fairy Geyser	FG_16	36-38	9.1	<1	N	Splash mat, subsurface mat	Chloroflexi (Red)
Bath Lake Vista	BLVA_5	56-57	6.2	117	Y	Flow channel, near source	Chloroflexi (Green)
Bath Lake Vista	BLVA_20	54-56	6.2	117	Y	Flow channel, near source	Chloroflexi (Green), Gammaproteobacteria (*Thermochromatium* sp.)
Obsidian Pool Prime[3]	OPP_17	56	5.7	<1	N	Large Pool Glass slide	Bacterial OP Divisions
Dragon Spring	DS_9	68-72	3.1	80	Y	Flow channel	Aquificales (*Hydrogenobaculum*)
100 Spring Plain	OSP_14	72-74	3.5	10	N	Flow channel	Aquificales (*Hydrogenobaculum*)
Mammoth Hot Spr.	MHS_10	70-72	6.5	70	Y	Flow channel	Aquificales (*Sulfurihydrogenibium*)
Calcite Springs	CS_12	74-76	7.8	105	Y	Flow channel	Aquificales (*Sulfurihydrogenibium*)
Octopus Spring	OS_11	80-82	7.9	<1	N	Flow channel	Aquificales (*Thermocrinis*)
Bechler Spring	BCH_13	80-82	7.8	<1	N	Flow channel	Aquificales (*Thermocrinis*)
Crater Hills	CH_1	76	2.6	<3	Y	Sediment, Turbulent Pool	Crenarchaeota (Sulfolobales)
Nymph Lake	NL_2	88	~4	<3	Y	Sediment, Turbulent Pool	Crenarchaeota (Sulfolobales)
Monarch Geyser	MG_3	78-80	4.0	25	Y	Sediment, Pool Edge	Crenarchaeota (Desulfurococcales, Thermoproteales)
Cistern Spring	CIS_19	78-80	4.4	20	Y	Sediment, Pool Edge	Crenarchaeota (Desulfurococcales, Thermoproteales)
Joseph's Coat Spr.	JCHS_4	80	6.1	20	Y	Sediment Pool Edge	Crenarchaeota (Desulfurococcales, Thermoproteales)
Washburn Spring	WS_18	76	6.4	160	Y	Sediment Pool Edge	Crenarchaeota Aquificales
100 Spring Plain	OSP_8	72	3.4	<3	N	Flow channel, Fe-oxide mat	Crenarchaeota Mixed, Novel Archaea

The three major habitat types sampled include phototrophic mats (green), filamentous "streamer" communities (blue), and archaeal-dominated, elemental sulfur-rich or Fe-oxide sediments (yellow and red, respectively).

[1] Physical context refers to location of sample: Examples include: dissected sub-surface layer (under-mat) within a phototrophic mat; high-velocity channels containing filamentous "streamer" communities or down-gradient Fe-oxide mats; S-rich sediment in various geothermal pools

[2] Target phylum as discussed in submitted "RI-Files" and based on prior molecular, culture, microscopic and/or other investigations performed by individual investigators included in this project (Table S1 in Supplementary Material).

[3] The sample from Obsidian Pool Prime (OPP_17) represents a different habitat type and was obtained from a large thermal pool after colonization of a glass plate. Physicochemical characteristics place this sample closest to "phototrophic mats" in the current study.

OSP_8 (250 versus 45 μM), but due to the consistent geothermal delivery of low levels of $H_2S(g)$ and the ubiquity of solid-phase elemental sulfur in CH_1, the consumption of any O_2 is likely driven by reduced phases of sulfur rather than by Fe(II). Consequently, Fe-oxides are not present in the thermal pool of CH_1, although evidence of these phases was observed where the air-water interface meets rocks along the edge of the pool.

The study design captures a significant swath of high-temperature habitats in YNP. It is important to integrate the primary variables responsible for the separation of these habitat types together with analysis and interpretation of the metagenomes. The decision-tree (**Figure 3**) is also useful for identifying habitat types not included in the study, such as acidic geothermal habitats less than 55–60°C, in which members of the Cyanidales (red algae) are commonly observed (Toplin et al., 2008), and in which moderately thermophilic heterotrophic bacteria are known to increase in abundance (Macur et al., 2004; Kozubal et al., 2012). The sample collected at Obsidian Pool Prime (OPP_17) represents a different habitat context relative to other sites included in the study and was obtained from 2 months growth on a large ($625 \, cm^2$) glass slide suspended in the water column at 56°C (pH = 5.7). OPP is a large pool (∼0.5 ha) receiving mixed water inputs, and is the only sample in the study to be collected from the photic zone of a large water-body, albeit heavily influenced by geothermal inputs. Although the physical context is considerably different than other springs and pools included in the study, characteristics of OPP_17 place it near the phototrophic sites in the decision-tree (**Figure 3**).

SEQUENCING OVERVIEW AND MAJOR ASSEMBLIES

One of the primary aims of the study was to utilize the distinct geochemical differences across these 20 sites as a basis for understanding the distribution and function of different thermophiles. Moreover, given the uncertainties in predicting *in situ* metabolism solely using thermodynamic favorability or inference based on distantly related cultured isolates, it is useful to determine the actual organisms and genes present in high-temperature habitats as a means of constraining the metabolic possibilities and focusing our efforts on specific electron transfer reactions or catabolic

pathways. The total metagenome sequence (Sanger) per site ranged from 20 to 60 Mb (**Figure 4**), but the ratio of *assembled* sequence to total sequence varied considerably across these microbial communities. The total assembled sequence (i.e., the sum of total contig length) reflects the collapse of redundant sequence reads into contigs that correspond to the predominant populations present (**Table 1**). Consequently, for communities dominated by a single phylotype (e.g., MHS_10, NL_2), the total assembled sequence was only ∼2 Mbp, and there were few singlet reads remaining after assembly (nearly 90% of the sequences assembled into contigs in these cases). In contrast, more assembled sequence was obtained in more diverse communities with a greater number of dominant phylotypes (e.g., CP_7, BLVA, OS_11, WS_18), but at lower contig coverage. Moreover, the significant number of singlet reads remaining after assembly in many sites suggested that the sequencing coverage was insufficient to allow genome assembly of potentially important members of these communities.

The predominant phylotypes present in these three sample groups include expected microorganisms that have been the focus of prior studies at these exact locations (**Table 1**), as well as novel microorganisms in both the *Bacteria* and *Archaea*. Members of the Chloroflexi, Cyanobacteria, Chlorobi, Firmicutes, Bacteroidetes, Acidobacteria, and Proteobacteria dominated the phototrophic mats (**Figure 5**). The only sulfidic phototrophic site (BLVA) included in the study was sampled twice (∼8 months apart), before and during a bloom of purple sulfur bacteria (Gamma-proteobacteria), and both samples contained a strong signature of green Chloroflexi. Each of the in-channel "streamer" communities were dominated by one of three major Aquificales lineages specialized to variations in pH and sulfide (**Figure 5**). However, there was considerable variation in the other organisms present across the six Aquificales communities. The low-pH
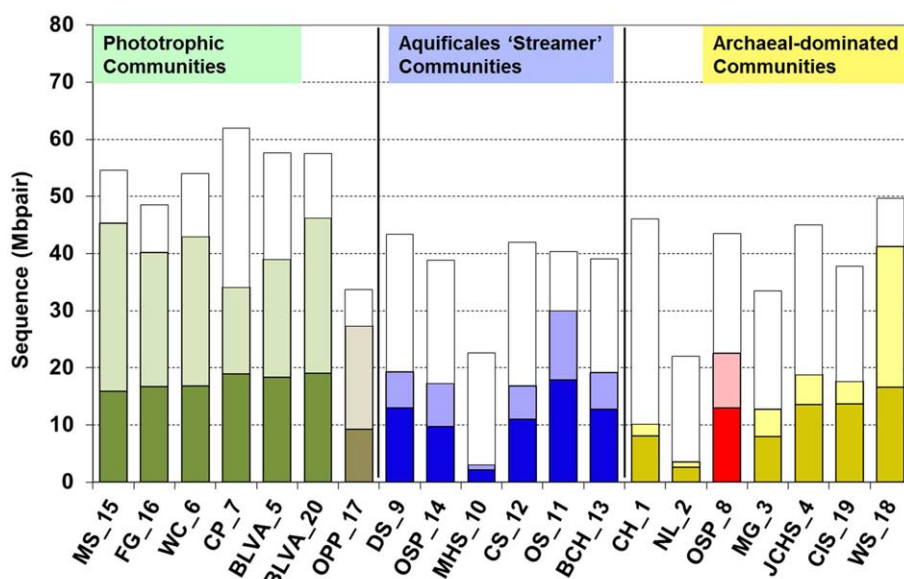
FIGURE 4 | Total amount of Sanger sequence (Mbpair) obtained for each site (top of open bars). The sum of assembled sequences (e.g., sum of all contig lengths) as well as that for all singleton sequences is shown as dark and light stacked bars, respectively [e.g., 62 Mbp of total sequence was obtained for CP_7, which resulted in 19 Mbp of assembled sequence (total contig length) and ∼14 Mbp of singleton sequences].
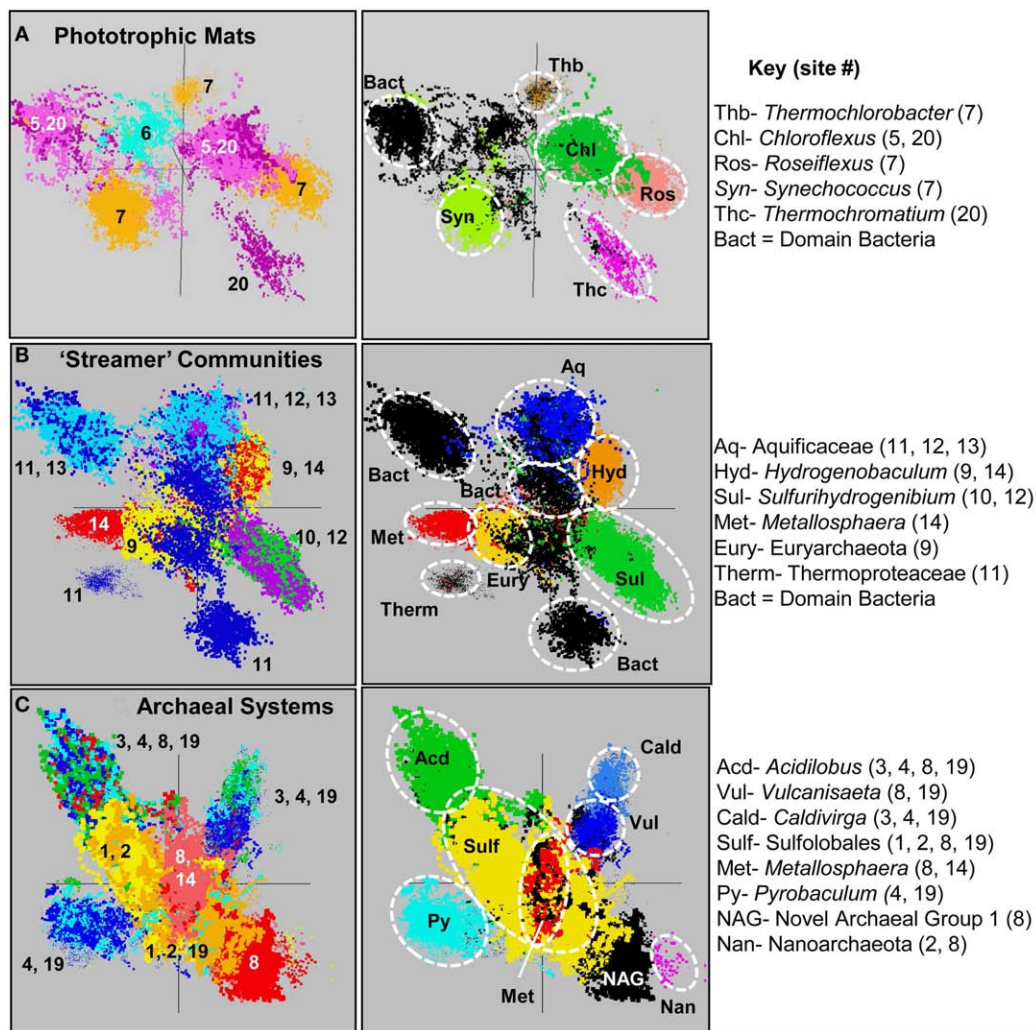
FIGURE 5 | Principal components analysis of oligonucleotide word frequencies (i.e., k-mers) of assembled metagenome sequence data for each group of sites. **(A)** Phototrophic, **(B)** "Streamer," and **(C)** Archaeal. Left Column: Assembled sequence data colored by site [*Phototrophic Mats*: BLVA_5, 20 = light-purple, dark-purple; WC_6 = light-blue; CP_7 = brown; MS_15 and FG_16 did not contain a significant number of contigs > 10,000 kb. "*Streamer*" *Communities*: DS_9 = yellow;

OSP_14 = red; MHS_10 = green; CS_12 = violet; OS_11 = dark-blue; BCH_13 = light-blue. *Archaeal Systems*: CH_1 = gold; NL_2 = yellow; MG_3 = green; CIS_19 = light-blue; JCHS_4 = dark-blue; OSP_8 = red; OSP_14 = pink). Sites OPP_17 and WS_18 not included. Right Column: Phylogenetic assignment of sequence clusters (identical PCA orientation) based on comparative sequence analysis of all contigs to reference genomes (APIS, Badger et al., 2006).

Aquificales communities (DS_9, OSP_14) contained significant sub-populations of different archaea (separated by the amount of sulfide), as did sites with higher pH values (OS_11, CS_12, BCH_13). In contrast, the MHS_10 sample contained no archaeal sequence and was dominated by one genus within the Aquificales (*Sulfurihydrogenibium*-like).

The high-temperature sulfur sediments and Fe-oxide mat samples were dominated by archaeal sequence reads (i.e., >85%) and contributed an extensive diversity of archaeal protein families (TIGRFAMS) not currently represented in environmental metagenome data, with the exception of replicate sites included in a smaller study of YNP chemotrophic communities (Inskeep et al., 2010). The sites dominated by Sulfolobales (CH_1, NL_2) represented a low-pH extreme and these communities are compared

in more detail to other sulfur sediments at moderate, but slightly acidic pH values (pH 4–6), which are dominated by members of the orders Desulfurococcales and Thermoproteales (sites MG_3, JCHS_4, CIS_19) (**Figure 5**). In contrast, the acidic Fe-mats (OSP_8, 14) contained Fe(II)-oxidizing Sulfolobales (*Metallosphaera* sp.) and novel archaeal populations observed only in the absence of significant levels of dissolved sulfide (**Table 1**; **Figure 5**).

RELATIVE DISTRIBUTION OF PROTEIN FAMILIES
To provide an independent assessment of the three major site groups, we analyzed broad functional differences among the 20 sites using relative TIGRFAM protein family abundances present in the metagenomes. TIGRFAM protein families represent a

relatively complete set of manually curated prokaryotic protein family models, which makes them suitable for the type of comparison we sought to perform. Abundance data, which were weighted by the scaffold assembly coverage, were obtained by counting protein copy numbers in assembled scaffolds for each site. Principal Component Analysis (PCA) of the TIGRFAM abundance data (**Figure 6**) showed that the first two principal components accounted for 80.1% of the functional variation among sites, and a two component plot (PC1 versus PC2) clearly separates the three major types of sites included in the study (phototrophic mats, "filamentous streamer communities," and archaeal-dominated S or Fe sediments). In particular, Component 2 appeared to measure the relative fraction of bacterial versus archaeal sequence in the metagenome. For example, although sites DS_9 and OSP_14 contain a dominant Aquificales population, these sites were more similar to the archaeal group because of the presence of subdominant (<30%) archaeal populations (albeit different archaeal phylotypes in each site). Although site OS_11 also contained a major Aquificales population, the community contained significant sub-populations of novel bacterial groups that were also found in phototrophic mats (e.g., Bacteroidetes, Firmicutes), and thus fell closer to the phototrophic group. Site WS_18 contained at least three major archaeal populations and was included with the archaeal group, but this sediment sample also contained at least two major bacterial populations that plotted away from the main archaeal cluster in TIGRFAM PCA space (**Figure 6**).

Comparing groups, the Aquificales "streamer" communities and archaeal-rich habitats yielded broader functional diversity than phototrophic sites because the chemotrophic sites included both bacteria and archaea while the phototrophic sites were dominated by bacteria. The phototrophic mats contained very few sequences attributable to archaea or members of the Aquificales. Component 1 reflects the amount of Aquificales versus phototrophs (e.g., important for bacterial sites) as well as the amount of sequence data attributable to members of the Sulfolobales (especially relevant for archaeal sites), in which the content of Sulfolobales decreased from left to right (see PC1, **Figure 6**). Component 3 separates two extreme sites (CH_1 and NL_2) that were dominated by Sulfolobales (**Figure 5**) from the majority of other sites, although PC3 explains only 5.9% of the variation across sites.

To identify what functional categories underlie the differences among sites, we obtained average TIGRFAM (Selengut et al., 2007) abundances for each functional category and site, and clustered these values using two-way hierarchical clustering (**Figure 7**).

FIGURE 6 | Principal components analysis of normalized TIGRFAM protein family abundance data across all 20 sites. The three panels show pairwise plots of the first three principal components (PC1 and PC2 account for 80% of the variation across sites, while PC3 only represents ∼6%). The first and second components separate sites into the three main habitat types studied (site labels and color groups are described in **Table 1**).

FIGURE 7 | Two-way hierarchical clustering of normalized TIGRFAM protein family abundance data averaged across intermediate-level TIGRFAM functional categories. The data was standardized (subtract mean and divide by standard deviation) across sites before clustering so that the color scale units are in standard deviations from the mean across sites. Yellow colors correspond to values that are higher than the site mean and blue colors to values that are lower than the mean.

The site clustering tree based on TIGRFAM abundances showed remarkable similarity to the decision-tree (**Figure 3**), which indicated that the environmental variables separating the major habitat types (especially pH, temperature, and reduced sulfur versus oxygen) are key determinants of microbial community structure and function. The TIGRFAM categories that varied most significantly among sites included a broad array of proteins important in central metabolism, cell replication, motility, photosynthesis, electron transport, and other metabolic functions (**Figure 7**). The observed site separation within a TIGRFAM category could result from

several factors that would not be evident without further investigation. These include the possibility that house-keeping functions simply reflect distinct phylogenetic differences observed across sites, and/or alternatively, that the range in abundance within a category may reflect the presence versus absence of specific metabolic capabilities contributed by phyla unique to a site. Consequently, the separation of sites using broad functional categories (**Figure 7**) accounts for both phylogenetic and functional differences, and provides strong support for the major site groupings emphasized in the three accompanying articles (Inskeep et al., 2013a,b; Takacs-Vesbach et al., 2013).

Specific TIGRFAM categories that varied in abundance across site groups included differences in functional categories such as nitrogen fixation and photosynthesis, expected to be highly represented in phototrophic sites (**Figure 7**). In contrast, TIGRFAMS that included cofactor biosynthesis (folic acid and lipoate), surface structures, fatty-acid biosynthesis, and chemotaxis/motility were more abundant in "streamer" communities. The archaeal sites exhibited a greater abundance of TIGRFAMs that included RNA processing, amino acid biosynthesis, nitrogen metabolism, aerobic respiration, and detoxification (**Figure 7**). In some cases, differences in TIGRFAM abundance across sites may have resulted from phylogenetic differences; for example, archaeal genes encoding a specific function may not be recognized as part of a TIGR-FAM category that has been established primarily from bacterial genomes. TIGRFAM categories that result in the greatest separation among the 20 sites included signal transduction, nitrogen fixation and regulatory functions, which represent a greater proportion of total sequences in phototrophic sites versus either Aquificales or archaeal communities (**Figure 7**; Figure S2 in Supplementary Material shows two of these TIGRFAM categories in greater detail).

More fine-grained functional differences among sites were obtained by only considering TIGRFAMs in a specific category such as "Electron Transport" (**Figure 8**). The site clustering based on the relative abundance of different electron transport domains provided a different view of the variation in functional attributes across sites. The distribution of specific respiratory complexes across sites [e.g., heme Cu oxidases (HCO), cytochrome *bd*-ubiquinol oxidases, blue Cu proteins, NiFe-hydrogenases, and nitrite/nitrate reductases] correlates more closely with geochemical parameters (e.g., oxygen, sulfur, hydrogen). Consequently, the distribution of different electron transport proteins across sites did not result in identical site clustering obtained using broader TIGRFAM categories (**Figure 7**). These observations are consistent with the fact that each group (i.e., phototrophic, archaeal, streamer communities) contained a range of sites with variable levels of sulfide or oxygen. Consequently, the factors contributing to functional variation within each of the three main site groups require additional dissection to appreciate how changes in these attributes are correlated with specific phylotypes (Inskeep et al., 2013a,b; Takacs-Vesbach et al., 2013).

The distribution and breadth of protein families observed in the current study represents a significant contribution to total protein diversity observed in all metagenomes currently found in the Integrated Microbial Genomes and Metagenomes (IMG/M) database (Markowitz et al., 2012) (as of spring 2012). This is due primarily to the fact that the current study focused on high-temperature communities rich in Aquificales, other deeply rooted novel bacteria, and numerous lineages of archaea. The functional diversity of the 20 Yellowstone geothermal samples included in this study was compared to other types of microbiomes as represented by the metagenomic sequence sets available in IMG/M, and based upon the TIGRFAM protein family abundance profiles for the assembled metagenomes. YNP samples are clearly separated from soil, water, gut, and other types of microbial communities in the public database (Figure S3 in Supplementary Material). As might be expected, the archaeal-dominated sites are most distant from the majority of public metagenomes, as few metagenome studies have targeted communities with significant archaeal populations. For example, the Crater Hills (CH_1) and Nymph Lake (NL_2) sites contained 1–2 predominant Sulfolobales populations and represented one extreme in the TIGRFAM PCA plots. Moreover, nearly identical samples obtained ∼2 years prior to this study (Inskeep et al., 2010) from four of the sites discussed here (CH_1, JCHS_4, CS_12, MHS_10) grouped with their expected replicate using PCA of TIGRFAM categories. Phototrophic mats and Aquificales "streamer" communities, which are dominated by eubacteria, fall closer to publicly available environmental metagenomes. The first principal component (which accounted for 24.8% of total variability in the TIGRFAM protein family abundance data across all sites in the database) almost entirely represents variation that exists within the YNP samples alone. The dramatic segregation of TIGRFAMS from YNP metagenomes versus previously sequenced mesophilic community metagenomes is indicative of the extensive contributions attributable to previously undescribed functional diversity.

SUMMARY

The metagenomes and corresponding metadata discussed here and in the accompanying articles provides a significant foundation of phylogenetic and metabolic information relevant to a wide range of geothermal ecosystems in YNP. Although the total sequence data obtained on an individual site basis was not sufficient to achieve assembly of all major phylotypes present in these distinct habitats, this study makes a considerable step toward understanding how microbial community structure and metabolic potential vary across a wide range of environmental parameters. A comparative distribution of protein families identified in the metagenome sequence reflects the major differences in phylogenetic structure among three primary groups of sites studied: phototrophic mats, Aquificales "streamer"-communities, and archaeal-dominated sediments. The systematic selection of geochemically distinct sites provides an enormous opportunity to link individual phylotypes (and associated metabolic attributes) with specific physicochemical properties of the habitats they occupy.

MATERIALS AND METHODS
GEOCHEMICAL ANALYSIS
Twenty geothermal sites in Yellowstone National Park (**Table 1**) were sampled and characterized in 2007–2008 (Figure S1 in Supplementary Material). Parallel samples of the bulk aqueous phase ($<0.2\,\mu$m) and sediment intimately associated with the microbial community were obtained simultaneously and analyzed using a
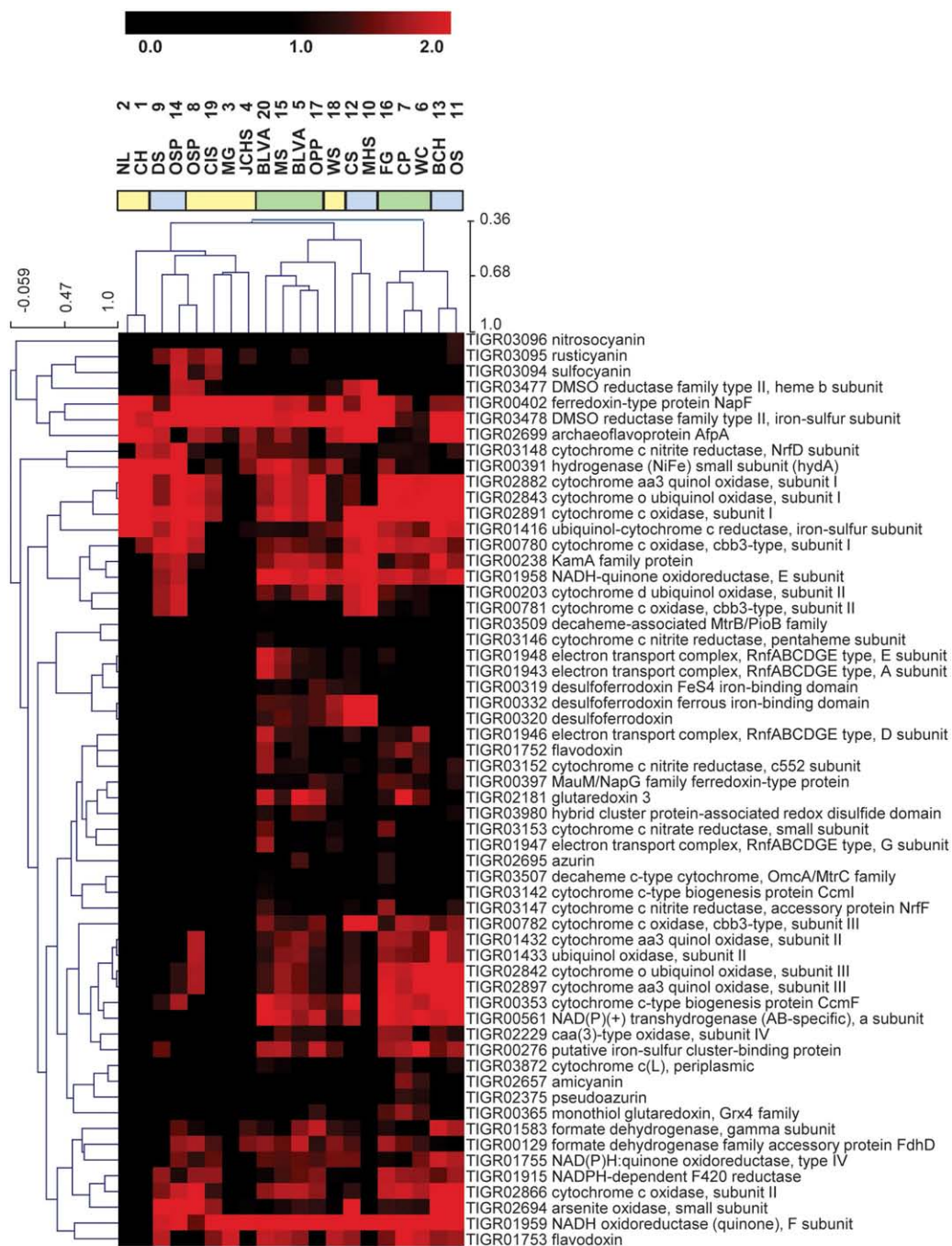
FIGURE 8 | Hierarchical clustering of metagenomes using relative TIGRFAM family abundances of genes in the TIGRFAM category "Electron Transport". Note that the abundance data has not been standardized, but rather log2-transformed (absolute abundances are shown).

combination of field and laboratory methods. Temperature, pH, and redox-sensitive species (Fe^{2+}/Fe^{3+}; total dissolved sulfide; dissolved O_2) were determined using field methods (with some exceptions given the number and diversity of sites studied) as described in more detail in previous reports (Langner et al., 2001; Macur et al., 2004). Total dissolved ions were determined using inductively coupled plasma (ICP) spectrometry and ion

chromatography (for all major cations, anions and trace elements). Dissolved gases (CO_2, H_2, CH_4) were determined using closed headspace gas chromatography (Inskeep et al., 2005) of sealed serum bottles ($<0.2\,\mu m$) obtained in the field without air contact. A complete dataset of geochemical information corresponding to these samples is provided (Table S2 in Supplementary Material). Selected sediment and microbial mat samples were analyzed using

scanning electron microscopy (Phillips Field Emission-SEM) in combination with energy-dispersive analysis of x-rays (EDAX) as well as x-ray diffraction (XRD). Thermodynamic calculations adjusted for temperature effects and performed using site-specific activities of dissolved and solid-phase constituents (Amend and Shock, 2001) showed that numerous oxidation-reduction reactions are exergonic (i.e., energy-yielding) in these types of geothermal systems and could support chemolitho- or chemoorganotrophic metabolisms (Amend et al., 2003; Inskeep et al., 2005; Shock et al., 2005). The oxidation of energy-rich, reduced constituents such as H_2, CH_4, H_2S, S^0, and As(III) is extremely favorable when geothermal waters are exposed to atmospheric O_2, or in some cases using alternate electron acceptors such as nitrate, ferric Fe, sulfate, or elemental S^0. Dissolved Fe(II) and ammonium (NH_4) concentrations can be very high in certain YNP geothermal systems, and offer yet another set of exergonic reactions that serve as a potential geochemical niche for chemotrophic organisms.

DNA EXTRACTION AND LIBRARY CONSTRUCTION

A standard DNA extraction protocol was used for the majority of samples, however, several sulfur sediments were ultimately subjected to various extraction kits to generate sufficient DNA yields for library construction. Our main emphasis was to obtain representative, unbiased environmental DNA for construction of small-insert libraries. Briefly, 3–25 g wet samples were extracted with 1 ml of Buffer A (200 mM Tris, pH 8; 50 mM EDTA; 200 mM NaCl; 2 mM sodium citrate; 10 mM $CaCl_2$) with lysozyme (1 mg/ml final concentration) for 1.5 h at 37°C. Proteinase K (final concentration 1 mg/ml) and SDS [final concentration 0.3% (w/v)] were then added and incubated for 0.5 h at 37°C. This first lysate was removed and the samples were re-extracted using bead-beating protocols. The two lysates were combined and extracted with phenol-chloroform, and the resulting DNA was re-precipitated in ethanol, treated with RNAase and quantified by gel electrophoresis and staining. Our intent was to avoid biasing the samples against organisms that may be difficult to lyse with chemical methods. The extraction kits (e.g., MoBio) used on several of the sulfur-containing sediment samples also included a modest physical lysis step. Across all 20 sites, the extracted DNA ranged in size from ∼4 to 12 kbp, and for many samples these procedures resulted in higher MW DNA ranging from 9 to 20 kbp.

RANDOM SHOTGUN SEQUENCING

Short-insert (∼3 kbp) pUC18 libraries were constructed from all sites, and preliminary sequencing of 5–10 megabases (Sanger), was performed as a quality control to determine if the initial sequencing results [via MEGAN (Huson et al., 2007) and blastx analysis] were consistent with the sample origin. All samples passed this quality checkpoint and were further sequenced to produce a total of 40–60 Mbp per site (**Figure 4**), with the exception of two sites. Both MHS_10 and NL_2 contained only one major population type that was covered sufficiently with ∼20 Mbp of Sanger sequence and these sites were also included for a half plate of 454 titanium pyro-sequencing (∼225 Mbp per sample), along with two additional samples (MG_3, JCHS_4). When combined with the total amount of Sanger sequencing performed (871 Mbp), this collectively represented nearly 2 Gbp of random shotgun sequence.

16S rRNA genes were also amplified from the DNA of each site using universal primers specific for bacteria and archaea, and one 384-well plate was sequenced from each successful library. Results from 16S rRNA gene sequencing were generally consistent with the phylogenetic signatures observed in the metagenome data, with the exception of one site (NL_2), presumably due to PCR amplification bias. For NL_2, the majority of 16S rRNA clones were Thermoproteales-like, but the random shotgun sequence data showed a predominant Sulfolobales-like population.

SEQUENCE ASSEMBLY

Metagenome assembly was conducted using two approaches (Celera and PGA), which resulted in reasonably similar overall assembly statistics. Due to the slightly better scaffold construction and prior history using the Celera assembler (Rusch et al., 2007; Inskeep et al., 2010), these assemblies were subjected to more detailed phylogenetic and functional analyses. However, the presence of genes or protein families described here are found in both assemblies, and the major results discussed here are not significantly different using either set of assembled data. Automated tools in IMG were used for gene identification and annotation for both sets of assembled data, and both sets of data are available on IMG/M. The following parameters were used for the Celera Assembler (Version 4.0): doOverlapTrimming = 0, doFragmentCorrection = 0, globalErrorRate = 12, utgErrorRate = 150, utgBubblePopping = 1, and useBogUnitig = 0. For PGA assemblies (Zhao et al., 2008), the following parameters were employed: OverlapLen = 30; Percent = 0.75; Clearance = 30; ClipIdn = 77; ClipQual = 10; CutoffScore = 400; EndOverhang = 800; InOverhang = 500; MinCovRep = 50; MinLinks = 2; MinSat = 3; NumIter = 50; PenalizeN = 1; QualOverLim = 400; QualScoreCutoff = 200; QualSumLim = 3500; SimDiFac = 30; Verbosity = 1.

As mentioned above, four sites (NL_2, MG_3, JCHS_4, MHS_10) received both Sanger sequence and one-half plate of 454 titanium pyro-sequencing (also available on IMG/M). Due to the success of Sanger sequencing in generating significant assemblies from these four sites, the additional pyro-sequence is not discussed in great detail in the subsequent manuscripts. In all four of these sites, contig coverage increased significantly when the pyro-sequence reads were included. However, the analyses discussed here are mostly derived from the Sanger sequence data, in order to maintain a consistent analytical approach for the data from all 20 sites.

PHYLOGENETIC ANALYSIS OF ENVIRONMENTAL SEQUENCE DATA

Phylogenetic analysis of random shotgun sequence reads was performed using a number of different approaches, and these yielded convergent information regarding the predominant phylotypes present in these geothermal sites. Individual sequence reads were analyzed using G + C content (mol%) distribution coupled with blastx and MEGAN (Huson et al., 2007) analysis. Genome-level phylogenetic analysis was accomplished using fragment recruitment of environmental sequence data to reference microbial genomes (Rusch et al., 2007). At the time of writing, the database contained reference microbial genomes for ∼1500 bacteria and 100 archaea, however, only a handful of microbial genomes

currently serve as appropriate references for the indigenous organisms within these communities. Assembled metagenome sequences were also analyzed using three dimensional PCA plots of nucleotide word frequencies (Teeling et al., 2004) with a simultaneous phylogenetic classification based on APIS or on a blast-based classification (Badger et al., 2006; Rusch et al., 2007). Metagenome sequence reported here can be viewed with these utilities at http://gos.jcvi.org/openAccess/scatterPlotViewer.html.

HIGH-LEVEL TIGRFAM ANALYSIS OF ENVIRONMENTAL SEQUENCE DATA

Assembled scaffolds were annotated as described in Inskeep et al. (2010) and predicted proteins from the scaffolds were assigned to TIGRFAM protein families (Selengut et al., 2007) using HMMER 3 (Eddy, 2011) with e-value cutoff of $1e-6$. TIGRFAM family counts for each scaffold were multiplied by the average coverage for the scaffold and the coverage weighted counts across all scaffolds for a particular site were summed to obtain estimated total family counts for a site. Different sites were normalized so that the total count over all TIGRFAM protein families was kept constant. TIGRFAM families were categorized using a two-level functional classification described on the TIGRFAM website[2]. PCA and statistical analysis of site group differences was performed using the STAMP v2.0 software (Parks and Beiko, 2010). The ANOVA test and Benjamini-Hochberg FDR correction implemented in STAMP was used to test for differences between multiple site groups. Two-way clustering was done on row-standardized (across sites) average TIGRFAM category abundance data using the Euclidean distance metric and complete-linkage hierarchical clustering. The MeV 4.8 (Saeed et al., 2003) software package was used for clustering and visualization. For comparison with other metagenomes in the IMG system, TIGRFAM profiles were generated for all the sites in this study and other metagenomic data sets downloaded using the IMG web interface.

DETAILED FUNCTIONAL ANALYSIS OF ENVIRONMENTAL METAGENOMIC SEQUENCE DATA

The assembled metagenome sequence data was also screened for specific functional genes corresponding to known or putative pathways in material and energy transfer. We were specifically interested in assessing metabolic potential for chemolithoautotrophy (CO_2 fixation and electron transfer) in high-temperature geothermal systems. Query DNA sequences known to code for proteins important in the oxidation of reduced chemical constituents or the reduction of a terminal acceptor (Table S3 in Supplementary Material) were used to search the environmental sequence data. Environmental sequence fragments exhibiting sequence similarity (e-values $< 10^{-10}$) to query sequences were then reanalyzed using blastp, and assessed individually using phylogenetic analysis of deduced protein sequences against known relatives, as well as fragment length relative to query length. False positives were minimized using this screening process. This included (i) sequences matching the correct protein family of the query sequence, but not the exact query sequence (e.g., Mo-pterin oxidoreductases versus a specific protein within this family); (ii)

sequences matching a query sequence due to regions of sequence similarity, but were clearly associated with a gene or gene cluster with different function; and (iii) sequences that returned mis-annotated blastp relatives. It is also possible that our inventory of metabolic potential missed sequences related to a specific query gene. For example, some genes found in the metagenome data were of insufficient length relative to a specific query sequence ($<40\%$) to make a definitive assignment. Moreover, the lower depth of coverage ($<1\times$) of sub-dominant phylotypes precluded a complete functional analysis of these organisms.

SEQUENCE AVAILABILITY

All annotated metagenome sequence assemblies (Celera/PGA) are available through the DOE-JGI IMG/M website[3] under IMG taxon OID numbers as follows (site order is identical to that presented in **Table 1**): Phototroph Sites [YNPSite06 (2022920004/2013515000), Site07 (2022920013/2014031006), Site15 (2022920016/2015219002), Site16 (2022920018/2016842 003), Site05 (2022920003/2013954000), Site20 (2022920020/2016 842008), and Site17 (2022920021/2016842005).]; Aquificales Sites [YNPSite09 (2022920010/2014031004), Site14 (2022920007/2013 954001), Site10 (2022920015/2015391001), Site12 (2022920011/20 14031005), Site11 (2022920012/2014031007), and Site13 (2022920 006/2013515002).]; Archaeal Sites [YNPSite01 (2022920009/2014 031002), Site02 (2022920014/2015219001, 2016842002), Site03 (2022920002/2014031003, 2016842001), Site19 (2022920017/2015 219000), Site04 (2022920008/2013843003), Site18 (2022920019/20 16842004), and Site08 (2022920005/2013515001).].

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Microbial_Physiology_and_Metabolism/10.3389/fmicb.2013.00067/abstract

Table S1 | Contributions of NSF Research Coordination Network Steering Committee and Working Group members to the Yellowstone Metagenome Community Sequencing Project (DOE_JGI CSP 787081).

Table S2 | Geochemical parameters measured in the bulk aqueous ($<0.2\,\mu$m) phase in 20 different geothermal systems sampled for metagenome analysis.

Table S3 | List of gene sequences and corresponding accession numbers used to query the assembled environmental sequence data for assessing

---

[2]http://www.jcvi.org/cgi-bin/tigrfams/index.cgi

[3]http://img.jgi.doe.gov/m

potential metabolic attributes associated with the predominant phylotypes found within these geothermal sites.

and phototrophic mats (green). **(A)** Signal transduction; **(B)** Regulatory functions (ANOVA test with Benjamini-Hochberg FDR multiple testing correction).

Figure S1 | Additional site photographs emphasizing landscape context of geothermal habitats and field sampling efforts (included as a separate file containing 53 annotated photographs).

Figure S2 | Top ranked TIGRFAM categories based on their ability to differentiate the three types of sites discussed in this study: aquificales "streamer" communities (blue), archaeal-dominated sediments (yellow),

Figure S3 | Principal components analysis of normalized TIGRFAM protein family abundance data across all 20 sites as well as a subset of diverse metagenomic datasets from IMG. The metagenomes are colored based on rough classification of the environments from which the sample originates: *Yellowstone sites in this study* = red; *Yellowstone sites from other studies* = orange; *sediments* = turquoise; *soils* = pink; *aquatic* = light-blue; *fungus-garden* = green; *animal-gut* = brown; *enrichment community* = light-orange; *other samples* = gray.

REFERENCES

Amend, J. P., Rogers, K. L., Shock, E. L., Gurrieri, S., and Inguaggiato, S. (2003). Energetics of chemolithotrophy in the hydrothermal system of Vulcano Island, southern Italy. *Geobiology* 1, 37–58.

Amend, J. P., and Shock, E. L. (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiol. Rev.* 25, 175–243.

Badger, J. H., Hoover, T. R., Brun, Y. V., Weiner, R. M., Laub, M. T., Alexandre, G., et al. (2006). Comparative genomic evidence for a close relationship between the dimorphic prosthecate bacteria *Hyphomonas neptunium* and *Caulobacter crescentus*. *J. Bacteriol.* 188, 6841–6850.

Baker, B. J., Tyson, G. W., Webb, R. I., Flanagan, J., Hugenholtz, P., Allen, E. E., et al. (2006). Lineages of acidophilic archaea revealed by community genomic analysis. *Science* 314, 1933–1935.

Ball, J. W., McCleskey, R. B., and Nordstrom, D. K. (2010). *Water-chemistry Data for Selected Springs, Geysers, and Streams in Yellowstone National Park, Wyoming, 2006-2008*. U.S. Geological Survey Open-File Report 2010-1192, 109. Available at: http://pubs.usgs.gov/of/2010/1192/.

Ball, J. W., McCleskey, R. B., Nordstrom, D. K., Holloway, J. M., and Verplanck, P. L. (2002). *Water-Chemistry Data for Selected Springs, Geysers, and Streams in Yellowstone National Park, Wyoming 1999-2000*. U. S. Geological Survey Open File Report 02-382, Boulder, CO.

Ball, J. W., Nordstrom, D. K., Cunningham, K. M., Schoonen, M. A. A., Xu, Y., and DeMonge, J. M. (1998). *Water-Chemistry and Onsite Sulfur Speciation Data for Selected Springs in Yellowstone National Park, Wyoming, 1994–1995*. U.S. Geological Survey Open-File Report 98-574, Denver.

Boomer, S. M., Lodge, D. P., Dutton, B. E., and Pierson, B. (2002). Molecular characterization of novel red green nonsulfur bacteria from five distinct hot spring communities in Yellowstone National Park. *Appl. Environ. Microbiol.* 68, 346–355.

Brock, T. D. (1978). *Thermophilic Microorganisms and Life at High Temperatures*. New York, NY: Springer-Verlag.

Canfield, D. E. (2005). The early history of atmospheric oxygen: homage to Robert A. Garrels. *Annu. Rev. Earth Planet. Sci.* 33, 1–36.

DeLong, E. F., Preston, C. M., Mincer, T., Rich, V., Hallam, S. J., Frigaard, N. U., et al. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311, 496–503.

Dick, G. J., Andersson, A. F., Baker, B. J., Simmons, S. L., Yelton, A. P., and Banfield, J. F. (2009). Community-wide analysis of microbial genome sequence signatures. *Genome Biol.* 10, R85.

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi:10.1371/journal.pcbi.1002195

Falkowski, P. G., Fenchel, T., and Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science* 320, 1034–1039.

Fournier, R. O. (1989). Geochemistry and dynamics of the Yellowstone National Park hydrothermal system. *Annu. Rev. Earth Planet. Sci.* 17, 13–53.

Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Res.* 17, 377–386.

Inskeep, W., Nordstrom, D. K., Mogk, D. W., Rodman, A. W., Fouke, B. W., Durães, N., et al. (2009). *Secondary Minerals Associated with Geothermal Features of Yellowstone National Park, Clay Mineral Society Workshop Lectures Series*, Vol. 17, ed. P. A. Schroeder (Chantilly, VA: The Clay Minerals Society), 1–29.

Inskeep, W. P., Ackerman, G. G., Taylor, W. P., Kozubal, M., Korf, S., and Macur, R. E. (2005). On the energetics of chemolithotrophy in nonequilibrium systems: case studies of geothermal springs in Yellowstone National Park. *Geobiology* 3, 297–317.

Inskeep, W. P., Jay, Z. J., Herrgard, M., Kozubal, M. A., Rusch, D. B., Tringe, S. G., et al. (2013a). Phylogenetic and functional analysis of metagenome sequence from high-temperature archaeal habitats demonstrate linkages between metabolic potential and geochemistry. *Front. Microbiol.* 4:95. doi:10.3389/fmicb.2013.00095

Inskeep, W. P., Klatt, C. G., Herrgard, M. J., Jay, Z. J., Rusch, D. B., Tringe, S. G., et al. (2013b). Community Structure and Function of High-temperature Chlorophototrophic Microbial Mats Inhabiting Diverse Geothermal Environments. *Front. Microbiol.* 4:106. doi:10.3389/fmicb.2013.00106

Inskeep, W. P., Rusch, D. B., Jay, Z. J., Herrgard, M. J., Kozubal, M. A., Richardson, T. H., et al. (2010). Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. *PLoS ONE* 5:e9773. doi:10.1371/journal.pone.0009773

Klatt, C. G., Wood, J. M., Rusch, D. B., Bateson, M. M., Hamamura, N., Heidelberg, J. F., et al. (2011). Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. *ISME J.* 5, 1262–1278.

Konhauser, K. (2006). *Introduction to Geomicrobiology*. Blackwell Science [ISBN 978-0-632-05454-1].

Konhauser, K. (2009). Deepening the early oxygen debate. *Nat. Geosci.* 2, 241–242.

Kozubal, M. A., Macur, R. E., Jay, Z. J., Beam, J. P., Malfatti, S. A., Tringe, S. G., 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. *Front. Microbiol.* 3:109. doi:10.3389/fmicb.2012.00109

Langner, H. W., Jackson, C. R., McDermott, T. R., and Inskeep, W. P. (2001). Rapid oxidation of arsenite in a hot spring ecosystem, Yellowstone National Park. *Environ. Sci. Technol.* 35, 3302–3309.

Liu, Z., Klatt, C. G., Wood, J. M., Rusch, D. B., Ludwig, M., Wittekindt, N., et al. (2011). Metatranscriptomic analyses of chlorophototrophs of a hot spring microbial mat. *ISME J.* 5, 1279–1290.

Macur, R. E., Langner, H. W., Kocar, B. D., and Inskeep, W. P. (2004). Linking geochemical processes with microbial community analysis: successional dynamics in an arsenic-rich, acid-sulphate-chloride geothermal spring. *Geobiology* 2, 163–177.

Madigan, M. T., Martinko, J. M., Stahl, D., and Clark, D. P. (2012). *Brock Biology of Microorganisms*, 13th Edn. Boston, MA: Benjamin Cummings.

Markowitz, V. M., Chen, I. A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., et al. (2012). IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res.* 40, D115–D122.

McCleskey, R. B., Ball, J. W., Nordstrom, D. K., Holloway, J. M., and Taylor, H. E. (2005). *Water-Chemistry Data for Selected Hot Springs, Geysers, and Streams in Yellowstone National Park, Wyoming, 2001-2001*. Boulder, CO: USGS.

Nealson, K., and Ghiorse, W. (2001). *Geobiology: Exploring the interface between the biosphere and the geosphere*. Washington D.C: American Academy of Microbiology.

Newman, D. K., and Banfield, J. F. (2002). Geomicrobiology: how molecular-scale interactions underpin biogeochemical systems. *Science* 296, 1071–1076.

Nordstrom, D. K., Ball, J. W., and McCleskey, R. B. (2005). "Ground water to surface water: Chemistry of thermal outflows in Yellowstone

National Park," in *Geothermal Biology and Geochemistry of Yellowstone National Park*, eds W. P. Inskeep, and T. R. McDermott (Bozeman, MT: Thermal Biology Institute, Montana State University), 73–94.

Nordstrom, D. K., McCleskey, R. B., and Ball, J. W. (2009). Sulfur geochemistry of hydrothermal waters in Yellowstone National Park: IV Acid-sulfate waters. *Appl. Geochem.* 24, 191–207.

Parks, D. H., and Beiko, R. G. (2010). Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 26, 715–721.

Ramsing, N. B., Ferris, M. J., and Ward, D. M. (2000). Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter-thick photic zone of a hot spring cyanobacterial mat. *Appl. Environ. Microbiol.* 66, 1038–1049.

Reysenbach, A. L., Banta, A. B., Boone, D. R., Cary, S. C., and Luther, G. W. (2000). Microbial essentials at hydrothermal vents. *Nature* 404, 835.

Reysenbach, A.-L., and Shock, E. (2002). Merging genomes with geochemistry in hydrothermal ecosystems. *Science* 296, 1077–1082.

Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., et al. (2007). The *Sorcerer II* Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biol.* 5:e77. doi:10.1371/journal.pbio.0050077

Rye, R. O., and Truesdell, A. H. (2007). "The question of recharge to the deep thermal reservoir underlying the geysers and hot springs of Yellowstone National Park," in *Integrated Geoscience Studies in the Greater Yellowstone Area: Volcanic, Tectonic, and Hydrothermal Processes in the YellowstoneGeoecosystem* ed. L. A. Morgan (U.S. Geol. Surv. Prof. Pap), 1717, 235–270.

Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., et al. (2003). A free, open-source system for microarray data management and analysis. *BioTechniques* 34, 374–378.

Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W. C., et al. (2007). TIGR-FAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Res.* 35, D260–D264.

Sessions, A. L., Doughty, D. M., Welander, P. V., Summons, R. E., and Newman, D. K. (2009). The continuing puzzle of the great oxidation event. *Curr. Biol.* 28, R567–R574.

Shock, E. L., Holland, M., Meyer-Dombard, D. R., and Amend, J. P. (2005). "Geochemical sources of energy for microbial metabolism in hydrothermal ecosystems: Obsidian Pool, Yellowstone National Park," in *Geothermal Biology and Geochemistry in YNP*, eds W. P. Inskeep and T. R. McDermott (Bozeman, MT: Montana State University Publications), 95-110.

Takacs-Vesbach, C., Inskeep, W. P., Jay, Z. J., Herrgard, M., Rusch, D. B., Tringe, S. G., et al. (2013). Metagenome sequence analysis of filamentous microbial communities obtained from geochemically distinct geothermal channels reveals specialization of three Aquificales lineages. *Front. Microbiol.* 4:84. doi:10.3389/fmicb.2013.00084

Teeling, H., Meyerdierks, A., Bauer, M., Amann, R., and Glöckner, F. O. (2004). Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environ. Microbiol.* 6, 938–947.

Toplin, J. A., Norris, T. B., Lehr, C. R., McDermott, T. R., and Castenholz, R. W. (2008). Biogeographic and phylogenetic diversity of thermoacidophilic Cyanidiales in Yellowstone National Park, Japan, and New Zealand. *Appl. Environ. Microbiol.* 74, 2822–2833.

Tringe, S. G., von Mering, C., Kobayashil, A., Salamov, A. A., Chen, K., Chang, H. W., et al. (2005). Comparative metagenomics of microbial communities. *Science* 308, 554–557.

Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., Richardson, P. M., et al. (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428, 37–43.

van der Meer, M. T. J., Schouten, S., Bateson, M. M., Nubel, U., Wieland, A., Kuhl, M., et al. (2005). Diel variations in carbon metabolism by green nonsulfur-like bacteria in alkaline siliceous hot spring microbial mats from Yellowstone National Park. *Appl. Environ. Microbiol.* 71, 3978–3986.

Ward, D. M., Bateson, M. M., Ferris, M. J., Kühl, M., Wieland, A., Koeppel, A., et al. (2006). Cyanobacterial ecotypes in the microbial mat community of mushroom spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 1997–2008.

Ward, D. M., Ferris, M. J., Nold, S. C., and Bateson, M. M. (1998). A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiol. Mol. Biol. Rev.* 62, 1353–1370.

Zhao, F., Zhao, F., Li, T., and Bryant, D. A. (2008). A new pheromone trail-based genetic algorithm for comparative genome assembly. *Nucleic Acids Res.* 36, 3455–3462.

Metagenome sequence analysis of filamentous microbial communities obtained from geochemically distinct geothermal channels reveals specialization of three Aquificales lineages

Cristina Takacs-Vesbach[1]*, William P. Inskeep[2]*, Zackary J. Jay[2], Markus J. Herrgard[3], Douglas B. Rusch[4], Susannah G. Tringe[5], Mark A. Kozubal[2], Natsuko Hamamura[6], Richard E. Macur[2], Bruce W. Fouke[7], Anna-Louise Reysenbach[8], Timothy R. McDermott[2], Ryan deM. Jennings[2], Nicolas W. Hengartner[9] and Gary Xie[10]

[1] Department of Biology, University of New Mexico, Albuquerque, NM, USA
[2] Department of Land Resources and Environmental Sciences, Thermal Biology Institute, Montana State University, Bozeman, MT, USA
[3] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark
[4] Center for Genomics and Bioinformatics, Indiana University, Bloomington, IN, USA
[5] Department of Energy-Joint Genome Institute, Walnut Creek, CA, USA
[6] Center for Marine Environmental Studies, Ehime University, Matsuyama, Ehime, Japan
[7] Roy J. Carver Biotechnology Center, University of Illinois, Urbana, IL, USA
[8] Department of Biology, Portland State University, Portland, OR, USA
[9] Computer, Computational, and Statistical Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, USA
[10] Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA

The Aquificales are thermophilic microorganisms that inhabit hydrothermal systems worldwide and are considered one of the earliest lineages of the domain *Bacteria*. We analyzed metagenome sequence obtained from six thermal "filamentous streamer" communities (∼40 Mbp per site), which targeted three different groups of Aquificales found in Yellowstone National Park (YNP). Unassembled metagenome sequence and PCR-amplified 16S rRNA gene libraries revealed that acidic, sulfidic sites were dominated by *Hydrogenobaculum* (Aquificaceae) populations, whereas the circum-neutral pH (6.5–7.8) sites containing dissolved sulfide were dominated by *Sulfurihydrogenibium* spp. (Hydrogenothermaceae). *Thermocrinis* (Aquificaceae) populations were found primarily in the circum-neutral sites with undetectable sulfide, and to a lesser extent in one sulfidic system at pH 8. Phylogenetic analysis of assembled sequence containing 16S rRNA genes as well as conserved protein-encoding genes revealed that the composition and function of these communities varied across geochemical conditions. Each Aquificales lineage contained genes for CO_2 fixation by the reverse-TCA cycle, but only the *Sulfurihydrogenibium* populations perform citrate cleavage using ATP citrate lyase (Acl). The Aquificaceae populations use an alternative pathway catalyzed by two separate enzymes, citryl-CoA synthetase (Ccs), and citryl-CoA lyase (Ccl). All three Aquificales lineages contained evidence of aerobic respiration, albeit due to completely different types of heme Cu oxidases (subunit I) involved in oxygen reduction. The distribution of Aquificales populations and differences among functional genes involved in energy generation and electron transport is consistent with the hypothesis that geochemical parameters (e.g., pH, sulfide, H_2, O_2) have resulted in niche specialization among members of the Aquificales.

Keywords: thermophiles, functional genomics, phylogeny, autotrophic processes, sulfide oxidation

INTRODUCTION

The order Aquificales represents a group of thermophilic microorganisms that inhabit marine and terrestrial hydrothermal systems worldwide (Ferrera et al., 2007). This lineage is of significant interest because its members are believed to comprise the deepest lineage of the domain *Bacteria* (Coenye and Vandamme, 2004; Barion et al., 2007), although alternative evolutionary histories have been suggested (Griffiths and Gupta, 2004; Boussau et al., 2008; Zhaxybayeva et al., 2009). The Aquificales include two predominant families, the Hydrogenothermaceae and Aquificaceae, and both are well-represented in different geothermal features of Yellowstone National Park (YNP) (Reysenbach et al., 2005). Members of the Hydrogenothermaceae include the genus *Sulfurihydrogenibium*, which inhabit circum-neutral sulfidic springs in YNP (Hugenholtz et al., 1998; Reysenbach et al., 2000b, 2005). The Aquificaceae comprise two divergent groups:

Hydrogenobaculum spp. are found predominantly in low-pH systems (pH < 4) (Jackson et al., 2001; Langner et al., 2001; Macur et al., 2004; D'Imperio et al., 2007, 2008; Hamamura et al., 2009), while *Thermocrinis*-like organisms generally exhibit higher-pH ranges (pH 6–9), overlapping with the near circum-neutral optimum for *Sulfurihydrogenibium* (Reysenbach et al., 1994; Blank et al., 2002).

Given our limited understanding of their phylogenetic history and metabolic potential, numerous questions remain regarding the ecology and evolution of the Aquificales. For example, the Aquificales are important primary producers in hydrothermal systems (Harmsen et al., 1997; Yamamoto et al., 1998; Reysenbach et al., 1999, 2000a; Blank et al., 2002; Eder and Huber, 2002; Inagaki et al., 2003; Spear et al., 2005) and several members of this lineage have been shown to use the reductive tricarboxylic acid (r-TCA) cycle for the fixation of carbon dioxide (Shiba et al., 1985; Beh et al., 1993; Ferrera et al., 2007). However, several members of the order are also capable of heterotrophy (Huber et al., 1998; Nakagawa et al., 2005; Caldwell et al., 2010), which makes their ecological role as possible primary producers unclear. Moreover, the diversity and variation of r-TCA-specific enzymes and substrates used by cultured members has led to speculation regarding the origin and the distribution of this pathway within the phylum (Hugler et al., 2007).

The Aquificales were originally named for the ability of the type strain, *Aquifex pyrophilus,* to oxidize molecular hydrogen to water (Huber et al., 1992). Consequently, hydrogen oxidation was generalized to the entire order based on the phenotype of a few cultivated members (Reysenbach and Cady, 2001; Donahoe-Christiansen et al., 2004; Huber and Eder, 2006; D'Imperio et al.,

2008). Additionally, the abundance of molecular hydrogen in hydrothermal systems was taken as evidence for the predominance of this metabolism in thermal features from YNP (Spear et al., 2005). However, the recent isolation of several new terrestrial species indicates some organisms within this group are unable to grow on hydrogen and do not contain Group I Ni-Fe hydrogenases (Reysenbach et al., 2009). Diverse metabolisms have been detected among cultured members of the Aquificales including the use of H_2, elemental sulfur, and thiosulfate as energy sources, and although generally aerobic, some members are microaerophilic and/or utilize nitrate as an electron acceptor (D'Imperio et al., 2008; Reysenbach et al., 2009). Thus, while the metabolic potential within the Aquificales is largely known from a few well-studied isolates, there is not an equal distribution of representatives from each of the families in culture, and their physiological capabilities remain unknown, especially under the environmental conditions that are characteristic of their natural habitats.

Much of what is known about the distribution of the Aquificales is based on molecular diversity studies across different habitat types (Reysenbach et al., 1994, 2000b; Hugenholtz et al., 1998; Stohr et al., 2001; Van Dover et al., 2001; Inagaki et al., 2003), although their ecology has been inferred largely from the geochemical conditions they inhabit (Fouke et al., 2000; Spear et al., 2005; Hall et al., 2008; Hamamura et al., 2009) and the physiology of the few cultivated members (Jahnke et al., 2001; Takai et al., 2002; Reysenbach et al., 2009). Based on molecular diversity surveys of 16S rRNA and metabolic genes, different Aquificales lineages do not generally share the same habitats in YNP (Reysenbach et al., 2005), although exceptions have been noted where *Thermocrinis-* and *Sulfurihydrogenibium*-like organisms have been

FIGURE 1 | Site photographs of Aquificales "streamer" communities sampled from Yellowstone National Park. The sites represent diverse geochemical environments as noted from the different mineralogy apparent visually [DS_9 (elemental sulfur); OSP_14 (Fe-oxides); MHS_10 (calcium carbonate); CS_12 (pyrite, amorphous Fe-sulfides); OS_11 (none); BCH_13 (none, Fe(III)-staining on silica)].

found together in sulfidic circum-neutral pH channels (Hall et al., 2008; Hamamura et al., 2009). Therefore, the distinct separation of different Aquificales lineages as a function of geochemical conditions provides a unique opportunity to study the evolutionary and ecological history of diverse members of this group *in situ*. Here we provide a phylogenetic and functional analysis of metagenome sequence obtained from six Aquificales "streamer" communities, representing two replicate communities of each of the three major Aquificales lineages found in high-velocity outflow channels of YNP. The distribution and metabolic potential of Aquificales in these distinct habitat types was correlated with the geochemical attributes and conditions measured in these same locations.

RESULTS

ENVIRONMENTAL AND GEOCHEMICAL CONTEXT

Each high-temperature microbial community (including associated biomineralized solid-phases) was sampled from the primary flow-path (e.g., Veysey et al., 2008) of geothermal outflow channels with high velocities ranging from 0.2 to 0.5 ms^{-1}. The sites were chosen to obtain a range in pH (3–8) and other geochemical attributes such as dissolved oxygen, dissolved sulfide, and/or predominant solid-phases associated with each microbial community (**Figure 1**; **Table 1**). Moreover, the physical and geochemical characteristics of these sites provide a representative subset of major Aquificales habitat types common in YNP (Reysenbach et al., 2005). Filamentous Aquificales "streamer" communities are generally found along the primary flow paths of thermal channels and colonize hydrodynamic regimes of shallow (<2 cm deep), high-velocity, turbulent (Reynold's Number > 100,000; Fouke, 2011) spring water. These zones exhibit rapid outgassing of dissolved gases such as CO_2 and H_2S, and in-gassing of oxygen (Inskeep et al., 2005; Kandianis et al., 2008; Fouke, 2011).

The acidic (pH 3–3.5) sites in Norris Geyser Basin (NGB) represent two common Aquificales habitats in YNP. The relatively high concentration of $H_2S(aq)$ in *Dragon Spring* (DS_9) (80–100 μM) results in the deposition of copious amounts of elemental

sulfur due to both biotic and abiotic oxidation (D'Imperio et al., 2008), whereas lower concentrations of $H_2S(aq)$ found in the *One Hundred Spring Plain* (OSP_14) site (<10 μM) are not sufficient to form elemental sulfur. In contrast, the "streamer" communities at OSP_14 form at a transition from reduced to oxygenated source waters and subsequent deposition of Fe-oxides (**Figure 1**). Scanning electron microscopy, elemental analysis, and electron diffraction confirm that elemental sulfur is the dominant solid phase in DS_9, and amorphous Fe(III)-oxide is the predominant phase associated with the community at OSP_14 (Langner et al., 2001; Inskeep et al., 2005) (**Figure 1**), although the gray discoloration of these Fe-oxides likely results from interaction of low concentrations of H_2S with ferric oxides.

The outflow channels at *Mammoth Hot Springs* (MHS_10) and *Calcite Springs* (CS_12) both contain high concentrations of dissolved sulfide (>100 μM), but other geochemical differences result in different solid-phases biomineralized within the "streamer" fabric (**Figure 1**). The hydrothermal fluids discharging at MHS intersect the Madison Limestone formation (Fournier, 1989; Fouke, 2011) and as a result, contain high concentrations of dissolved inorganic carbon (DIC), Ca, and to a lesser extent Mg. Upon discharge, $CaCO_3$ is biomineralized as aragonite needles within and along microbial filaments forming the "streamer" fabric (**Figure 1**) (Fouke et al., 2003; Kandianis et al., 2008; Fouke, 2011). In contrast, the black, filamentous structures at CS_12 are comprised primarily of pyrite formed along and within intertwined microbial filaments. The higher-pH, dissolved Fe, and total dissolved sulfide at *Calcite Spring* all favor the precipitation of iron sulfides relative to MHS_10 where no pyrite is observed. Rhombohedral crystals of elemental sulfur were found within the "streamers" of both MHS_10 and CS_12; however, the precipitation of elemental sulfur due to reaction of dissolved sulfide with oxygen is favored at low-pH (Xu et al., 1998; Nordstrom et al., 2005) when $H_2S(aq)/HS^- \gg 1$ (the pKa of H_2S at 70°C is ∼6.8; Amend and Shock, 2001).

Higher-pH (pH 7.5–8) "streamer" communities were also sampled from the outflow channels of low sulfide, alkaline siliceous

Table 1 | Sample location, aqueous geochemical parameters[1] and predominant solid-phases[2] associated with six Aquificales communities in Yellowstone National Park (YNP).

Location	T °C	pH	I	DIC	DS	S$_2$O$_3$	DO	As	Fe	CH$_4$	H$_2$	Solid Phases	Coordinates
			-- mM --		------------ μM ------------					---- nM ----			
Dragon Spring (DS_9)	70-72	3.1	18	1.3	80	<1	< 3	32	40	300	67	S^0	44° 39' 12.108" N. Lat 110° 28' 39.601" W. Lon
One Hundred Spring (OSP_14)	73	3.5	17	0.2	10	<1	22	28	37	50	40	Fe(AsO$_4$)(OH)$_3$ (amorphous)	44° 43' 53.4" N. Lat 110° 42' 40.9" W. Lon
Mammoth Hot Springs (MHS_10)	70-72	6.5	32	16.6	70	13	< 3	20	0.4	<10	14	CaCO$_3$ (aragonite), S^0	44° 58' 9.915" N. Lat 110° 42' 35.368" W. Lon
Calcite Springs (CS_12)	76	7.8	16	0.71	105	55	< 3	21	3.2	< 10	25	FeS$_2$, S^0	44° 54' 17.46" N. Lat 110° 24' 14.52" W. Lon
Bechler Spring (BCH_13)	82	7.8	11	7.4	<2	<1	-	2.4	0.7	-	-	Minor, Fe-oxide staining on SiO$_2$	44° 17' 9.431" N. Lat 110° 52' 42.521" W. Lon
Octopus Spring (OCT_11)	82	7.9	16	14.7	<2	<1	34	27	0.2	300	10	Minor, SiO$_2$	44° 32' 2.701" N. Lat 110° 47' 52.402" W. Lon

[1]*I, ionic strength calculated from aqueous geochemical modeling at sample temperature; DIC, dissolved inorganic C; DS, dissolved sulfide; S$_2$O$_3$, thiosulfate; DO, dissolved oxygen; CH$_4$ and H$_2$ values are for aqueous species.*

[2]*Predominant solid phases determined using scanning electron microscopy (FE-SEM) coupled with energy dispersive analysis of x-rays (EDAX) and x-ray diffraction (XRD).*

springs (*Octopus* and *Bechler Springs* (OS_11 and BCH_13). Under these conditions, filamentous communities within the flow channels at temperatures ranging from 78 to 82°C appear light-pink to gelatinous, and exhibit no obvious deposition of reaction products that may have formed from the oxidation of reduced species (e.g., no Fe-oxides, elemental sulfur or pyrite solid-phases are formed). This is consistent with the geochemical characteristics of these systems indicating low concentrations of reduced species such as sulfide, hydrogen and Fe(II), and mainly comprised of dissolved Na, Cl, and Si (**Table 1**). The predominant solid-phase(s) deposited along the evaporative margins of these channels are various forms of silica (although Fe-oxide staining is evident in BCH_13).

TAXONOMIC CLASSIFICATION OF METAGENOME SEQUENCE

We first report on analysis of individual metagenome sequences (average length ∼800 bp), then summarize results from metagenome assemblies. The G + C content (%) was determined for individual Sanger sequences (∼800 bp per read) from each of the six sites and classified taxonomically using "blastx" against the NCBI database (**Figure 2**). In addition, individual sequences were compared to available reference genomes using fragment-recruitment analysis (nucleotide) against more than 1400 reference genomes (tools developed during the Global Ocean Survey; Rusch et al., 2007) (**Figure A1** in Appendix).

Classification of unassembled DNA sequence reads (**Figure 2**) showed that the acidic sites (DS_9 and OSP_14) were both dominated by *Hydrogenobaculum* populations with an average G + C content (%) of 34.7%. However, these two samples were considerably different with respect to the archaeal populations present. The elemental sulfur site (DS_9) contained novel populations within the Euryarchaeota and Thaumarchaeota, whereas the Fe-rich streamers (OSP_14) contained a *Metallosphaera yellowstonensis*-like population known to be an important Fe(II)-oxidizer in these habitats (Kozubal et al., 2011). Two circum-neutral sites (pH 6.5 and 7.8), which both contain high concentrations of dissolved sulfide, provided an excellent comparison to the lower pH (3–3.5) sulfidic habitats. A major shift from *Hydrogenobaculum*-like organisms at pH 3–3.4 (DS_9, OSP_14) to *Sulfurihydrogenibium*-like organisms at pH 6.5 and 7.8 (MHS_10 and CS_12) was observed across these environments. Nearly 90% of the unassembled Sanger sequence reads from MHS_10 were related to *Sulfurihydrogenibium* spp. (**Figure 2**). The G + C content distribution plots of sequences from MHS_10 and CS_12 show the importance of *Sulfurihydrogenibium*-like organisms with an average G + C value of 32.4%. A second major Aquificales population is evident in CS_12 with an average G + C content of 46.5%, and is closely related to the major Aquificales organisms present in OS_11 and BCH_13 (**Figure 2**). The relative abundance of Aquificales lineages observed using all random sequences was generally very similar to the distribution of 16S rRNA gene sequences PCR-amplified using universal bacterial primers (**Figure A4** in Appendix).

The outflow channels of *Octopus* (OS_11) and *Bechler* (BCH_13) *Springs* exhibit similar pH values to CS_12 (pH 7.5–8),

FIGURE 2 | Frequency plots of G + C content (%) of random shotgun sequence reads obtained from six Aquificales habitats in Yellowstone National Park (YNP). Phylogenetic classification of each sequence read (∼800 bp) was performed using MEGAN (based on "blastx"), and provides one method for visualizing the predominant phylotypes that contribute to metagenome sequence in these environments (gray = total reads; orange = *Hydrogenobaculum;* red = *Metallosphaera;* green = *Sulfurihydrogenibium;* violet = *Thermus;* blue = Aquificaceae; light-blue = *Hydrogenivirga;* light-orange = Domain *Bacteria;* light-green = Thermoproteaceae).

but do not contain significant concentrations of dissolved sulfide (below detection). The predominant Aquificales sequences in OS_11 and BCH_13 have nearly identical G + C content distributions (average G + C = 46.7%) and taxonomic assignment (*Thermocrinis*-like), and were also similar to the sub-dominant Aquificales population in CS_12 (**Figure 2**). Past work on the distribution of 16S rRNA genes in OS_11, BCH_13, and similar channels have also shown that these Aquificales are related to *Thermocrinis* spp. (Hugenholtz et al., 1998; Reysenbach et al., 2005; Hall et al., 2008). The only genome sequence available within this genus was *T. albus*, and this genome did not serve as an adequate reference for metagenome sequence from these sites (e.g., see fragment recruitment, **Figure A1** in Appendix). Consequently, the YNP *Thermocrinis*-like populations are not well-represented by currently available reference genomes, and phylogenetic assignment of these population(s) in the G + C content frequency plot (**Figure 2**) was made at the family level (Aquificaceae). In contrast, the genome sequence of *Hydrogenobaculum* sp. Y04AAS1 is a reasonably good reference for populations present in DS_9 and OSP_14, while the *Sulfurihydrogenibium* spp. genomes (strain Y03AOP1 or *S. yellowstonense;* Reysenbach et al., 2009) serve as good references for metagenome sequence from MHS_10 and CS_12 (**Figure A1** in Appendix). The recently released *Hydrogenobaculum* spp. genomes isolated from *Dragon Spring* (Romano et al., 2013) were not compared to DS_9 assemblies here, but these reference genomes are likely superior (relative to strain Y04AAS1) for direct comparison to *Hydrogenobaculum*-like sequence from DS_9.

PHYLOGENETIC ANALYSIS OF METAGENOME SEQUENCE ASSEMBLIES

The *de novo* assemblies from each Aquificales community resulted in significant consensus sequence of indigenous populations present in these environments, regardless of whether the Celera (Rusch et al., 2007) or the PGA assembler (Zhao et al., 2008) was used. A similar number and distribution of contigs was generated from the two assemblies. In several cases, the amount of consensus sequence corresponding to the major Aquificales populations approached the expected size of reference Aquificales genomes (Reysenbach et al., 2009). Ordination of nucleotide word frequency analysis (NWF-PCA, or k-mer analysis) (Teeling et al., 2004) has been shown to be an excellent criterion for determining whether contigs generated during assembly may belong to the same species, and was used to evaluate all Aquificales contigs greater than 2000 bp (**Figure 3**). Similar Aquificales populations were observed in each of the replicate field sites: DS_9 (yellow) and OSP_14 (red); MHS_10 (green) and CS_12 (violet), and OS_11 (dark-blue) and BCH_13 (light-blue), respectively. Moreover, clear separation in Aquificales sequence was observed across sites with major geochemical differences (**Figure 3A**).

Phylogenetic analysis of these sequence clusters using the Automated Phylogenetic Identification System (APIS; Badger et al., 2006) in identical PCA orientation shows that the majority of Aquificales contigs from sites DS_9 and OSP_14 were most closely related to *Hydrogenobaculum* sp. Y04AAS1 (**Figure 3B**). Sequence clusters from MHS_10 consistently showed highest similarity to either the *Sulfurihydrogenibium* sp. Y03AOP1 or *S. yellowstonense* reference genomes. CS_12 contained Aquificales contigs most

FIGURE 3 | Nucleotide word frequency PCA plots of metagenome assemblies from six Aquificales streamer communities in YNP. **(A)** Data colored by site (black circles): DS_9 = yellow and OSP_14 = red; MHS_10 = green and CS_12 = violet; OS_11 = dark-blue and BCH_13 = light-blue. **(B)** Identical PCA orientation with phylogenetic analysis to the closest reference genomes (white-dashed circles): orange = *Hydrogenobaculum,* green = *Sulfurihydrogenibium,* dark-blue = Aquificaceae (no reference genomes adequately describe the similar *Thermocrinis*-like Aquificales phylotypes observed in OS_11 and BCH_13).

similar to the *Sulfurihydrogenibium* sp. genomes, as well as a separate population showing greater identity to the *Aquifex aeolicus* VF5 genome (i.e., Aquificaceae). In the absence of a reference strain with higher sequence identity, the *Aquifex aeolicus* VF5 genome served to represent this *Thermocrinis*-like population (see also **Figure 2**).

Contigs from the higher-pH (pH ∼8), non-sulfidic flow channels (OS_11 and BCH_13) cluster together and were also classified at the family level, Aquificaceae (**Figure 3**). As discussed above, the Aquificales populations within OS_11 and BCH_13 are actually related to *Thermocrinis* spp., as determined from phylogenetic analysis of 16S rRNA and other genes present within each of the metagenome assemblies (**Figure 4**). However, there are currently no adequate reference strains for comparison to these YNP populations. For example, the recently released *T. albus* genome is not a good reference for the YNP *Thermocrinis* populations from OS_11 or BCH_13, based on poor nucleotide identity and lack of similar gene content (**Figure A2** in Appendix).

The composition of each community was also evaluated using phylogenetic analysis of 16S rRNA genes detected within the assembled sequence. Of 407 16S rRNA genes identified by the integrated metagenome data management and comparative analysis system (IMG/M, Markowitz et al., 2012) from all six sites, 64 were of sufficient length (>1100 bp) for phylogenetic analysis. Aquificales 16S rRNA gene sequences dominated the dataset (**Figure 4**) and grouped with *Hydrogenobaculum* (DS_9 and OSP_14), *Sulfurihydrogenibium* (MHS_10 and CS_12), or *Thermocrinis* spp. (OS_11, BCH_13, CS_12). Unclassified 16S rRNA gene sequences grouped with novel DNA sequences reported in previous studies (Reysenbach et al., 1994; Blank et al., 2002; Hall et al., 2008) and represent members of potentially deeply rooted candidate phyla (Hall et al., 2008). Finally, PCR-amplified bacterial and archaeal 16S rRNA gene clone libraries largely agreed with the diversity detected within the assembled metagenome sequence (**Figures A3**

FIGURE 4 | Phylogenetic tree of Aquificales 16S rRNA gene sequences observed within assembled metagenome sequence from six YNP geothermal outflow channels [DS_9 (orange), OSP_14 (red), MHS_10 (green), CS_12 (violet), OS_11 (dark-blue), BCH_13 (light-blue) [neighbor-joining tree with 1000 bootstraps].
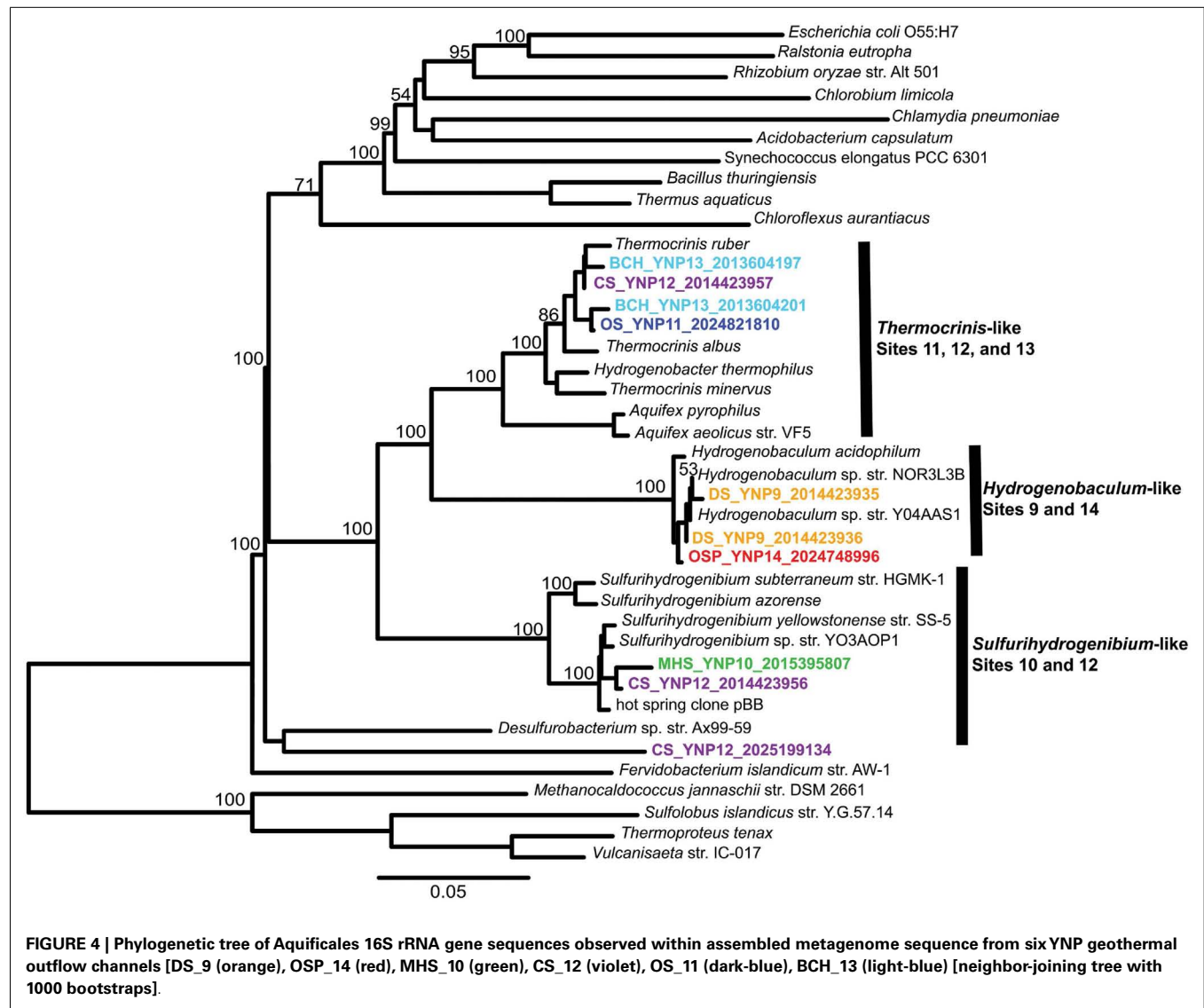
and **A4** in Appendix). Although archaeal primers failed to amplify from several samples, metagenome sequence contained significant levels of different archaea (**Figure 2**).

To broaden our phylogenetic analysis, the community composition was also investigated by assigning phylotypes of 31 conserved protein-encoding marker genes using AMPHORA (Wu and Eisen, 2008). Of the 159,636 predicted genes among the six metagenomes, 994 were protein markers detected by AMPHORA and of sufficient length for phylogenetic analysis. At least 29 of the 31 housekeeping genes included in the AMPHORA database were detected in each of the six sites. The majority of marker genes for each site were assigned to members of the Aquificales, which reflects the greater amount of assembled Aquificales-like sequence across these sites, relative to other organisms (**Figure A5** in Appendix). Significant amounts of assembled sequence corresponding to members of the Firmicutes, Alphaproteobacteria, Thermotogae, Bacteroidetes, Spirochetes, Proteobacteria, Actinobacteria,

Deinococcus-Thermus, Acidobacteria, Gammaproteobacteria, Planctomycetes, Chlamydiae, Betaproteobacteria, Fusobacteria, and Epsilonproteobacteria, as well as unclassified *Bacteria* and *Archaea* was predicted using AMPHORA (**Figure A5** in Appendix). A minor number of sequence matches to Chloroflexi, Chlorobi, and/or Cyanobacteria may reflect exogenous inputs from lower temperatures. The version of AMPHORA used here suggests greater diversity than indicated by other more direct phylogenetic analyses of either individual sequences, assembled contigs, or PCR-amplified 16S rRNA genes (e.g., **Figures 2** and **3**; **Figures A2–A4** in Appendix), and is due in part to a limited number of appropriate reference genomes included in the AMPHORA database, especially from thermophiles and other deeply rooted, uncharacterized lineages. This prohibits accurate phylogenetic placement of a subset of these sequences. Consequently, the actual abundance of each population in these samples is better understood in terms of total sequence reads (i.e., **Figure 2**).

Population-level community richness can be inferred from the average abundance of single-copy genes detected within a specific population, assuming sufficient coverage (Table S1 in Supplementary Material). Given that a majority of the AMPHORA protein markers were detected for the Aquificales populations present in each site, the abundance of single-copy genes should provide a reasonable estimate of population-level heterogeneity. AMPHORA markers suggested that MHS_10 is dominated by a single *Sulfurihydrogenibium* sp., while CS_12 averaged five distinct *Sulfurihydrogenibium*-related copies of each single-copy gene (Table S1 in Supplementary Material). The acidic communities (DS_9 and OSP_14) may be comprised of three distinct *Hydrogenobaculum*-like populations, whereas OS_11 and BCH_13 contained at least eight or nine novel *Thermocrinis*-related populations. The number of Aquificales 16S rRNA genes identified in the metagenome assemblies (**Figure 4**) did not always agree with the number of Aquificales-related single-copy genes detected using AMPHORA. For example, 13 *Sulfurihydrogenibium*-like 16S rRNA genes were detected in MHS, yet the average abundance of single-copy genes suggest that this site is dominated by only one *Sulfurihydrogenibium*-like population. This discrepancy may be attributed to the fact that 16S rRNA genes do not often assemble with the same consistency as other housekeeping genes (Rusch et al., 2007), and/or to the possible presence of multiple rRNA operon copies (sequenced *Sulfurihydrogenibium* genomes have two to four annotated 16S rRNA gene copies). Although the sequencing depth obtained in the current study is not sufficient to accurately assess these communities at the ecotype level (Ward et al., 2006), it is clear that sequence variants of highly related Aquificales populations are evident within each site.

FUNCTIONAL ANALYSIS OF "STREAMER" COMMUNITIES IN YNP

The functional gene content among the Aquificales "streamer" communities was compared using multivariate statistical analysis of protein family (TIGRFAM) abundances based on all predicted proteins from assembled metagenome sequence. To study different driving forces behind functional diversity, the analysis was performed using both the full set of all TIGRFAM protein families and a set of TIGRFAMS associated with electron transport (ET) functions. Both analyses reveal contributions from the predominant Aquificales lineage(s) across sites as well as the variable co-community members observed within each individual site (e.g., **Figure 2**). Site comparisons were made using PCA of all TIGRFAMS, which show strong support for the similarities between site pairs in acidic (DS_9, OSP_14) and circum-neutral sulfidic systems (MHS_10, CS_12) (**Figure 5A**). The first principal component explained over 57% of the total variation in protein family abundance among the different "streamer" communities and separated the two predominant families of Aquificales represented across these sites (Aquificaceae include both the *Hydrogenobaculum* sp. that dominate sites DS_9 and OSP_14, and the *Thermocrinis*-like populations in OS_11 and BCH_13). Also, these four sites all contained archaeal co-community members (e.g., ∼30% of sequence reads in OSP_14 and DS_9, ∼20% in OS_11, and ∼8% in BCH_13). Conversely, sites MHS_10 and CS_12 were dominated by *Sulfurihydrogenibium* populations in

the family Hydrogenothermaceae. Despite the similar *Thermocrinis*-like populations present in OS_11 and BCH_13, these communities exhibited considerable functional differences, which ultimately tracked with differences in co-community members. Several of the phylogenetically distinct bacterial populations observed in OS_11 and BCH_13 (see **Figure 2**), which are not well characterized, contribute to unexpected functional differences between these communities. Factor 2, which accounted for ∼22% of variation in relative TIGRFAM abundance across these six sites, separated the low-pH *Hydrogenobaculum* sites from higher-pH sites.

Principal components analysis (PCA) using a subset of "ET" TIGRFAMS shows a slightly different pattern in site separation (**Figure 5B**) that is more sensitive to specific electron donors and acceptors (i.e., respiratory pathways), and more consistent with geochemical differences across sites (**Table 1**). Factor 1, which represented 63% of variation across sites, separated sites based on pH (low-pH sites DS_9 and OSP_14 versus other four sites), while Factor 2 correlated strongly with the presence or absence of sulfide, and separated the more oxic sites OS_11 and BCH_13 from sulfidic sites. The third principal component (∼6% of functional variation) emphasized differences in community composition between the two oxic sites (OS_11 and BCH_13), and did not contribute to overall functional differences in the more sulfidic systems (DS_9, OSP_14, MHS_10, CS_12).

Hierarchical cluster analysis of the "ET" TIGRFAMs (**Figure 6**) show expected site pairing based on geochemical attributes identified in the study design (Inskeep et al., 2013), and can be used to identify specific protein families that explain the overall variation seen in PCA (**Figure 5B**). These TIGRFAMs included ET and terminal oxidase proteins specific to different respiratory pathways dependent on either sulfur, arsenite, hydrogen, and oxygen as well as other cytochromes, ferrodoxins, and flavoproteins (**Figure 6**; Table S2 in Supplementary Material). For example, while MHS_10 and CS_12 sites were both dominated by *Sulfurihydrogenibium* sp., the secondary *Thermocrinis* (family Aquificaceae*) and *Thermus* populations in CS_12 contributed a significant number of additional cytochrome families that are not present in *Sulfurihydrogenibium* sp. Similarly, the higher complexity community at OS_11 contained several additional gene families compared to the lower complexity *Thermocrinis*-dominated community at BCH_13.

Hierarchical cluster analysis using broader TIGRFAM profiles (**Figure A6** in Appendix) supports the relative similarity of acidic sites (DS_9 and OSP_14) and circum-neutral sulfidic sites (MHS_10 and CS_12); however, the higher-pH non-sulfidic (i.e., oxic) communities (OS_11, BCH_13) did not form a separate group. Despite similar *Thermocrinis* populations in these two communities (e.g., see **Figure 2**), the overall functional profiles are quite different and this is consistent with the diverse and novel bacterial phylotypes observed in these sites (especially OS_11). Detailed analysis of specific functional categories across sites reveal differences in the relative abundance of genes coding for numerous cellular functions including nucleotide and DNA metabolism, regulatory functions, energy metabolism, central C metabolism, mobile elements, transcription, cofactors, and transporters (**Figure A6** in Appendix).
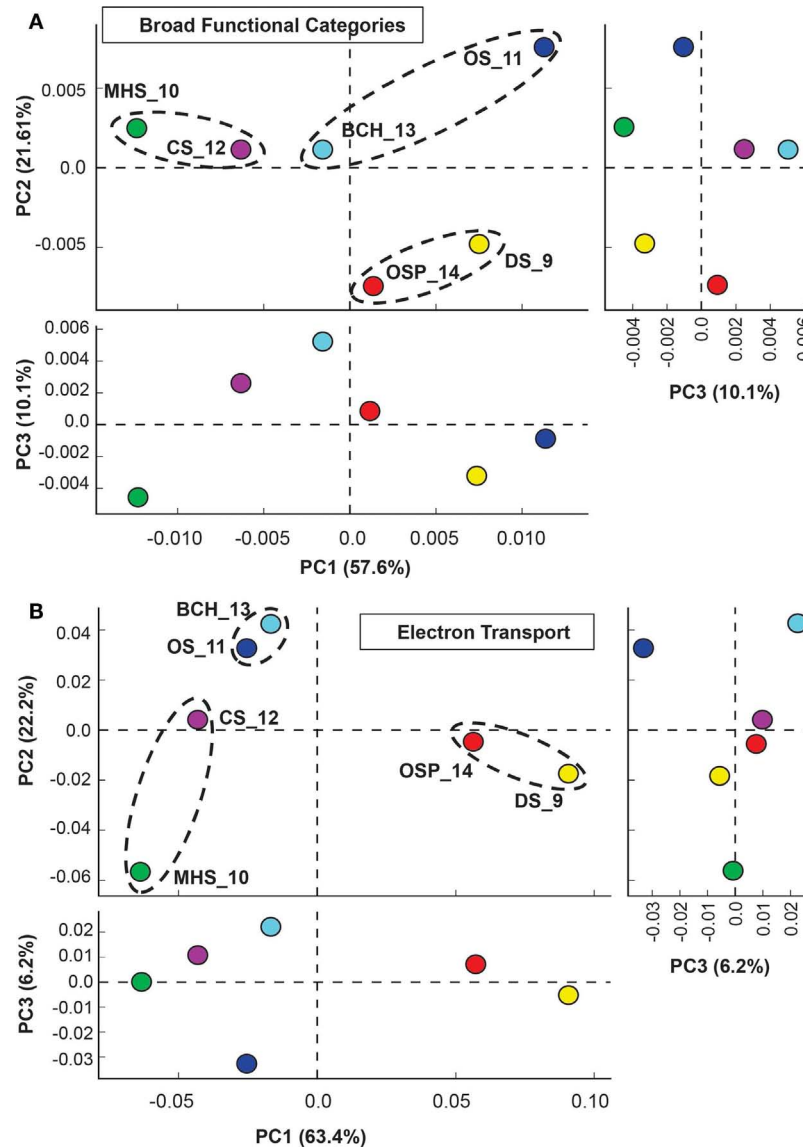
FIGURE 5 | Principal components analysis (PCA) of relative gene abundances across six Aquificales streamer communities. **(A)** All TIGRFAMs grouped into functional categories, and **(B)** Only those TIGRFAMs associated with the role category "Electron Transport." Site colors: DS_9 (yellow), OSP_14 (red), MHS_10 (green), CS_12 (violet), OS_11 (dark-blue), BCH_13 (light-blue). Circled sites represent "replicates" that contain one of three dominant Aquificales lineages (*Hydrogenobaculum*, *Sulfurihydrogenibium*, or *Thermocrinis*).

The TIGRFAMs with the greatest statistical site separation (lowest *p*-value) include the pyridine nucleotide biosynthesis and PTS signal transduction categories (see **Figure A7** in Appendix and Table S3 in Supplementary Material) for a complete list of TIGRFAM *p*-values). A coarse view of the importance of specific geochemical variables such as pH can be examined by looking at specific TIGRFAMs providing the highest statistical separation of sites based on selected variables. For example, the most pH-dependent TIGRFAMs include categories such as glutathione disulfide reductases, thioredoxin-disulfide reductases, formate dehydrogenases, and NADH-plastoquinone oxidoreductases (**Figure A8** in Appendix).

FUNCTIONAL DIFFERENCES AMONG AQUIFICALES LINEAGES IN YNP
Fixation of carbon dioxide

A detailed inventory of genes coding for carbon dioxide fixation and various oxidation/reduction pathways indicated major functional differences among the three primary Aquificales lineages observed in this study (**Table 2**). The reverse-TCA pathway is thought to be the earliest CO_2 fixation process used by microorganisms (Wachtershauser, 1990; Hugler et al., 2005). Although members of the Aquificales have been shown to utilize this pathway (Aoshima et al., 2004; Ferrera et al., 2007), the enzymes associated with the key step (citrate cleavage) differ among families of the cultured Aquificales. For example, *Sulfurihydrogenibium* spp. (Family

FIGURE 6 | Hierarchical cluster analysis of relative gene abundances in the TIGRFAM role category "Electron Transport" across six Aquificales streamer communities. TIGRFAMs with low variation across the sites were removed before the clustering to retain ~50 of the most variable families. Subunits of the protein complexes were only represented by one representative TIGRFAM family. Pearson correlation was used as the distance measure for average linkage agglomerative clustering. Sites cluster consistent with pH and the presence or absence of sulfide.

Table 2 | Summary of key metabolic genes identified in metagenome sequence of the predominant Aquificales populations present in six different environmental systems covering a wide range in pH, dissolved sulfide, and dissolved oxygen.

Process/Pathway	Marker Genes[1]	Dragon Sp. (DS)	100 Spring Plain (OSP)	Mammoth (MHS)	Calcite Sp. (CS)	Bechler (BCH)	Octopus Sp. (OCT)
Fixation of CO$_2$, Reductive TCA							
ATP citrate lyase	*aclB*	0	0	1	2	0	0
Citryl coA lyase	*ccl*	1	1	0	1	1	1
Citryl coA synthetase	*ccsA*	3	1	0	1	2	2
***Oxidation Reduction Reactions*[2]**							
Thiosulfate Oxidase	*tqoAB*	1	1	0	0	0	0
Hydrogen Oxidation (Group 1 Ni-Fe)	*hynS, hynL*	2	2	0	0	0	0
Arsenite Oxidation	*aroA*	1	1	0	0	1	1
Oxidation of Reduced S (Elemental S)	*hdrA/hdrB*	1,1	1,2	0,2	1,4	1,2	1,3
Oxidation of Reduced S (Sulfide)	*sqr*	2	2	2	6	1	1
Oxidation of Reduced S (S-transferase)	*rdh*	4	4	6	10	5	8
Sulfite Oxidase	*soxB?C*	1	1	1	3	1	1
Oxygen Reduction, Heme Cu Oxidases	*coxA*, other	1	1	2	1	1	2
Oxygen Reduction, bd-ubiquinol type		2	1	1	1	0	0
Dissimilatory Reduction of Sulfur	*psrA/sreA*[3]	2	2	1	1	0	0
Trace Element Detoxification	*arsB, arsC, arsM*	1,4,2	2,2,2	2,1,1	4,4,7	1,6,2	6,5,4
	merA	0	0	0	0	0	1
Heavy-metal translocating ATPases	*zntA*	3	4	0	8	3	0

[1] *Marker genes code for proteins with high specificity for possible pathway.*

[2] *No genes were found for nitrification (amoA), denitrification (e.g. nirK, nirS, norB), methanogenesis or methanotrophy (e.g. mcrA), sulfate reduction (e.g. dsrABI), or arsenate reduction (e.g. arrAB) within the Aquificales populations represented in these six sites.*

[3] *Includes Mo-pterin proteins similar to sreA and arrA.*

Hydrogenothermaceae) catalyze citrate cleavage using ATP citrate lyase, which includes a large and small subunit (AclA and AclB, respectively). Conversely, *Thermocrinis* spp. (Family Aquificaceae) catalyze citrate cleavage using two separate enzymes, citryl-CoA synthetase (Ccs) and citryl-CoA lyase (Ccl). We detected the *aclB* gene in both *Sulfurihydrogenibium* sites (MHS_10 and CS_12), but not in any sites containing strictly Aquificaceae (DS_9, OSP_14, OS_11 and BCH_13; **Figure 7**). Although CS_12 contained both *Thermocrinis* and *Sulfurihydrogenibium* populations, all *aclB* genes identified from this site grouped unambiguously with the *Sulfurihydrogenibium* sp. (**Figure 7**, bootstrap support = 92%).

Genes coding for citryl-CoA synthetase (*ccs*A) and citryl-CoA lyase (*ccl*) were found in both the *Hydrogenobaculum* (DS_9 and OSP_14) and *Thermocrinis*-like (OS_11, CS_12 and BCH_13) populations (**Figures 8A,B**). Therefore, these members of the Aquificaceae appear to fix CO_2 using this alternative citrate cleavage mechanism. Deduced protein sequences (CcsA and Ccl) from OS_11, BCH_13, DS_9, and OSP_14 grouped in distinct clades consistent with the pH differences among sites, as well as the different genera observed within this family (i.e., DS and OSP versus OS and BCH). The *ccs*A gene copy detected in CS_12 grouped with similar entries in OS_11 and BCH_13, and was confirmed to come from the sub-dominant *Thermocrinis*-like population in this site. *Sulfurihydrogenibium*-like populations in MHS_10 and CS_12 lacked citryl-CoA synthetase (*ccs*A) genes, and instead contained one copy of a succinyl-CoA synthetase (**Figure 8A**). Differences in the gene neighborhood between the citryl-CoA and succinyl-CoA synthetase pathways of the Aquificaceae versus Hydrogenothermaceae suggest that the *Sulfurihydrogenibium* copy of succinyl-CoA synthetase from MHS is not involved in CO_2 fixation. Annotation inconsistencies among these two fairly similar proteins (CcsA and succinyl-CoA synthetase) have made it difficult to make definitive assignments without visualizing the sequences in phylogenetic trees (**Figure 8A**) or other alignment tools, and these proteins are in fact thought to be related via gene duplication (Aoshima et al., 2004).

The r-TCA pathway has not actually been demonstrated in any cultured member of the *Hydrogenobaculum*, although field data on $^{14}CO_2$ incorporation suggests that members of these communities fix CO_2 at significant rates (Boyd et al., 2009). Genes found in DS_9 and OSP_14 are divergent relative to the *Thermocrinis* entries (**Figures 8A,B**), so it is unclear if these perform an identical function in both genera. However, other evidence that the r-TCA pathway is operative in the DS_9 and OSP_14 *Hydrogenobaculum*-like populations includes two enzymes required for the reductive pathway: 2-oxoglutarate:ferredoxin oxidoreductase and fumarate reductase. Both genes are present in the *Hydrogenobaculum* sp. Y04AAS1 genome as well as the DS_9 and OSP_14 metagenomes.

Oxidation of H₂, reduced sulfur, and arsenite. Other functional differences among the major Aquificales lineages include possible explanations for the specialization of these populations to specific geochemical environments. For example, *Hydrogenobaculum* populations from DS_9 and OSP_14 contain Group I Ni-Fe hydrogenases, but these genes are notably absent from the *Sulfurihydrogenibium* (MHS_10 and CS_12) and the *Thermocrinis* populations (OS_11 and BCH_13) (**Table 2**). The potential for H_2
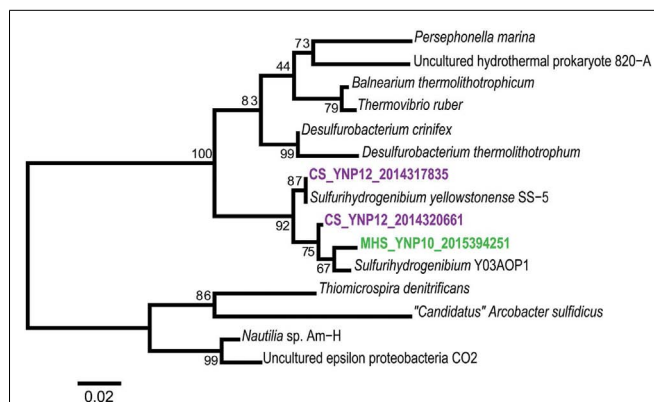
FIGURE 7 | Deduced protein tree of ATP citrate lyase (AclB), a key marker for CO₂ fixation via the reverse-TCA cycle found in *Sulfurihydrogenibium* sp. in MHS_10 and CS_12 (neighbor-joining tree with 1000 bootstraps).

to serve as an electron donor for metabolism appears limited to the acidic sites where concentrations of aqueous H_2 have been measured in the 50–100 nM range (Inskeep et al., 2005; Spear et al., 2005), while other sites contain lower H_2 (**Table 1**). The *Hydrogenobaculum*-like populations were also the only Aquificales to contain genes coding for thiosulfate oxidase (*tqoAB*), often implicated in oxidation of thiosulfate in the order Sulfolobales (Friedrich et al., 2005). However, it is known that thiosulfate concentrations are considerably higher in circum-neutral sulfidic sites, due to the greater stability of thiosulfate at intermediate pH (Xu et al., 1998; Nordstrom et al., 2005). The *Sulfurihydrogenibium*-like organisms may process thiosulfate through an abundance of rhodanese domain proteins known to be involved in sulfur-transferase reactions as well as SoxBC complexes (Friedrich et al., 2005). Moreover, all Aquificales lineages from each site contained either one or more copies of a highly conserved and syntenous gene complex thought to be important in the oxidation of reduced sulfur (i.e., sulfide and elemental S), and includes several hetero-disulfide reductases as well as other Fe-S proteins (*rhd, tusA, dsrE, hdrC, hdrB, hdrA, orf2, hdrC, hdrB*) (Ghosh and Dam, 2009). Each of the three lineages also contained *sqr* (sulfide:quinone reductase) genes (**Table 2**), which have been shown to code for proteins involved in the oxidation of dissolved sulfide to S^0 or polysulfide chains, followed by electron transfer to the quinone pool through a flavin adenine dinucleotide (FAD) cofactor (Cherney et al., 2010). Even the *Thermocrinis* populations from OS_11 and BCH_13 exhibited potential for the oxidation of sulfide and elemental S, although it is unlikely that sufficient sulfide exists in these geothermal channels to support the growth of active "streamer" communities. The habitat range of *Thermocrinis*-like organisms in YNP includes high-pH (7–9) sulfidic channels (Inskeep et al., 2005; Hall et al., 2008; Planer-Friedrich et al., 2009) and thus may explain the presence of these genes in *Thermocrinis* assemblies. The HDR gene complexes appear highly conserved across numerous Aquificales and Sulfolobales, as well as acidophilic bacteria such as *Acidithiobacillus ferrooxidans* (Ghosh and Dam, 2009; Inskeep et al., 2013).

FIGURE 8 | Deduced protein tree of (A) citryl-coA synthetase (CcsA), and (B) citryl-CoA lyase (Ccl) showing metagenome entries from Aquificales sites (neighbor-joining tree with 1000 bootstraps).
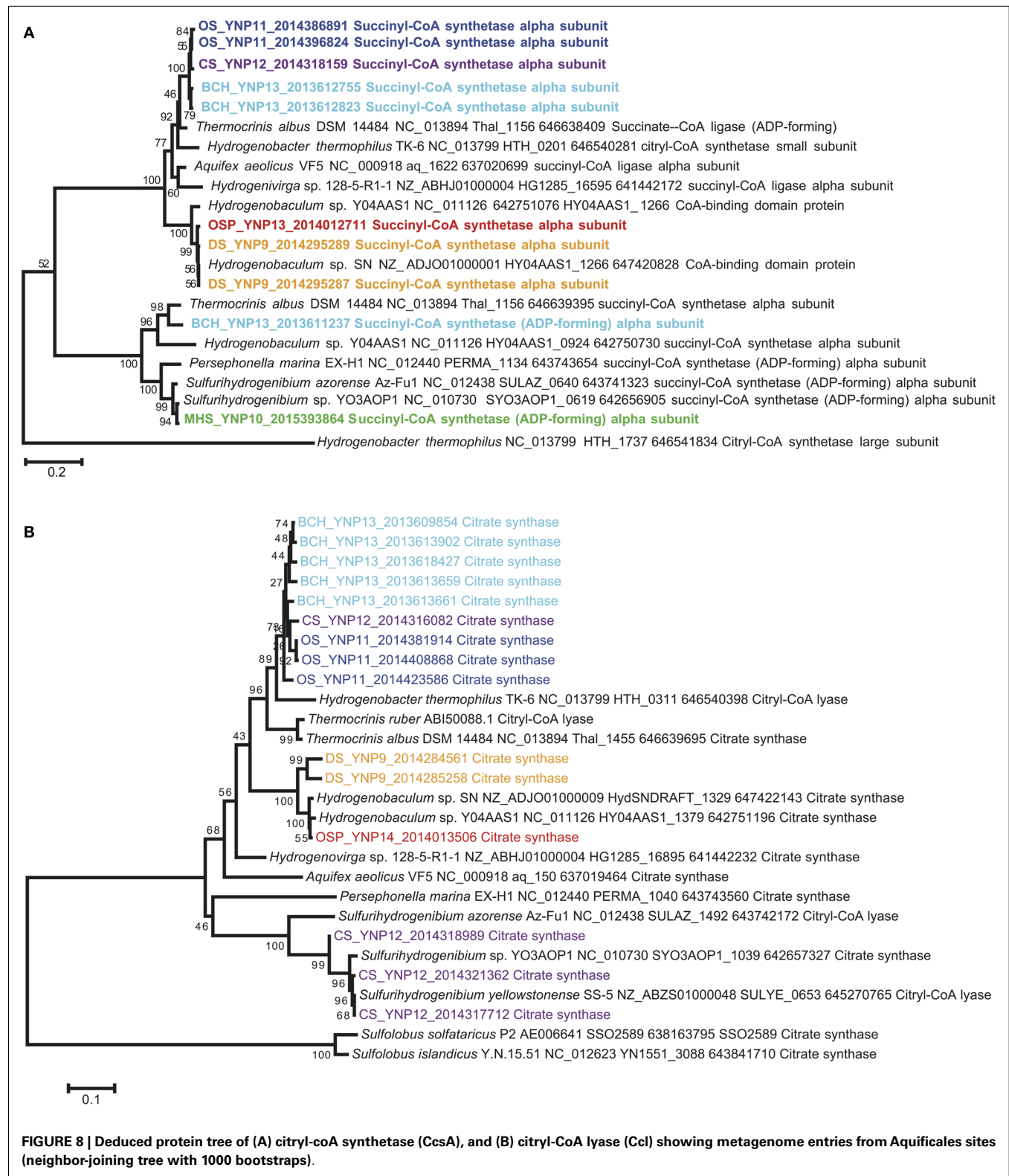
The oxidation of arsenite to arsenate is highly exergonic (ranging from 50 to 60 kJ/mol electron) in these geothermal systems (Inskeep et al., 2005) and has been shown to serve as a sole electron donor in several unrelated bacteria (D'Imperio et al., 2007; Santini et al., 2007). Consequently, it is interesting that both the *Hydrogenobaculum* and *Thermocrinis*-like organisms from sites

DS_9, OSP_14, OS_11, and BCH_13 contain full copies of the arsenite oxidase Mo-pterin subunit I (*aio*A, also abbreviated *aro*A, *aso*A, *aox*B in prior work). *Thermocrinis* and *Hydrogenobaculum* spp. oxidize considerable amounts of arsenite in acidic to circumneutral springs in YNP (Macur et al., 2004; Inskeep et al., 2005; Hamamura et al., 2009), corresponding to the oxygenation of geothermal outflow channels and correlation with *aio*A expression in these same habitats (Clingenpeel et al., 2009; Hamamura et al., 2009). It is possible that members of the Aquificales gain energy from the oxidation of arsenite *in situ*. However, this has not been established in culture (Donahoe-Christiansen et al., 2004). All Aquificales populations from the current study also contained evidence for arsenic detoxification, including the potential to reduce arsenate via ArsC and extrude arsenite via an efflux pump (AsrB), as well as to methylate arsenite via methyl transferases (ArsM) (Bentley and Chasteen, 2002; Mukhopadhyay et al., 2002).

Respiratory processes.   The presence of terminal oxidase complexes in each of the Aquificales populations suggests that these organisms all respire oxygen. However, the distribution of different types of subunit I heme Cu oxidases (HCOs) across the Aquificales lineages (**Figure 9**) is not consistent with the distribution of CO_2 fixation genes. This observation invokes a different evolutionary history of carbon dioxide fixation versus aerobic respiration among these Aquificales lineages. The *Sulfurihydrogenibium* and *Hydrogenobaculum*-like organisms contain similar Type C-Cbb3 HCOs (Garcia-Horsman et al., 1994) despite the phylogenetic distance between these organisms (i.e., different families). Conversely, the *Thermocrinis* spp. from OS_11 and BCH_13 contained multiple copies (at least two distinct copies per site) of Type A HCOs (Pereira et al., 2001), which are phylogenetically related to HCOs from the majority of aerobic bacteria.

Hydrogenobaculum (DS_9, OSP_14) and *Sulfurihydrogenibium*-like (MHS_10, CS_12) organisms were the dominant Aquificales in sites containing high levels of dissolved sulfide (low dissolved O_2). Consequently, the Cbb3 cytochromes associated with these populations likely bind O_2 with greater efficiency compared to Type A HCOs of the *Thermocrinis*-like populations, and is consistent with known properties of these proteins (Garcia-Horsman et al., 1994). The Cbb3-type heme-copper oxidases are found only in several groups of bacteria, especially the Aquificales and Proteobacteria (312 of 365 sequences; Sousa et al., 2011), and appear to represent the evolution of a separate respiratory complex in low O_2 environments (Garcia-Horsman et al., 1994). In addition, the *Hydrogenobaculum* and *Sulfurihydrogenibium*-like organisms have genes coding for a *bd*-ubiquinol oxidase (*cydAB*), also thought to function under lower O_2 concentrations (Jünemann, 1997). The *Thermocrinis*-like organisms in OS_11 and BCH_13 show no evidence of the *bd*-ubiquinol oxidases, consistent with the higher oxygen levels and low sulfide in these sites.

A detailed survey of other respiratory processes (dissimilatory reduction) suggests that the reduction of elemental sulfur and/or polysulfide is important in some members of the Aquificales. Although some sequenced Aquificales isolates contain a nitric oxide reductase (NorB) (one of the steps required for complete denitrification from nitrate to N_2), *nar*G, *nor*B, *nir*K, or *nir*S like sequences were not found in the YNP Aquificales. Moreover, no evidence was found for dissimilatory reduction of sulfate or sulfite (*dsr*AB), arsenate (*arr*AB), or CO_2 (*mcr*A, methanogenesis) in any of the Aquificales lineages. However, the organisms found in sulfidic channels (i.e., *Hydrogenobaculum* and *Sulfurihydrogenibium*) all contain sulfur reductases (*sre*A) or polysulfide reductase (*psr*A), as well as tetrathionate reductases (*ttr*A, another DMSO-Mopterin). Consequently, the Aquificales lineages that inhabit sulfidic channels under low O_2 tensions may require electron shuttling to reduced sulfur instead of, or in addition to, O_2. Genes for sulfur reduction were notably absent in the *Thermocrinis*-like populations of OS_11 and BCH_13, which correlates with the lack of elemental sulfur in these springs.

DISCUSSION

Phylogenetic and functional analysis of metagenome sequence from three major types of high-velocity, filamentous "streamer" communities revealed three lineages of Aquificales, whose metabolic potential correlated primarily with pH and sulfide and/or elemental sulfur. Sites with low-pH (pH 3–3.5) and high-sulfide contained *Hydrogenobaculum* spp., whereas higher-pH sites were dominated by either *Sulfurihydrogenibium* spp. (high-sulfide) or *Thermocrinis*-like (low sulfide) populations. *Calcite Springs* (CS_12) also hosted a minority *Thermocrinis*-like population and was the only site here to contain two major Aquificales genera. This is consistent with previous 16S rRNA gene diversity surveys that have generally found only minor overlap in the distribution of different Aquificales across YNP geothermal environments (Reysenbach et al., 2005; Hall et al., 2008; Hamamura et al., 2009). *Thermocrinis* organisms have also been observed in sulfidic channels at higher-pH, near 9 (Planer-Friedrich et al., 2009).

Metagenome sequence assemblies for each of the three major Aquificales lineages resulted in total scaffold sizes that approach full genomes, and which represent "consensus sequence" or "pan-genomes" of these populations (Medini et al., 2005). Sequence variability of highly related populations within a field site may contribute to incomplete assembly, and "closure" of these *de novo* assemblies would require considerable manual effort, as well as additional sequencing to close gaps. Sequence heterogeneity within individual Aquificales populations was observed using AMPHORA to detect the number of single-copy genes. Although the streamer community from MHS_10 was dominated by what appears to be a fairly homogeneous population type of *Sulfurihydrogenibium* sp., other sites exhibited greater variability within the primary Aquificales population. For example, the *Thermocrinis*-like populations from *Octopus*, *Bechler*, and/or *Calcite Springs* all revealed higher numbers of numerous single-copy genes, which suggests greater heterogeneity of these populations *in situ*. Additional sequence coverage of these populations may result in less single-copy gene variability, and future efforts will be necessary to clarify sources of this variability. Importantly, the sequence assemblies generated in this study provide a foundation for future efforts to determine the types and rates of genetic change in these same sites.

In addition to the abundant Aquificales populations, *de novo* assemblies were also obtained for several novel bacterial and archaeal lineages, although at lower coverage. These co-community members provide an interesting comparative study in

FIGURE 9 | Protein tree of heme-copper oxidases (subunit I of terminal oxidase complex). Metagenome entries are highlighted by site, and labels correspond to the major types of heme-copper oxidases observed in different Aquificales "streamer" communities. Similar heme-copper oxidases are found in sites OS_11 and BCH_13, and these are significantly different than the terminal oxidases found in *Hydrogenobaculum* and or *Sulfurihydrogenibium* from sites DS_9, OSP_14, MHS_10 and CS_12 (neighbor-joining tree constructed using nitric oxide reductase (NorB) as the out group; 1000 bootstraps).

their own right given that each Aquificales community exhibited a different assemblage of interacting community members. For example, the acidic sites contained different archaeal populations representing specialization on Fe(II) (e.g., *Metallosphaera*-like) versus reduced sulfur (Thaumarchaeota- and Thermoplasmatales-like), while the higher-pH (pH ∼8) sites contained several novel bacterial lineages, and a population of Thermoproteaceae (G + C ∼ 60%) in OS_11 and BCH_13. Some of the novel bacterial populations present in OS_11 and BCH_13 were related to unclassified 16S rRNA gene sequences described previously (Reysenbach et al., 1994; Blank et al., 2002; Hall et al., 2008). Metagenomics provides a promising opportunity to gain insight into the metabolic potential of these novel populations, although additional sequencing will be required to bring the coverage up to levels suitable for building reasonable *de novo* assemblies (e.g., >2× coverage).

Global protein (TIGRFAM) analysis coupled with PCA and hierarchical clustering of all Aquificales "streamer" communities demonstrated important linkages among geochemistry, the presence of distinct phylotypes, and their metabolic genes. Clustering of TIGRFAMs specific to ET demonstrated specific linkages between individual phylotypes and respiratory pathways that are consistent with the strong influence of geochemistry on community structure (especially pH and sulfide). Perhaps one of the more interesting findings in the study is the degree to which co-community members varied across these high-velocity stream-channel habitats. Even site pairs containing the same Aquificales phylotype contained considerably different interacting populations. Factors controlling the variation in community structure in different sites containing the same Aquificales phylotype can also be shown to track with geochemical conditions. Such is the case when comparing the hypoxic elemental sulfur habitats (DS_9) to the more oxic Fe(III)-oxide (OSP_14) streamer communities, samples that both contained a very similar *Hydrogenobaculum* population. The archaeal co-community members in DS_9 are likely anaerobic (or microaerophilic) populations compared to the aerobic Fe(II)-oxidizing Sulfolobales in OSP_14.

Detailed analysis of functional genes present in the three Aquificales lineages revealed several examples of divergence that were likely driven by environmental selection and lateral transfer events. These events have resulted in inconsistent patterns in phylogeny between highly conserved genes and those coding for various functional processes. Characteristics of the r-TCA cycle make it a good candidate for analysis of early steps important in the evolution of autotrophy. The pathway's auto-catalytic nature and role in central C metabolism may reflect its importance in the evolution of associated anabolic and oxidative pathways (Hugler et al., 2005). The key ATP-dependent step of acetyl CoA synthesis is catalyzed by different proteins in the Hydrogenothermaceae versus Aquificaceae (AclB versus CcsA and Ccl, respectively; Aoshima et al., 2004), suggesting different evolutionary histories of these lineages with respect to CO_2 fixation. The only Aquificales in this study to contain *acl*B were the *Sulfurihydrogenibium*-like organisms (Family Hydrogenothermaceae) present in MHS and CS (**Figure 7**). The *Hydrogenobaculum* and *Thermocrinis* populations (Family Aquificaceae) in the other sites contain *ccs*A and *ccl* genes rather than *acl*B (**Figures 8A,B**). The r-TCA pathway catalyzed by Ccs and Ccl is believed to be the ancestral pathway among the Aquificales

(Hugler et al., 2005), however, our understanding of this pathway is not complete as evinced by inconsistent annotation of both the citryl-CoA synthetase (*ccs*A) and the citryl-CoA lyase (*ccl*).

Although the Aquificales occupy diverse geochemical habitats, they generally flourish in zones of shallow, high-velocity, turbulent spring water where dynamic mixing occurs to create disequilibria between hypoxic and oxic conditions. The exchange of atmospheric oxygen and subsequent reaction with dissolved sulfide present in thermal waters is one of the more important geochemical processes occurring within sulfidic geothermal channels (Inskeep et al., 2005; Nordstrom et al., 2005). The disequilibrium between oxygen and sulfide establishes conditions suitable for microaerophiles growing on sulfide, elemental sulfur or thiosulfate (Reysenbach et al., 2009). Metagenome sequence of the YNP Aquificales clearly indicates the potential for the oxidation of reduced sulfur species using a variety of S-oxidation pathways coupled with Type C-Cbb3 or *bd*-ubiquinol terminal oxidase complexes, especially in the *Hydrogenobaculum* and *Sulfurihydrogenibium*-like organisms detected in sulfidic systems. The *Thermocrinis* organisms present in OS_11 and BCH_13 contain Type A-HCOs (**Figure 9**), indicating their functional divergence from the other Aquificales genera with respect to oxygen. There is evidence that some Aquificales have copies of both types of HCOs as is noted for *Hydrogenivirga* spp. Consequently, it is possible that subsequent evolution in specific habitat types separated lineages containing only the Type C or the Type A HCO. Interestingly, the *Hydrogenobaculum* and *Sulfurihydrogenibium* populations both contain the cbb3-Type C HCOs even though they are members of different families. The *Hydrogenobaculum* and *Thermocrinis* organisms are from the same family, but do not share the same respiratory complexes, although they do share similarity in r-TCA proteins important in CO_2 fixation. The major difference in metabolic processing of CO_2 and O_2 between these lineages provides an excellent opportunity for relating their evolutionary histories to paleobiological events and the timing of radiation relative to the "Great Oxidation Event" (Canfield, 2005; Anbar et al., 2007; Konhauser, 2009).

The two higher-pH (pH ∼8) non-sulfidic sites (OS_11, BCH_13) contained a similar *Thermocrinis*-like Aquificales, however, the microbial community structure was considerably different and the OS_11 streamer community contained at least three additional novel bacterial assemblies compared to BCH_13. The inorganic constituents of these two springs were reasonably similar and they both supported a similar *Thermocrinis*-like population. Clearly, additional geochemical and or geophysical factors, not considered in the current study, contribute to these differences in community structure across apparently similar sites. Variations in dissolved and/or particulate organic carbon across sites may play an important role in modifying community composition. However, the organic compounds contributing to the measured total dissolved organic carbon (DOC), as well as solid-phases of C, have not been characterized. Although the concentration of DOC was higher in OS_11 than BCH_13, this association is not supported with any detailed linkages among specific organic compounds and microbial diversity at the current time. Further characterization of organic constituents present in geothermal systems will be necessary to determine if variations in organic solutes contained in

geothermal source waters also influence community structure and function across different habitat types. Moreover, additional sites are necessary to understand possible differences in *Thermocrinis*-like populations present in high-sulfide environments (e.g., CS_12) compared to those in low sulfide systems (e.g., OCT_11, BCH_13).

MATERIALS AND METHODS

SITE SELECTION, SAMPLE COLLECTION, AND PROCESSING
Six Aquificales "streamer" communities (**Figure 1**) were sampled from high-velocity, in-channel habitats during 2007–2008, and were chosen to replicate at least three major lineages of Aquificales known to exist in YNP across a pH range from 3 to 9 (i.e., *Hydrogenobaculum*, *Sulfurihydrogenibium*, and *Thermocrinis* spp.). The research sites chosen for study have been the subject of significant prior characterization and include: *Dragon Spring* (DS_9), *One Hundred Springs Plain* (OSP_14), *Mammoth Hot Spring* (MHS_10), *Calcite Springs* (CS_12), *Octopus Spring* (OS_11), and *Bechler Springs* (BCH_13) (e.g., Fouke et al., 2003; Inskeep and McDermott, 2005; Inskeep et al., 2005; Reysenbach et al., 2005; Fouke, 2011). Each microbial community and associated solid phase was sampled aseptically, stored in 50 mL sterile Falcon tubes on dry ice, and transported to a −80°C freezer (MSU) until DNA extraction.

Parallel samples of the bulk aqueous phase ($<0.2\,\mu m$) associated with the microbial community were obtained simultaneously and analyzed using a combination of field and laboratory methods. As described in more detail in other reports (Inskeep et al., 2004; Macur et al., 2004; Hall et al., 2008), pH, temperature, and other redox sensitive species (Fe^{2+}/Fe^{3+}; As^{III}/As^{V}; total dissolved sulfide; dissolved O_2) were determined using field methods. Total dissolved ions were determined using inductively coupled plasma (ICP) spectrometry and ion chromatography (for all major cations, anions, and trace elements). Dissolved gases (CO_2, H_2, CH_4) (all sites but BCH) were determined using closed head-space gas chromatography (Inskeep et al., 2005) of sealed serum-bottle samples obtained in the field. A subset of these sites have been sampled many times with excellent replication (Langner et al., 2001; Inskeep et al., 2005, 2010; Reysenbach et al., 2005; Fouke, 2011). The location and primary physicochemical characteristics obtained during sampling are provided here (**Table 1**), and additional geochemical data are provided as supplemental information (Table S2 in Supplementary Material, Inskeep et al., 2013).

DNA EXTRACTION AND LIBRARY CONSTRUCTION
DNA was extracted from all samples using a standardized protocol (Inskeep et al., 2013) to minimize variation in composition across sample type due to extraction method or technician. Our main emphasis was to obtain representative, unbiased community DNA as template for construction of small insert libraries. Briefly, approximately 3 g wet samples were extracted in 1 ml of Buffer A (200 mM Tris, pH 8; 50 mM EDTA; 200 mM NaCl; 2 mM sodium citrate; 10 mM $CaCl_2$) with lysozyme (1 mg/ml final concentration) for 1.5 h at 37°C. Proteinase K (final concentration 1mg/ml) and SDS [final concentration 0.3% (w/v)] were added and incubated for 0.5 h at 37°C. This first lysate was removed, and the samples were re-extracted using bead-beating

protocols. The two lysates were then combined and extracted with phenol-chloroform, and the resulting DNA was re-precipitated in ethanol, treated with RNAase and quantified by gel electrophoresis and staining. Small insert (puc13) libraries were constructed and transformed, then sequenced using Sanger sequencing to generate approximately 40 Mbp per site (∼800 bp reads), with the exception of MHS_10 that only received ∼20 Mbp of Sanger sequence, due in part to the simplicity of the community and the fact that MHS_10 also received a half-plate of 454 pyrosequencing. For consistency, this manuscript focuses on the Sanger data across each of the six sites; assemblies using the pyro-sequence data for MHS_10 did not result in improved contig size or increases in total assembled data, although it did improve the coverage of this phylotype to $>50\times$. Full-length 16S rRNA genes were also PCR-amplified and cloned from the DNA of each site using universal primers specific for *Bacteria* and *Archaea*, and one 384-well plate was sequenced from each successful library. Unique PCR-amplified 16S rRNA gene sequences (<97% DNA similarity) were chimera-checked using Bellerophon (Huber et al., 2004).

METAGENOME SEQUENCE ANALYSIS
Unassembled metagenomic sequence reads were plotted as a function of %G + C content and taxonomic assignment based on best "blastx" hits using MEGAN (Huson et al., 2007). Only a handful of microbial genomes currently serve as appropriate references for the indigenous organisms within these chemotrophic communities, consequently, many of the taxonomic assignments were given at family or domain level. Genome-level analysis of metagenome data was performed using fragment recruitment of unassembled sequence reads to reference microbial genomes (Rusch et al., 2007). At the time of writing, this database contained reference microbial genomes for ∼1500 bacteria and 100 archaea.

Random shotgun DNA sequence (∼40–50 Mb Sanger per site) was assembled using both the Celera (Version 4.0, Rusch et al., 2007) and PGA (Zhao et al., 2008) assemblers as described previously in Inskeep et al. (2013). Briefly, the analyses presented here was based on the Celera assemblies of our metagenomic data built using the following parameters: doOverlapTrimming = 0, doFragmentCorrection = 0, globalErrorRate = 12, utgErrorRate = 150, utgBubblePopping = 1, and useBogUnitig = 0. For PGA assemblies (Zhao et al., 2008), the following parameters were employed: OverlapLen = 30; Percent = 0.75; Clearance = 30; ClipIdn = 77; ClipQual = 10; CutoffScore = 400; EndOverhang = 800; InOverhang = 500; MinCovRep = 50; MinLinks = 2; MinSat = 3; NumIter = 50; PenalizeN = 1; QualOverLim = 400; QualScoreCutoff = 200; QualSumLim = 3500; SimDiFac = 30; Verbosity = 1. Assemblies obtained from Celera and PGA were gene-called and annotated using the Department of Energy-Joint Genome Institute IMG/M pipeline (Markowitz et al., 2012). All annotated metagenome sequence assemblies (Celera/PGA) discussed in the current manuscript are available through the DOE-JGI IMG/M (Markowitz et al., 2012) website (http://img.jgi.doe.gov/m) under IMG taxon OID numbers as follows: YNPSite09 (2022920010/2014031004), Site14 (2022920007/ 2013954001), Site10 (2022920015/2015391001),

Site12 (2022920011/2014031005), Site11 (2022920012/2014031007), and Site13 (2022920006/2013515002).

Assembled metagenome sequence (e.g., contigs and scaffolds > 2 kb) was analyzed using three dimensional PCA scatterplots of nucleotide word frequencies (Teeling et al., 2004; Inskeep et al., 2010) to evaluate consensus assembled sequence of dominant phylotypes (e.g., **Figure 3A**). The sequence clusters were also viewed (**Figure 3B**) with a simultaneous blast-based taxonomic classification (Rusch et al., 2007) or the Automated Phylogenetic Inference System (APIS; Badger et al., 2006). Briefly, APIS is a system for automatic creation and summarizing of phylogenetic trees for each protein encoded by a genome or metagenomic dataset.

PHYLOGENETIC MARKER AND SINGLE-COPY GENES

The community composition of each site was also investigated by phylogenetic analysis of 16S rRNA genes detected among the six assembled metagenomes in IMG/M. Of the 407 sequences annotated as 16S rRNA genes, 64 were of sufficient length (>1100 bp) for robust phylogenetic analysis. Smaller fragments were not included in the alignments or phylogenetic tree to maximize robust phylogenetic placement of the primary community members. 16S rRNA gene sequences were aligned in Green genes (DeSantis et al., 2006), imported into the ARB program (Ludwig et al., 2004), and manually adjusted according to conserved regions of the gene and the established secondary structure to ensure that only homologous regions were compared. Initial phylogenetic analysis was performed in PAUP* (version 40.b10; Sinauer Associates, Sunderland, MA, USA) and tree topology was explored using parsimony, neighbor-joining, and maximum-likelihood analyses. Final 16S rRNA gene trees (**Figure 4**; **Figure A3** in Appendix) were created by neighbor-joining analysis with a maximum-likelihood correction. For the Aquificales-specific tree (**Figure 4**), a heuristic search was performed with tree bisection-reconnection (TBR) branch swapping in PAUP*. Transition/transversion ratio and nucleotide frequencies were estimated according to the F84 model and bootstrap values were determined from 1000 re-samplings of the dataset. 16S rRNA genes from metagenome assemblies and unique (<97% DNA similarity) PCR-amplified 16S rRNA genes from the original sample DNA (up to 384 sequences per library were amplified using universal bacterial and archaeal primers and chimera-checked using Bellerophon; Huber et al., 2004) were added to a neighbor-joining 16S rRNA gene tree (**Figure A3** in Appendix) using the parsimony tool in ARB.

Community composition was also investigated by identifying the phylogeny of 31 conserved protein-encoding marker genes using AMPHORA (Wu and Eisen, 2008). Assembled contigs and singleton reads from the six sites were analyzed as described for the Sargasso Sea dataset in Wu and Eisen (2008). Phylotypes were identified from the first internal node (n1) whose bootstrap support exceeded 70%. Phylogenetic classifications that had less than 70% bootstrap support were resolved further by "blastx identity." For example, because there are few archaeal representatives included in the AMPHORA database, significant bootstrap support was not often detected for archaeal query sequences. Similarly, deeply divergent sequences were poorly supported and were unidentified in AMPHORA, thus requiring further investigation using "blastx." Aquificales single-copy genes that were shared among

all datasets ($n = 911$ total distributed among 29 protein markers) were identified and population richness was determined as the average number of single-copy genes present in each dataset (Table S1 in Supplementary Material).

TIGRFAM ANALYSIS

Assembled metagenome sequence from each of the "streamer" communities was annotated as described in Inskeep et al. (2010, 2013) and predicted proteins from the scaffolds were assigned TIGRFAM protein families (Selengut et al., 2007) using HMMER 3 (Eddy, 2011) with E-value cutoff of 1e-6. PCA and statistical analysis of site group differences was performed using the STAMP v2.0 software (Parks and Beiko, 2010). Briefly, the White's non-parametric T-test and ANOVA tests were used to test for differences between two site groups and multiple site groups, respectively. Two-way clustering was done using row-standardized (across sites) average TIGRFAM category abundance data using the Euclidean distance metric and complete-linkage hierarchical clustering in MeV 4.8 (Saeed et al., 2003) software.

FUNCTIONAL COMPARISONS AMONG AQUIFICALES LINEAGES IN YNP

The assembled metagenome sequence data was screened for specific functional genes corresponding to known and or putative pathways involved in biosynthesis and energy transfer. Specifically, we were interested in assessing metabolic potential for chemolithoautotrophy (CO_2 fixation and electron transfer genes) in high-temperature geothermal systems. Query DNA sequences known to code for proteins important in the oxidation of reduced chemical constituents or the reduction of a terminal acceptor were used to search the assembled metagenome sequence data using "blastx" routines (full list of gene sequences and accession numbers given Table S3 in Supplementary Material, in Appendix, Inskeep et al., 2013). IMG/M was used as an additional method for identifying CO_2 fixation and other functional genes, and for gene neighborhood analysis. Metagenome sequences exhibiting homology (E-values $< 10^{-10}$) to query sequences were then carefully assessed by manually examining amino acid sequence alignments (for fragments of sufficient length relative to query sequence) and subsequent phylogenetic analysis of deduced protein sequences against known relatives. False positives were eliminated by this screening process and included (i) sequences matching the correct protein family of the query sequence, but not the exact query sequence (e.g., Mo-pterin oxidoreductases versus a specific protein within this family), (ii) sequences that match a query sequence due to homologous regions, but are clearly associated with a gene or gene cluster with different function, and (iii) sequences that returned mis-annotated "blastn" relatives. It is also possible that our inventory of metabolic potential has missed sequences related to a specific query gene. For example, some homologous genes found in the metagenome data were of insufficient length relative to a specific query sequence to make a definitive assignment. Clearly, the metagenomes obtained here do not represent complete sequence for all sub-dominant populations in these sites, thus the functional analysis also cannot be considered complete for these representatives. Phylogenetic analysis was performed on amino acid sequences (aligned in MUSCLE; Edgar, 2004) of select functional genes in MEGA 5 (Tamura et al.,

2011) using maximum-likelihood analysis with bootstrapping (1000 replicates).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Microbial_Physiology_and_Metabolism/10.3389/fmicb.2013.00084/abstract

Table S1 | Number of Aquificales-like single-copy genes identified in assembled metagenome sequence using AMPHORA (Wu and Eisen, 2008).

Table S2 | TIGRFAM electron transport gene family counts across six Aquificales streamer communities and results for comparison of low-pH and high-pH sites using White's non-parametric *T*-test.

Table S3 | TIGRFAM functional category gene family counts across six Aquificales streamer communities and results for comparison of taxonomically distinct sites (Hydrogenobaculum-dominated, Sulfurihydrogenibium-dominated, Thermocrinis (Aquificaceae)-dominated) using ANOVA.

REFERENCES

Amend, J. P., and Shock, E. L. (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiol. Rev.* 25, 175–243.

Anbar, A., Duan, Y., Lyons, T., Arnold, G., Kendall, B., Creaser, R., et al. (2007). A whiff of oxygen before the great oxidation event? *Science* 317, 1903.

Aoshima, M., Ishii, M., and Igarashi, Y. (2004). A novel enzyme, citryl-CoA lyase, catalysing the second step of the citrate cleavage reaction in *Hydrogenobacter thermophilus* TK-6. *Mol. Microbiol.* 52, 763–770.

Badger, J. H., Hoover, T. R., Brun, Y. V., Weiner, R. M., Laub, M. T., Alexandre, G., et al. (2006). Comparative genomic evidence for a close relationship between the dimorphic prosthecate bacteria *Hyphomonas neptunium* and *Caulobacter crescentus*. *J. Bacteriol.* 188, 6841–6850.

Barion, S., Franchi, M., Gallori, E., and Di Giulio, M. (2007). The first lines of divergence in the Bacteria domain were the hyperthermophilic organisms, the Thermotogales and the Aquificales, and not the mesophilic Planctomycetales. *BioSystems* 87, 13–19.

Beh, M., Strauss, G., Huber, R., Stetter, K.-O., and Fuchs, G. (1993). Enzymes of the reductive citric acid cycle in the autotrophic eubacterium *Aquifex pyrophilus* and in the archaebacterium *Thermoproteus neutrophilus*. *Arch. Microbiol.* 160, 306–311.

Bentley, R., and Chasteen, T. G. (2002). Microbial methylation of metalloids: arsenic, antimony, and

bismuth. *Microbiol. Mol. Biol. Rev.* 66, 250–271.

Blank, C. E., Cady, S. L., and Pace, N. R. (2002). Microbial composition of near-boiling silica-depositing thermal springs throughout Yellowstone National Park. *Appl. Environ. Microbiol.* 68, 5123–5135.

Boussau, B., Gueguen, L., and Gouy, M. (2008). Accounting for horizontal gene transfers explains conflicting hypotheses regarding the position of Aquificales in the phylogeny of Bacteria. *BMC Evol. Biol.* 8:272. doi:10.1186/1471-2148-8-272

Boyd, E. S., Leavitt, W. D., and Geesey, G. G. (2009). CO_2 uptake and fixation by a thermoacidophilic microbial community attached to precipitated sulfur in a geothermal spring. *Appl. Environ. Microbiol.* 75, 4289–4296.

Caldwell, S. L., Liu, Y., Ferrera, I., Beveridge, T., and Reysenbach, A. L. (2010). *Thermocrinis minervae* sp. nov., a hydrogen- and sulfur-oxidizing, thermophilic member of the Aquificales from a Costa Rican terrestrial hot spring. *Int. J. Syst. Evol. Microbiol.* 60, 338–343.

Canfield, D. E. (2005). The early history of atmospheric oxygen: homage to Robert M. *Garrels. Annu. Rev. Earth Planet. Sci.* 33, 1–36.

Cherney, M. M., Zhang, Y., Solomonsona, M., Weinera, J. H., and James, N. G. (2010). Crystal structure of sulfide:quinone oxidoreductase from *Acidithiobacillus ferrooxidans*: insights into sulfidotrophic respiration and detoxification. *J. Molec. Biol.* 398, 292–305.

Clingenpeel, S. R., D'Imperio, S., Oduro, H., Druschel, G. K., and McDermott, T. R. (2009). Cloning and *in situ* expression studies of the *Hydrogenobaculum* arsenite oxidase

genes. *Appl. Environ. Microbiol.* 75, 3362–3365.

Coenye, T., and Vandamme, P. (2004). A genomic perspective on the relationship between the Aquificales and the epsilon-Proteobacteria. *Syst. Appl. Microbiol.* 27, 313–322.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Green genes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072.

D'Imperio, S., Lehr, C. R., Breary, M., and McDermott, T. R. (2007). Autecology of an arsenite chemolithotroph: sulfide constraints on function and distribution in a geothermal spring. *Appl. Environ. Microbiol.* 73, 7067–7074.

D'Imperio, S., Lehr, C. R., Oduro, H., Druschel, G., Kuhl, M., and McDermott, T. R. (2008). Relative importance of H_2 and H_2S as energy sources for primary production in geothermal springs. *Appl. Environ. Microbiol.* 74, 5802–5808.

Donahoe-Christiansen, J., D'Imperio, S., Jackson, C. R., Inskeep, W. P., and McDermott, T. R. (2004). Arsenite-oxidizing *Hydrogenobaculum* strain isolated from an acid-sulfate-chloride geothermal spring in Yellowstone National Park. *Appl. Environ. Microbiol.* 70, 1865–1868.

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi:10.1371/journal.pcbi.1002195

Eder, W., and Huber, R. (2002). New isolates and physiological properties of the Aquificales and description of *Thermocrinis*

albus sp. nov. *Extremophiles* 6, 309–318.

Edgar, R. C. (2004). Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Ferrera, I., Longhorn, S., Banta, A. B., Liu, Y., Preston, D., and Reysenbach, A. L. (2007). Diversity of 16S rRNA gene, ITS region and aclB gene of the Aquificales. *Extremophiles* 11, 57–64.

Fouke, B. W. (2011). Hot-spring systems geobiology: abiotic and biotic influences on travertine formation at Mammoth Hot Springs, Yellowstone National Park, USA. *Sedimentology* 58, 170–219.

Fouke, B. W., Bonheyo, G. T., Sanzenbacher, B., and Frias-Lopez, J. (2003). Partitioning of bacterial communities between travertine depositional facies at Mammoth Hot Springs, Yellowstone National Park, USA. *Can. J. Earth Sci.* 40, 1531–1548.

Fouke, B. W., Farmer, J. D., Des Marais, D. J., Pratt, L., Sturchio, N. C., Burns, P. C., et al. (2000). Depositional facies and aqueous-solid geochemistry of travertine-depositing hot springs (Angel Terrace, Mammoth Hot Springs, Yellowstone National Park, U.S.A.). *J. Sediment. Res. A Sediment. Petrol. Process.* 70, 565–585.

Fournier, R. O. (1989). Geochemistry and dynamics of the Yellowstone National Park hydrothermal system. *Annu. Rev. Earth Planet. Sci.* 17, 13–53.

Friedrich, C. G., Bardischewsky, F., Rother, D., Quentmeier, A., and Fischer, J. (2005). Prokaryotic sulfur oxidation. *Curr. Opin. Microbiol.* 8, 253–259.

Garcia-Horsman, J. A., Barquera, B., Rumbley, J., Ma, J., and Gennis, R. B. (1994). The superfamily of heme-copper respiratory oxidases. *J. Bacteriol.* 176, 5587.

Ghosh, W., and Dam, B. (2009). Biochemistry and molecular biology of lithotrophic sulfur oxidation by taxonomically and ecologically diverse bacteria and archaea. *FEMS Microbiol. Rev.* 33, 999–1043.

Griffiths, E., and Gupta, R. S. (2004). Signature sequences in diverse proteins provide evidence for the late divergence of the order Aquificales. *Int. Microbiol.* 7, 41–52.

Hall, J. R., Mitchell, K. R., Jackson-Weaver, O., Kooser, A. S., Cron, B. R., Crossey, L. J., et al. (2008). Molecular characterization of the diversity and distribution of a thermal spring microbial community by using rRNA and metabolic genes. *Appl. Environ. Microbiol.* 74, 4910–4922.

Hamamura, N., Macur, R. E., Korf, S., Ackerman, G., Taylor, W. P., Kozubal, M., et al. (2009). Linking microbial oxidation of arsenic with detection and phylogenetic analysis of arsenite oxidase genes in diverse geothermal environments. *Environ. Microbiol.* 11, 421–431.

Harmsen, H., Prieur, D., and Jeanthon, C. (1997). Group-specific 16S rRNA-targeted oligonucleotide probes to identify thermophilic bacteria in marine hydrothermal vents. *Appl. Environ. Microbiol.* 63, 4061–4068.

Huber, R., and Eder, W. (2006). "Aquificales," in *The Prokaryotes: A Handbook on the Biology of Bacteria*, 3rd Edn, Vol. 7, eds M. Dworkin, S. Falkow, E. Rosenberg, K. H. Schleifer, and E. Stackebrandt (New York: Springer), 925–928.

Huber, R., Eder, W., Heldwein, S., Wanner, G., Huber, H., Rachel, R., et al. (1998). *Thermocrinis ruber* gen. nov., sp. nov., a pink-filament-forming hyperthermophilic bacterium isolated from Yellowstone National Park. *Appl. Environ. Microbiol.* 64, 3576–3583.

Huber, R., Wilharm, T., Huber, D., Trincone, A., Burggraf, S., König, H., et al. (1992). *Aquifex pyrophilus* gen. nov. spec. nov., represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria. *Syst. Appl. Microbiol.* 15, 340–351.

Huber, T., Faulkner, G., and Hugenholtz, P. (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* 20, 2317–2319.

Hugenholtz, P., Pitulle, C., Hershberger, K. L., and Pace, N. R. (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *J. Bacteriol.* 180, 366–376.

Hugler, M., Huber, H., Molyneaux, S. J., Vetriani, C., and Sievert, S. M. (2007). Autotrophic CO$_2$ fixation via the reductive tricarboxylic acid cycle in different lineages within the phylum Aquificae: evidence for two ways of citrate cleavage. *Environ. Microbiol.* 9, 81–92.

Hugler, M., Wirsen, C. O., Fuchs, G., Taylor, C. D., and Sievert, S. M. (2005). Evidence for autotrophic CO$_2$ fixation via the reductive tricarboxylic acid cycle by members of the epsilon subdivision of Proteobacteria. *J. Bacteriol.* 187, 3020–3027.

Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. (2007). Megan analysis of metagenomic data. *Genome Res.* 17, 377–386.

Inagaki, F., Takai, K., Hirayama, H., Yamato, Y., Nealson, K. H., and Horikoshi, K. (2003). Distribution and phylogenetic diversity of the subsurface microbial community in a Japanese epithermal gold mine. *Extremophiles* 7, 307–317.

Inskeep, W. P., Ackerman, G. G., Taylor, W. P., Kozubal, M., Korf, S., and Macur, R. E. (2005). On the energetics of chemolithotrophy in nonequilibrium systems: case studies of geothermal springs in Yellowstone National Park. *Geobiology* 3, 297–317.

Inskeep, W. P., Jay, Z. J., Tringe, S. G., Herrgard, M., Rusch, D. B., et al. (2013). The YNP Metagenome Project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem. *Front. Microbiol.* 4:67. doi:10.3389/fmicb.2013.00067

Inskeep, W. P., Macur, R. E., Harrison, G., Bostick, B. C., and Fendorf, S. (2004). Biomineralization of As (V)-hydrous ferric oxyhydroxide in microbial mats of an acid-sulfate-chloride geothermal spring, Yellowstone National Park. *Geochim. Cosmochim. Acta* 68, 3141–3155.

Inskeep, W. P., Rusch, D. B., Jay, Z. J., Herrgard, M. J., Kozubal, M. A., Richardson, T. H., et al. (2010). Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. *PLoS ONE* 5:e9773. doi:10.1371/journal.pone.0009773

Inskeep, W. P., and McDermott, T. R. (2005). "Geomicrobiology of acid-sulfate-chloride geothermal springs

in Yellowstone National Park," in *Geothermal biology and geochemistry in Yellowstone National Park*, eds W. P. Inskeep & T. R. McDermott (Thermal Biology Institute, Montana State University, Bozeman, MT 59717), 143-162.

Jackson, C. R., Langner, H. W., Donahoe-Christiansen, J., Inskeep, W. P., and McDermott, T. R. (2001). Molecular analysis of microbial community structure in an arsenite-oxidizing acidic thermal spring. *Environ. Microbiol.* 3, 532–542.

Jahnke, L. L., Eder, W., Huber, R., Hope, J. M., Hinrichs, K. U., Hayes, J. M., et al. (2001). Signature lipids and stable carbon isotope analyses of Octopus Spring hyperthermophilic communities compared with those of Aquificales representatives. *Appl. Environ. Microbiol.* 67, 5179–5189.

Jünemann, S. (1997). Cytochrome bd terminal oxidase. *Biochim. Biophys. Acta* 1321, 107–127.

Kandianis, M. T., Fouke, B. W., Johnson, R. W., Veysey, J., and Inskeep, W. P. (2008). Microbial biomass: a catalyst for CaCO$_3$ precipitation in advection-dominated transport regimes. *Geol. Soc. Am. Bull.* 120, 442–450.

Konhauser, K. (2009). Biogeochemistry: deepening the early oxygen debate. *Nat. Geosci.* 2, 241–242.

Kozubal, M., Dlakic, M., Macur, R., and Inskeep, W. (2011). Terminal oxidase diversity and function in "*Metallosphaera yellowstonensis*": gene expression and protein modeling suggest mechanisms of Fe (II) oxidation in the Sulfolobales. *Appl. Environ. Microbiol.* 77, 1844–1853.

Langner, H. W., Jackson, C. R., McDermott, T. R., and Inskeep, W. P. (2001). Rapid oxidation of arsenite in a hot spring ecosystem, Yellowstone National Park. *Environ. Sci. Technol.* 35, 3302–3309.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004). ARB: a software environment for sequence data. *Nucleic Acids Res.* 32, 1363–1371.

Macur, R. E., Langner, H. W., Kocar, B. D., and Inskeep, W. P. (2004). Linking geochemical processes with microbial community analysis: successional dynamics in an arsenic-rich, acid-sulphate-chloride geothermal spring. *Geobiology* 2, 163–177.

Markowitz, V. M., Chen, I.-M. A., Chu, K., Szeto, E., Palaniappan, K., Grechkin, Y., et al. (2012). IMG/M: the integrated metagenome data management and comparative

analysis system. *Nucleic Acids Res.* 40, D123–D129.

Medini, D., Donati, C., Tettelin, H., Masignani, V., and Rappuoli, R. (2005). The microbial pan-genome. *Curr. Opin. Genet. Dev.* 15, 589–594.

Mukhopadhyay, R., Rosen, B. P., Phung, L. T., and Silver, S. (2002). Microbial arsenic: from geocycles to genes and enzymes. *FEMS Microbiol. Rev.* 26, 311–325.

Nakagawa, S., Shtaih, Z., Banta, A., Beveridge, T. J., Sako, Y., and 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*. *Int. J. Syst. Evol. Microbiol.* 55, 2263–2268.

Nordstrom, D. K., Ball, J. W., and McCleskey, R. B. (2005). "Ground water to surface water: chemistry of thermal outflows in Yellowstone National Park," in *Geothermal Biology and Geochemistry in Yellowstone National Park*, eds W. P. Inskeep and T. R. McDermott (Bozeman: Montana State University Thermal Biology Institute), 143–162.

Parks, D. H., and Beiko, R. G. (2010). Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 26, 715–721.

Pereira, M. M., Santana, M., and Teixeira, M. (2001). A novel scenario for the evolution of haem-copper oxygen reductases. *Biochim. Biophys. Acta* 1505, 185–208.

Planer-Friedrich, B., Fisher, J. C., Hollibaugh, J. T., Süß, E., and Wallschläger, D. (2009). Oxidative transformation of trithioarsenate along alkaline geothermal drainages-abiotic versus microbially mediated processes. *Geomicrobiol. J.* 26, 339–350.

Reysenbach, A.-L., Banta, A., Civello, S., Daly, J., Mitchell, K., Lalonde, S., et al. (2005). "The aquificales of Yellowstone National Park," in *Geothermal Biology and Geochemistry in Yellowstone National Park*, eds W. P. Inskeep and T. R. McDermott (Bozeman: Montana State University Thermal Biology Institute), 129–142.

Reysenbach, A. L., Banta, A. B., Boone, D. R., Cary, S. C., and Luther, G. W. (2000a). Microbial essentials at hydrothermal vents. *Nature* 404, 835.

Reysenbach, A. L., Ehringer, M., and Hershberger, K. (2000b). Microbial diversity at 83 degrees C in Calcite Springs, Yellowstone National Park: another environment where the Aquificales and "Korarchaeota" coexist. *Extremophiles* 4, 61–67.

Reysenbach, A. L., and Cady, S. L. (2001). Microbiology of ancient and modern hydrothermal systems. *Trends Microbiol.* 9, 79–86.

Reysenbach, A. L., Hamamura, N., Podar, M., Griffiths, E., Ferreira, S., Hochstein, R., et al. (2009). Complete and draft genome sequences of six members of the Aquificales. *J. Bacteriol.* 191, 1992–1993.

Reysenbach, A. L., Seitzinger, S., Kirshtein, J., and Mclaughlin, E. (1999). Molecular constraints on a high-temperature evolution of early life. *Biol. Bull.* 196, 367–371.

Reysenbach, A. L., Wickham, G. S., and Pace, N. R. (1994). Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. *Appl. Environ. Microbiol.* 60, 2113–2119.

Romano, C., D'Imperio, S., Woyke, T., Mavromatis, K., Lasken, R., Shock, E. L., et al. (2013). Comparative genomic analysis of phylogenetically closely-related *Hydrogenobaculum* sp. from Yellowstone National Park. *Appl. Environ. Microbiol.* 79, 2932–2943.

Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., et al. (2007). The Sorcerer II Global Ocean sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* 5:e77. doi:10.1371/journal.pbio.0050077

Saeed, A., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., et al. (2003). TM4: a free, open-source system for microarray data management and analysis. *BioTechniques* 34, 374.

Santini, J. M., Kappler, U., Ward, S. A., Honeychurch, M. J., Vanden Hoven, R. N., and Bernhardt, P. V. (2007). The NT-26 cytochrome c552 and its

role in arsenite oxidation. *Biochim. Biophys. Acta* 1767, 189–196.

Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W. C., et al. (2007). TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Res.* 35, D260–D264.

Shiba, H., Kawasumi, T., Igarashi, Y., Kodama, T., and Minoda, Y. (1985). The CO$_2$ assimilation via the reductive tricarboxylic acid cycle in an obligately autotrophic aerobic hydrogen-oxidizing bacterium, *Hydrogenobacter thermophilus. Arch. Microbiol.* 141, 198–203.

Sousa, F. L., Alves, R. J., Pereira-Leal, J. B., Teixeira, M., and Pereira, M. M. (2011). A bioinformatics classifier and database for heme-copper oxygen reductases. *PLoS ONE* 6:e19117. doi:10.1371/journal.pone.0019117

Spear, J. R., Walker, J. J., Mccollom, T. M., and Pace, N. R. (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2555–2560.

Stohr, R., Waberski, A., Volker, H., Tindall, B. J., and Thomm, M. (2001). *Hydrogenothermus marinus* gen. nov., sp. nov., a novel thermophilic hydrogen-oxidizing bacterium, recognition of *Calderobacterium hydrogenophilum* as a member of the genus *Hydrogenobacter* and proposal of the reclassification of *Hydrogenobacter acidophilus* as *Hydrogenobaculum acidophilum* gen. nov., comb. nov., in the phylum 'Hydrogenobacter/Aquifex'. *Int. J. Syst. Evol. Microbiol.* 51, 1853–1862.

Takai, K., Hirayama, H., Sakihama, Y., Inagaki, F., Yamato, Y., and Horikoshi, K. (2002). Isolation and metabolic characteristics of previously uncultured members of the order Aquificales in a subsurface gold mine. *Appl. Environ. Microbiol.* 68, 3046–3054.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). Mega5: molecular evolutionary genetics analysis

using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.

Teeling, H., Waldmann, J., Lombardot, T., Bauer, M., and Glockner, F. O. (2004). Tetra: a web-service and a stand-alone program for the analysis and comparison of tetranucleotide usage patterns in DNA sequences. *BMC Bioinformatics* 5:163. doi:10.1186/1471-2105-5-163

Van Dover, C. L., Humphris, S. E., Fornari, D., Cavanaugh, C. M., Collier, R., Goffredi, S. K., et al. (2001). Biogeography and ecological setting of Indian Ocean hydrothermal vents. *Science* 294, 818–823.

Veysey, J., Fouke, B. W., Kandianis, M. T., Schickel, T. J., Johnson, R. W., and Goldenfeld, N. (2008). Reconstruction of water temperature, pH, and flux of ancient hot springs from travertine depositional facies. *J. Sediment. Res. A Sediment. Petrol. Process.* 78, 69–76.

Wachtershauser, G. (1990). Evolution of the first metabolic cycles. *Proc. Natl. Acad. Sci. U.S.A.* 87, 200.

Ward, D. M., Bateson, M. M., Ferris, M. J., Kühl, M., Wieland, A., Koeppel, A., et al. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 1997–2008.

Wu, M., and Eisen, J. A. (2008). A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* 9, R151.

Xu, Y., Schoonen, M., Nordstrom, D., Cunningham, K., and Ball, J. (1998). Sulfur geochemistry of hydrothermal waters in Yellowstone National Park: I. The origin of thiosulfate in hot spring waters. *Geochim. Cosmochim. Acta* 62, 3729–3743.

Yamamoto, H., Hiraishi, A., Kato, K., Chiura, H. X., Maki, Y., and Shimizu, A. (1998). Phylogenetic evidence for the existence of novel thermophilic bacteria in hot

spring sulfur-turf microbial mats in Japan. *Appl. Environ. Microbiol.* 64, 1680–1687.

Zhao, F., Zhao, F., Li, T., and Bryant, D. A. (2008). A new pheromone trail-based genetic algorithm for comparative genome assembly. *Nucleic Acids Res.* 36, 3455–3462.

Zhaxybayeva, O., Swithers, K. S., Lapierre, P., Fournier, G. P., Bickhart, D. M., Deboy, R. T., et al. (2009). On the chimeric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5865–5870.

APPENDIX

FIGURE A1 | Recruitment of metagenome sequence reads (∼800 bp per read) from six Aquificales streamer samples to the most relevant reference genomes currently available in reference databases. Recruitment of metagenome sequence fragments across the complete genomes (x-axis) are shown from 50 to 100% nucleotide identity (y-axis). Site Colors: Yellow = *Dragon Spring* (DS_9), Red = *One Hundred Spring Plain* (OSP_14); Green = Mammoth Hot Springs (MHS_10), Violet = *Calcite Springs* (CS_12); Dark-Blue = *Octopus Spring* (OS_11), light-blue = *Bechler Springs* (BCH_13). Reference genomes: *Hydrogenobaculum* sp. Y04AAS1; *Sulfurihydrogenibium* sp. Y03AOP1 (with and without the 220 Mb of pyrosequence for MHS_10); *Thermocrinis albus* DSM 14484; *Hydrogenobacter thermophilus* TK-6; *Thermus aquaticus* HB8 (Plot constructed using JCVI bioinformatic utilities, Rusch et al., 2007).
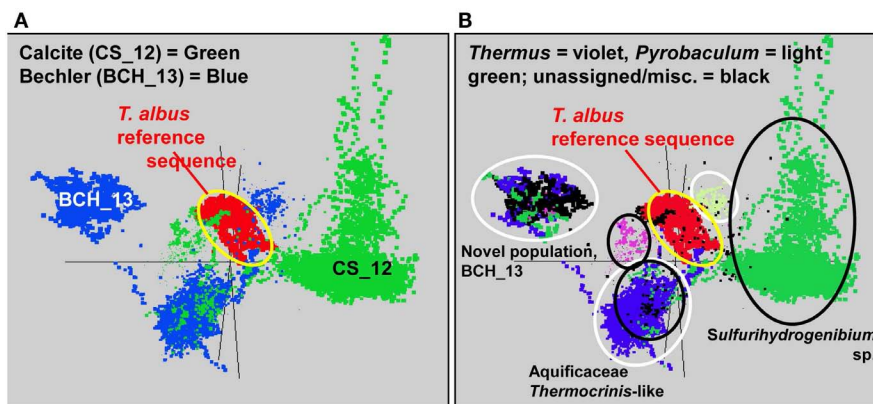
FIGURE A2 | Nucleotide word frequency PCA plots showing detailed analysis of Thermocrinis-like populations in Calcite and Bechler Springs compared to Thermocrinis albus reference sequence and other phyla present in these two communities. **(A)** Sequence data colored by site (CS_12 and BCH_13), with *T. albus* reference sequence (red) added for comparison., and **(B)** Sequence data in identical orientation now analyzed phylogenetically to reveal specific assembled sequence corresponding to particular population types within the two different sites (green = *Sulfurihydrogenibium* sp., blue = Aquificaceae, violet = *Thermus* sp.; light-green = Thermoproteaceae; unassigned = black). White circles indicate major assemblies from BCH_13, and black circles indicate major assemblies in CS_12.
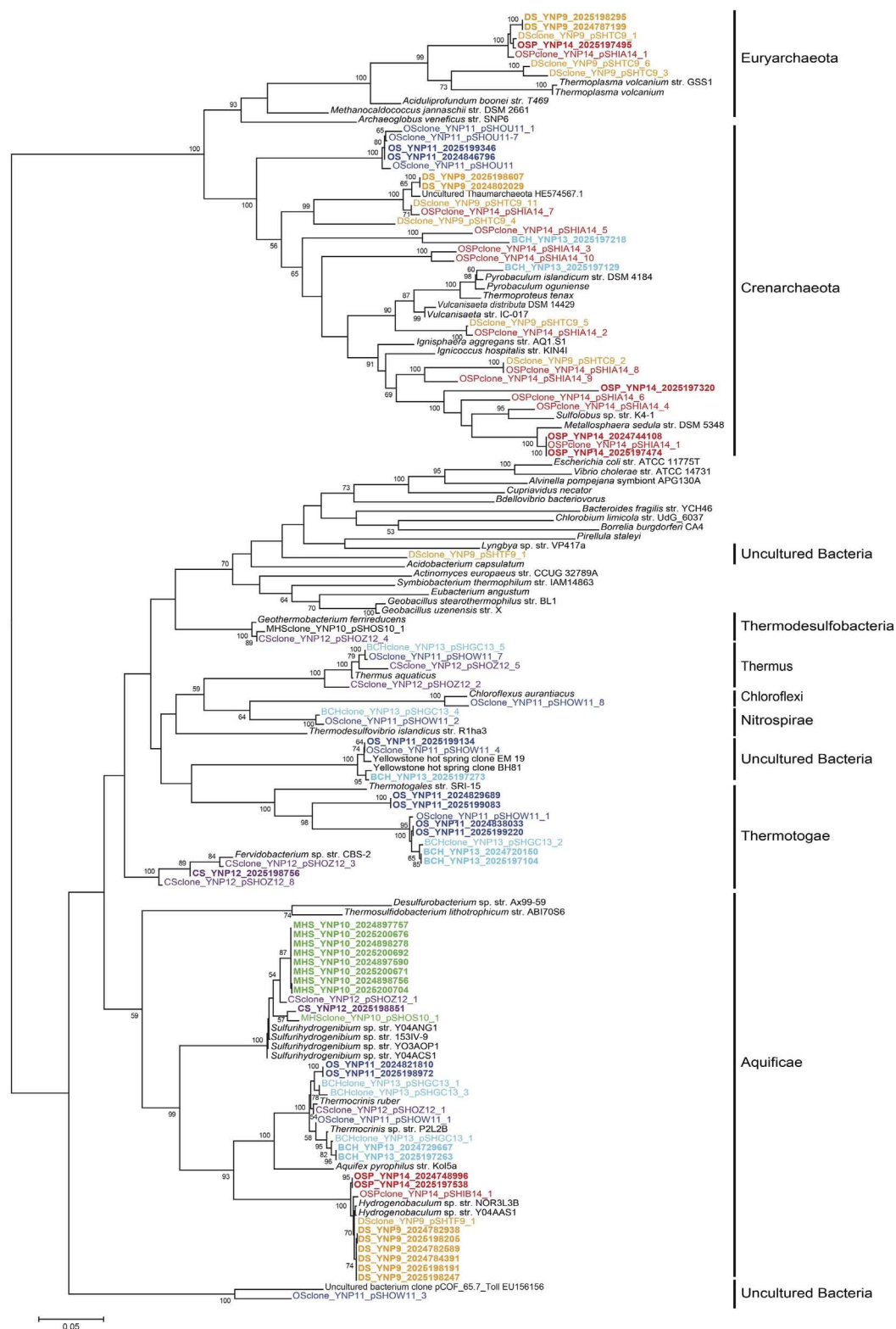
FIGURE A3 | Phylogenetic tree of all 16S rRNA genes observed in metagenome sequence (bold) combined with 16S rRNA genes amplified using PCR (up to 384 sequences per library) across six Aquificales streamer communities. Entries are colored by site: DG_9 = orange;

OSP_14 = red; MHS_10 = green; CS_12 = violet; OS_11 = dark-blue; BCH_13 = light-blue (core tree established using neighbor-joining methods with maximum-likelihood correction; metagenome entries added by parsimony).

FIGURE A4 | Distribution of PCR-amplified 16S rRNA genes observed in ribosomal panels across Aquificales streamer habitats. Universal archaeal and bacterial primers were used for each site and up to ~384 clones were sequenced per primer set.

FIGURE A5 | Consensus phylogenetic classification (AMPHORA; Wu and Eisen, 2008**) of assembled sequence using analysis of 31 housekeeping genes**. Although the Aquificales are the dominant members of each streamer sample, diverse and novel members of other bacterial and archaeal lineages are predicted to vary in abundance across sites.

FIGURE A6 | Hierarchical cluster analysis of relative gene abundances across six Aquificales streamer communities using all TIGRFAMs grouped into functional categories. Broad TIGRFAM categories include all cellular processes such as regulatory functions, energy metabolism, central C metabolism, mobile elements, transcription, cofactors and transporters. Data was standardized by functional category before clustering to avoid biasing analysis by a few categories with high gene abundance. Pearson correlation was used as the distance measure for average linkage agglomerative clustering.

FIGURE A7 | Functional categories with the most significant differences in relative gene abundance among the six Aquificales streamer communities [(A) biosynthesis of cofactors, prosthetic groups and carriers and (B) signal transduction: PTS; *p*-values = 0.0047 and 0.0057, respectively]. The relative proportion of sequences identified within each site are indicated (sites colored as before), and site pairs with similar geochemistry and Aquificales populations reveal replicate behavior in these specific TIGRFAMs.
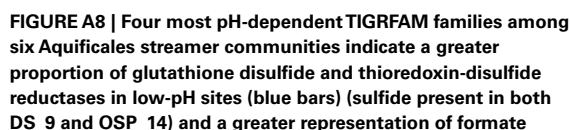
FIGURE A8 | Four most pH-dependent TIGRFAM families among six Aquificales streamer communities indicate a greater proportion of glutathione disulfide and thioredoxin-disulfide reductases in low-pH sites (blue bars) (sulfide present in both DS_9 and OSP_14) and a greater representation of formate dehydrogenases and NADH-plastoquinone oxidoreductases in higher-pH sites (orange bars) with (MHS_10, CS_12) or without (OS_11, BCH_13) sulfide. The relative proportion of sequences to all electron transport-associated genes is shown for each streamer community.

Phylogenetic and functional analysis of metagenome sequence from high-temperature archaeal habitats demonstrate linkages between metabolic potential and geochemistry

William P. Inskeep[1,2]*, Zackary J. Jay[1], Markus J. Herrgard[3], Mark A. Kozubal[1], Douglas B. Rusch[4], Susannah G. Tringe[5], Richard E. Macur[1], Ryan deM. Jennings[1], Eric S. Boyd[2,6], John R. Spear[7] and Francisco F. Roberto[8]

[1] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA
[2] Thermal Biology Institute, Montana State University, Bozeman, MT, USA
[3] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark
[4] Center for Genomics and Bioinformatics, Indiana University, Bloomington, IN, USA
[5] Department of Energy, Joint Genome Institute, Walnut Creek, CA, USA
[6] Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA
[7] Department of Civil and Environmental Engineering, Colorado School of Mines, Golden, CO, USA
[8] Newmont Mining Corporation, Englewood, CO, USA

Geothermal habitats in Yellowstone National Park (YNP) provide an unparalleled opportunity to understand the environmental factors that control the distribution of archaea in thermal habitats. Here we describe, analyze, and synthesize metagenomic and geochemical data collected from seven high-temperature sites that contain microbial communities dominated by archaea relative to bacteria. The specific objectives of the study were to use metagenome sequencing to determine the structure and functional capacity of thermophilic archaeal-dominated microbial communities across a pH range from 2.5 to 6.4 and to discuss specific examples where the metabolic potential correlated with measured environmental parameters and geochemical processes occurring *in situ*. Random shotgun metagenome sequence (~40–45 Mb Sanger sequencing per site) was obtained from environmental DNA extracted from high-temperature sediments and/or microbial mats and subjected to numerous phylogenetic and functional analyses. Analysis of individual sequences (e.g., MEGAN and G + C content) and assemblies from each habitat type revealed the presence of dominant archaeal populations in all environments, 10 of whose genomes were largely reconstructed from the sequence data. Analysis of protein family occurrence, particularly of those involved in energy conservation, electron transport, and autotrophic metabolism, revealed significant differences in metabolic strategies across sites consistent with differences in major geochemical attributes (e.g., sulfide, oxygen, pH). These observations provide an ecological basis for understanding the distribution of indigenous archaeal lineages across high-temperature systems of YNP.

Keywords: archaea, thermophilic archaea and bacteria, geochemistry, phylogeny, functional genomics

INTRODUCTION

Archaea are now recognized as the third domain of Life and are considered an ancestral link to the Eukarya (Woese and Fox, 1977; Woese et al., 1990). Although early research on organisms within this domain often focused on extreme thermophilic or halophilic organisms (Stetter, 2006), it is now established that archaea are not limited to extreme environments, and recent findings demonstrate the broad distribution of members of this domain across a wide range of environments including soil, human, marine, and aquatic habitats (Chaban et al., 2006; Auguet et al., 2009; Pester et al., 2011). The role of archaea in contemporary and past environments has been the subject of considerable debate, and members of this group have now been implicated as "missing links" in major global element cycling, including methanogenesis and nitrification in marine systems (DeLong, 2005; Schleper et al., 2005; Falkowski et al., 2008). The evolutionary history of these organisms has been defined in part by different environmental contexts involving variations in oxygen, iron, sulfur, and/or other major crustal or atmospheric components (e.g., CO_2, CH_4, H_2, NH_4). Metagenome sequencing of well-characterized, high-temperature geothermal systems with variable geochemistry provides a unique opportunity for understanding the metabolic attributes important to specific taxa within the *Archaea* (Inskeep et al., 2010). Contemporary thermophilic archaea in Yellowstone National Park (YNP) occupy a wide range of habitats with regard to dissolved oxygen, sulfide, iron, hydrogen, and pH (e.g., Inskeep

et al., 2005; Meyer-Dombard et al., 2005). Consequently, the field environments of YNP provide a natural laboratory where a subset of geochemical parameters varies across different springs, and selects for different types of microbial communities dominated by archaea.

Our understanding of archaeal diversity in Yellowstone hinges primarily on cultivation-based studies, although the past decade has seen an increase in genetic-based investigations. Much early interest centered on members of the order Sulfolobales (phylum Crenarchaeota), and several of these acidophiles were cultivated as the first recognized archaea (Brock et al., 1972; Brierley and Brierley, 1973). These organisms are distributed globally in hydrothermal vents or solfataras and have generally been cultivated at low pH (1–3) using reduced forms of S, Fe and complex C as electron donors under aerobic to microaerobic conditions. The oxidation of Fe(II) is less studied in the Sulfolobales, however, recent work in YNP shows that *Metallosphaera yellowstonensis* populations occupy acidic Fe(III)-oxide mats and contain genes required to oxidize Fe(II) via a *fox* terminal oxidase complex (Kozubal et al., 2008, 2011). The distribution of members of the orders Desulfurococcales and Thermoproteales has not been studied with great detail in YNP, however, these organisms have generally been observed in sulfidic sediments and in higher pH systems (pH 3–8) compared to the Sulfolobales (Barnes et al., 1994; Inskeep et al., 2005; Meyer-Dombard et al., 2005). Two acidophilic Desulfurococcales were isolated from hypoxic sulfur sediments in Norris Geyser Basin (YNP) as obligate anaerobes growing on complex carbon sources and elemental S as an electron acceptor (Boyd et al., 2007). In addition, we have recently obtained a *Pyrobaculum*-like isolate from YNP that also grows with elemental sulfur and complex carbon sources (Macur et al., 2013). However, the distribution of different types of Thermoproteales in YNP and their role in community function is not known. Other less understood members of the domain *Archaea* also occur in YNP geothermal systems, and include members of the Euryarchaeota (e.g., Segerer et al., 1988), Korarchaeota (Elkins et al., 2008; Miller-Coleman et al., 2012), Nanoarchaeota (Clingenpeel et al., 2011), and Thaumarchaeota (Brochier-Armanet et al., 2008; de la Torre et al., 2008; Hatzenpichler et al., 2008; Spang et al., 2010; Beam et al., 2011; Pester et al., 2011). Factors responsible for the distribution of these and other novel phyla to be discussed herein have not been determined, and in many cases, cultivated relatives of these organisms do not exist to inform on their physiology or function *in situ*.

Seven high-temperature geothermal systems were sampled to represent a range of geochemical conditions and to investigate effects of pH, dissolved gases (sulfide, oxygen, hydrogen), and Fe on the distribution and function of archaea in YNP. The primary objectives of this study were to (i) determine the community structure and function of high-temperature archaeal-dominated microbial communities across a pH range from 2.5 to 6.4 utilizing metagenome sequencing, (ii) compare inferred functional attributes across sites and different phylotypes using assembled metagenome sequence to obtain protein family abundances and functional gene content, and (iii) evaluate environmental parameters and geochemical processes across sites that may define the distribution patterns of thermophilic archaea in YNP.

RESULTS

GEOCHEMICAL CONTEXT AND ELEMENT CYCLING

The thermal (70–85°C) sites discussed here range in pH from 2.5 to 6.4 (**Table 1**), and represent several of the major chemotrophic habitat types observed in Yellowstone's geothermal basin. Six of the communities are hypoxic and contain elemental sulfur (and/or dissolved sulfide). Consequently, there is significant potential for chemotrophic metabolism based on sulfur oxidation-reduction reactions in these habitats (e.g., Amend and Shock, 2001; Inskeep et al., 2005). The acidic *Crater Hills* (CH_1) and *Nymph Lake* (NL_2) sites are highly turbulent pools that contain suspended solids of elemental sulfur and SiO_2, and only low levels of dissolved sulfide (e.g., $<5 \mu M$) (**Figure 1**). Metagenome sequence was also obtained from four mildly acidic, sulfidic sediments at *Monarch Geyser* (MG_3), *Cistern Spring* (CIS_19), *Joseph's Coat Hot Springs* (JCHS_4), and *Washburn Springs* (WS_18) (**Table 1**). The presence of dissolved sulfide results in the deposition of elemental sulfur in all of these sites (Xu et al., 1998, 2000; Macur et al., 2013); however, pyrite and amorphous Fe-sulfides are important solid phases at JCHS_4 and WS_18, and significant [>1% (w/w)] amounts of stibnite (Sb_2S_3) and orpiment (As_2S_3) are also present in sediments from JCHS_4 (**Table 1**; **Figure 1**). Metagenome analysis of a similar JCHS_4 sediment sample obtained 1 year prior to the current study (Inskeep et al., 2010) showed a community dominated by two different Thermoproteales populations and one member of the order Desulfurococcales. Prior 16S rRNA gene surveys have also suggested that crenarchaeal populations similar to those observed in JCHS are important in the sulfur sediments at CIS_19 and MG_3 (Macur et al., 2013). The predominant dissolved ions in the highly sulfidic (hypoxic) system at *Washburn Springs* (WS_18) are ammonium and sulfate ($\sim$24 mM NH_4, 17 mM sulfate), although high levels of dissolved hydrogen (450 nM), methane (14.9 μM), inorganic carbon (DIC) (5.5 mM), and organic carbon (DOC) (0.3 mM) were also measured in this study (**Table 1**; Table S2 in Supplementary Material, Inskeep et al., 2013).

The oxidation of dissolved sulfide with oxygen occurs via abiotic and/or biotic processes (Xu et al., 1998; Friedrich et al., 2005; Ghosh and Dam, 2009), and results in the deposition of elemental sulfur commonly observed within sulfidic systems of YNP (**Figure 1**). Moreover, the products of sulfide oxidation are pH dependent and vary due to kinetic favorability of specific reaction steps (Xu et al., 1998, 2000). Thiosulfate concentrations in geothermal channels are often higher at intermediate pH values (5.5–7) due to greater stability compared to low pH, where thiosulfate disproportionation to elemental sulfur and sulfite (SO_3^{2-}) can occur rapidly over time scales of seconds-minutes (Xu et al., 1998; Nordstrom et al., 2005). The impact of higher thiosulfate on archaeal communities is not known, but may provide flexibility for energy conservation through oxidative or reductive processes (Amend and Shock, 2001). The morphology of elemental S in these sulfidic sediments is generally rhombohedral, but spheres of variable diameter are also found, and cells have been observed adhering to these S minerals (e.g., Inskeep et al., 2010; Macur et al., 2013). The role of biota in the mineralization of FeS_2 and Sb_2S_3 in JCHS_4 is not known, but does not require a reductive step

Table 1 | Sample location, total dissolved (<0.2 μm) geochemical parameters, and predominant solid phases associated with high-temperature, archaeal-dominated microbial communities in Yellowstone National Park (YNP).

Location	T °C	pH	I	SO$_4$	NH$_4$	DIC	DS	DO	S$_2$O$_3$	As	Fe	CH$_4$	H$_2$	Solid Phases[2]	YNP Thermal Inventory	Coordinates
			-- mM --							μM		nM				
Crater Hills (CH_1)	76	2.5	18	8.7	7.2	1.4	2	<3	<3	2	230	0.34	67	S^0, SiO$_2$	'Alice Spring' CHANN041	44° 39' 11.823" N. Lat 110° 29' 5.067" W. Lon
Nymph Lake (NL_2)	88	~4[3]	17	1.3	0.23	0.6	3	<3	<3	9	1.5	<0.01	84	S^0, SiO$_2$	West of Nymph Lake	44° 45' .8.354" N. Lat 110° 43' 31.413" W. Lon
100 Springs Plain (OSP_8)	72	3.3	14	1.1	0.09	0.2	<1	30	<3	28	36	<0.01	34	Fe(AsO$_4$)(OH)$_x$	NHSP115	44° 43' 58.953" N. Lat 110° 42' 32.374" W. Lon
Monarch Geyser (MG_3)	80	4.0[3]	16	1.3	0.28	1.4	25	<3	55	19	3.6	0.61	375	S^0, SiO$_2$	NBB218	44° 43' 27.453" N. Lat 110° 42' 20.207" W. Lon
Cistern Spring (CIS_19)	78	4.4[3]	11	0.7	0.30	1.2	20	<3	200	19	0.2	0.30	460	S^0, SiO$_2$, KAl$_3$SO$_4$(OH)$_6$	NBB201	44° 43' 23.092" N. Lat 110° 42' 14.489" W. Lon
Joseph's Coat Hot Springs (JCHS_4)	80	6.1	23	3.8	8.4	0.2	20	<3	700	140	1.6	0.19	65	Sb$_2$S$_3$, FeS$_2$, As$_2$S$_3$, S^0, SiO$_2$	Scorodite Spring JCS083	44° 44' 20.984" N. Lat 110° 19' 28.179" W. Lon
Washburn Springs (WS_18)	76	6.4	42	16.7	24	5.5	160	<3	<3	1.6	1.1	14.9	450	S^0, FeS$_2$, SiO$_2$	Washburn Spring	44° 45' 53.62" N. Lat 110° 25' 49.26 "W. Lon

I, ionic strength calculated from aqueous geochemical modeling at sample temperature; DIC, dissolved inorganic C; DS, dissolved sulfide; DO, dissolved oxygen; S$_2$O$_3^{2-}$, thiosulfate; DO, dissolved oxygen.

[1]Predominant solid phases determined using scanning electron microscopy (FE-SEM) coupled with energy dispersive analysis of X-rays (EDAX) and X-ray diffraction (XRD).

[2]pH values at these sites have been noted to vary (MG_3; pH ~ 4–4.5; CIS_19; pH ~ 4.4–4.9). Also, see USGS reports (e.g., Ball et al., 1998, 2002; McCleskey et al., 2005).
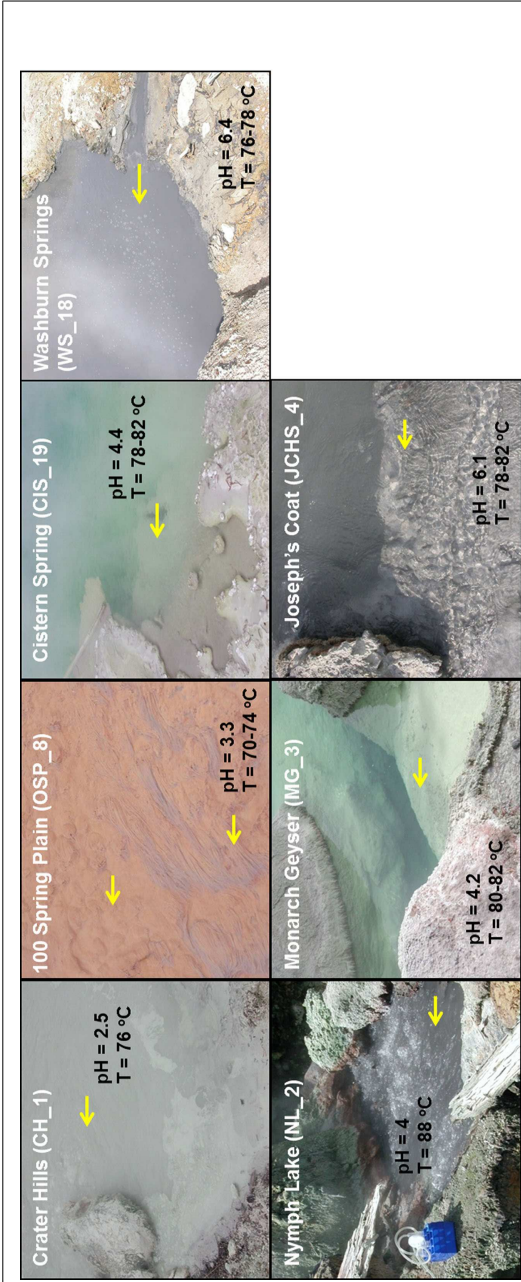
FIGURE 1 | Site photographs of high-temperature chemotrophic systems in Yellowstone National Park (YNP) selected for metagenome sequencing and described in the current study. The sites cover a range in geochemical conditions including sulfur-rich sediments from pH 2.5 (CH_1) to 6.4 (WS_18) and higher oxygen environments where the oxidation of ferrous-iron results in the formation of Fe(III)-oxide microbial mats (OSP_8) [pH and temperature measured on site; yellow arrows indicate sampling locations; all site locations referenced in **Table 1**].

since reduced constituents necessary for the formation of these minerals are already present in the source waters [i.e., Fe(II), DS; **Table 1**].

To contrast geochemical systems heavily influenced by sulfide and elemental S, two microbial mat samples were obtained from an acidic spring (pH 3.3–3.5) in the *One Hundred Spring Plain* (OSP), and included both a filamentous "streamer" community (OSP_14) and an Fe(III)-oxide microbial community (OSP_8) from the same geochemical environment. The "streamer" communities are infrequently distributed on top of the Fe-oxide mats from 70 to 80°C in shallow (∼1 cm), high-velocity outflow channels (**Figure 1**) (Takacs-Vesbach et al., 2013). The Fe-oxide mats (OSP_8) form as a result of oxidation of dissolved Fe(II) and subsequent deposition of amorphous, high-arsenate, Fe(III)-oxides (Inskeep et al., 2004; Macur et al., 2004). Prior metagenome and mRNA analysis of Fe(III)-oxide samples from *Beowulf Spring* confirmed the importance of *Metallosphaera*-like organisms and Fe(II)-oxidizing genes within these systems (Inskeep et al., 2010; Kozubal et al., 2011, 2012). The high-temperature Fe(III)-oxide mineralizing environments contain higher oxygen contents and support significantly greater archaeal diversity than low-pH (i.e., pH ∼ 2–6) sulfidic systems (Inskeep et al., 2005, 2010; Kozubal et al., 2012).

PHYLOGENETIC ANALYSIS OF METAGENOME SEQUENCE

Analysis of individual sequences (average length ∼800 bp) across these chemotrophic environments revealed systems inhabited by as few as one dominant population type (e.g., NL_2) to those containing significant archaeal diversity (e.g., OSP_8, WS_18). Combined phylogenetic (MEGAN) and G + C content (%) analysis of all individual sequences revealed the predominant phylotypes represented in each site (**Figure 2**). Sequences from the acidic and sulfidic sites (CH_1 and NL_2) were dominated by members of the order Sulfolobales (family *Sulfolobaceae*). The single dominant population type in NL_2 with a G + C content of ∼53.5% (referred to here as Type 1 Sulfolobales) was also one of two main population types present in CH_1 (**Figure 2**). A second *Sulfolobaceae* population in CH_1 (Type 2 Sulfolobales, G + C = 38%) was also found in the sulfur sediments at *Cistern Spring* (CIS_19) (along with less-dominant Type 1 populations). Conversely, *Monarch Geyser* (MG_3) contained a smaller number of Sulfolobales-like sequence reads with G + C contents near 60% (**Figure 2**). Sequence reads in CH_1 and NL_2 were classified at two different levels of phylogenetic resolution (family and genus) to illustrate that the total archaeal reads (gray) were nearly all related to the family *Sulfolobaceae,* but that considerably fewer sequences were highly related to the reference genomes of *Sulfolobus* spp. (**Figure 2**).

Metagenome sequences from the less acidic, sulfidic sediments of *Monarch Geyser* (MG_3), *Cistern Spring* (CIS_19), and *Joseph's Coat Springs* (JCHS_4) were dominated by populations within the orders Thermoproteales and Desulfurococcales (phylum Crenarchaeota). The G + C content (%) of the predominant Desulfurococcales population was consistently ∼59% across all sites (**Figure 2**), while the two Thermoproteales populations exhibited G + C contents of either ∼48% (Type 1, light-blue;

Caldivirga/Vulcanisaeta; e.g., Itoh et al., 1999, 2002), or ∼62.5% (Type 2, dark-blue; *Pyrobaculum* clade; e.g., Volkl et al., 1993; Fitz-Gibbon et al., 2002). The sulfidic sediments at WS_18 (pH 6.4) also contained representatives of the Thermoproteales (e.g., a Type 2 *Pyrobaculum*-like population, and a Type 3 *Thermofilum*-like population), as well as significant fractions of *Sulfurihydrogenibium* spp. (Aquificales), Thermodesulfobacteria, and members of the Korarchaeota (G + C ∼ 40–50%) (**Figure A1** in Appendix).

The acidic Fe(III)-oxide microbial mat from *One Hundred Spring Plain* (OSP_8) contained considerable archaeal diversity, which correlated with low sulfide (**Table 1**) and higher levels of dissolved oxygen [i.e., 30–40 μM O_2(aq)]. At least five distinct populations are evident in the G + C (%) distribution plot (**Figure 2**); phylogenetic analysis showed that these peaks corresponded to phylum Euryarchaeota (G + C ∼ 31%), *Vulcanisaeta/Caldivirga* sp. (G + C ∼ 47.8%), *Metallosphaera* sp. (G + C ∼ 48.2%), *Acidilobus* sp. (G + C ∼ 57.5%), and a novel archaeal population with a G + C content of 32.5 ± 2%. The sequences of the novel archaeal Group I population have been proposed (Kozubal et al., 2013) to represent a new phylum-level lineage in the *Archaea*, the Geoarchaeota. The sequence reads similar to *Aeropyrum*-like populations are actually more closely related to the recently released genome of *Acidilobus saccharovorans* (Prokofeva et al., 2009; Mardanov et al., 2010) and draft genome sequence for *A. sulfurireducens* (Boyd et al., 2007; Inskeep et al., 2010), but the *Aeropyrum* reference correctly identifies this population as a member of the order Desulfurococcales.

PHYLOGENETIC ANALYSIS OF METAGENOME SEQUENCE ASSEMBLIES

Assembly of individual sequence reads resulted in large contigs and scaffolds for several of the predominant archaeal populations present in these sites. Sequence assemblies for each sample were evaluated using nucleotide word frequencies (NWF) (Teeling et al., 2004) combined with Principal Components Analysis (PCA; **Figure 3**). This technique can often resolve organisms at the genus-species level because of unique sequence character including G + C content (%) and codon usage bias (Teeling et al., 2004; Inskeep et al., 2010). Assemblies from archaeal sites were separable to a large extent using principal component analysis (**Figure 3A**). The PCA plot is also shown with corresponding phylogenetic analysis of contigs Automated Phylogenetic Inference System (APIS; Badger et al., 2006) at the order-level (**Figure 3B**), and the genus-level (**Figure 3C**). The majority of predominant phylotypes present across these communities were delineated using this approach.

The predominant sequence assemblies corresponded to the major peaks observed in the G + C distribution plots (**Figure 2**). For example, the acidic sites were dominated by Sulfolobales populations, and three major lineages within this order were identified across sites (Table S2 in Supplementary Material). The acidic elemental sulfur-rich sediments (CH_1 and NL_2) contained one and two major Sulfolobales types, respectively. These populations are related to *Sulfolobus* spp. (Type 1, G + C content = 52–54%) and *Stygiolobus* spp. (Type 2, G + C content = 38%) based on full-length genes identified in the assembled sequence (including the 16S rRNA gene) (**Figure 4**; Tables S1 and S2 in Supplemental Material). The higher pH (pH 4–6) sulfur sediments from MG_3,
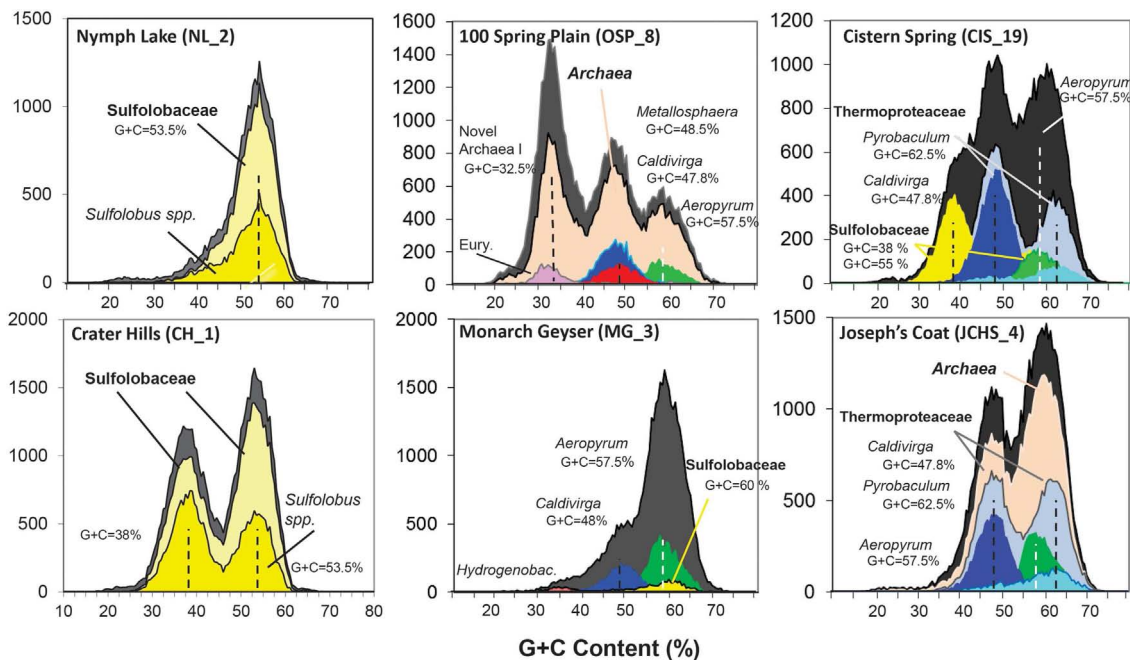
FIGURE 2 | Frequency plot of the G + C content (%) of random shotgun sequence reads (Sanger) obtained from archaeal habitats in Yellowstone National Park (YNP). Phylogenetic classification of each sequence (∼800 bp) was performed using MEGAN ("blastx"), which shows the predominant populations that contribute to metagenome sequence in these environments (dark-gray = total reads, pink = domain *Archaea* (shown in sites OSP_8 and JCHS_4 only), yellow = *Sulfolobaceae* (identity to *Sulfolobus* sp. also shown), red = *Metallosphaera* sp., green = *Aeropyrum pernix*, violet = Euryarchaeota, light-gray = Thermoproteaceae, dark-blue = *Caldivirga* sp., light-blue = *Pyrobaculum* sp.). Site WS_18 is shown in **Figure A1** in Appendix.
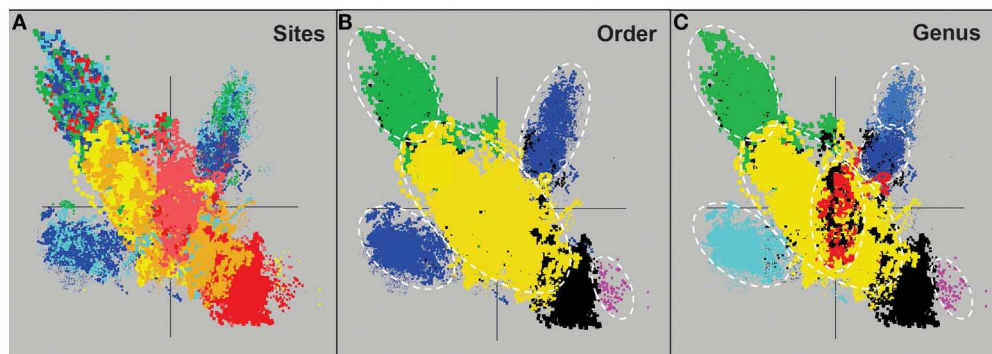
FIGURE 3 | Nucleotide word frequency PCA plots of metagenome assemblies (>5 kb contigs) from high-temperature, archaeal-dominated geothermal environments in YNP. **(A)** Sites: *Crater Hills* (CH_1) = gold; *Nymph Lake* (NL_2) = yellow; *Monarch Geyser* (MG_3) = green; *Joseph's Coat Hot Springs* (JCHS_4) = dark-blue; *Cistern Spring* (CIS_19) = light-blue; *One Hundred Spring Plain* "streamer" (OSP_14) = light-red; *One Hundred Spring Plain* Fe-oxide mat (OSP_8) = red; **(B)** Order Level: The identical PCA orientation was maintained as presented in A, but with phylogenetic analysis of contigs to the closest reference genome (yellow = Sulfolobales; green = Desulfurococcales; blue = Thermoproteales; purple = Nanoarchaeum; black = unassigned); and **(C)** Genus Level (yellow = *Sulfolobus*; red = *Metallosphaera*; green = *Aeropyrum*; light-blue = *Pyrobaculum*; dark-blue = *Vulcanisaeta*; gray-blue = *Caldivirga*; purple = *Nanoarchaeum*; black = unassigned).

CIS_19, and JCHS_4 also contain variable amounts of sequence corresponding to Sulfolobales populations (**Figures 3A,B**); the G + C peak in CIS_19 at 36–38% was similar to the *Stygiolobus* population in CH_1 (**Figure 2**). Conversely, the Fe-oxide mats (OSP_8) and Fe-oxide "streamer" community (OSP_14) were the only sites to contain sequence data corresponding to

M. yellowstonensis-like populations (**Figure 3C**). This aerobic organism has been shown to oxidize Fe(II) using a different terminal oxidase complex (*fox*) than used for S oxidation (*dox*) (Bathe and Norris, 2007) and can generate sufficient energy for growth by oxidizing large amounts of Fe(II) (Kozubal et al., 2011).

The main sequence clusters observed in MG_3, CIS_19, and JCHS_4 were related to members of the orders Desulfurococcales and Thermoproteales (**Figure 3**). Phylogenetic analysis consistently showed two major types of Thermoproteales in both CIS_19 and JCHS_4 (and to a lesser extent in MG_3 and OSP_8) (**Figure 3B**), corresponding to *Caldivirga/Vulcanisaeta*-like (Type 1 Thermoproteales) and *Pyrobaculum*-like (Type 2 Thermoproteales) organisms (**Figure 3C**). Similar *Caldivirga/Vulcanisaeta* and similar *Acidilobus*-like populations were also observed in the non-sulfidic, Fe-oxide mat (OSP_8) (based on NWF PCA, sequence similarity, G + C content, and functional analysis). Consequently, although these populations were clearly the main community members detected in hypoxic sulfur-rich sediments from pH 4 to 6, (Jay et al., 2011) they also appeared in Fe-oxide mats where sulfide and elemental sulfur are generally below detection.

At least four major archaeal populations were identified in the Fe-oxide mat (OSP_8) using G + C content and NWF PCA analysis: *M. yellowstonensis*, *Vulcanisaeta* spp., *Acidilobus* spp. and a "novel archaeal Group I" population belonging to the proposed phylum, Geoarchaeota (Kozubal et al., 2013). The deeply rooted phylogenetic position of the 16S rRNA gene (**Figure 4**) was consistent with analysis of other single-copy genes (e.g., RNA polymerases, gyrases, transcriptional factors, etc.) identified within the assembled sequence (Table S4 in Supplementary Material). The metagenome sequence of NAG1 was only distantly related to other reference genomes; amino acid identities relative to currently available reference genomes generally ranged from 40 to 60%, and closest relatives of individual genes included members of the domain *Archaea* as well as *Bacteria* (Kozubal et al., 2013). The Geoarchaeota (NAG1) population was the most abundant community member in OSP_8, which resulted in excellent contig coverage (average ~6×), and a total scaffold length of ~1.7 Mb in only eight scaffolds (Kozubal et al., 2013). The Fe-oxide community (OSP_8) also contained several other archaea (although at lower coverage) including relatives of the Euryarchaeota (distantly related to the Thermoplasmatales), Nanoarchaeota, Crenarchaeota (i.e., other Sulfolobales), as well as the Candidate phylum Thaumarchaeota (Brochier-Armanet et al., 2008; Beam et al., 2011).

Nanoarchaeal sequence

Assembled sequence distantly related to *Nanoarchaeum equitans* (Huber et al., 2002; Waters et al., 2003) was found in several archaeal-dominated microbial communities (**Figure 4**; Table S1 in Supplemental Material). Partial 16S rRNA gene sequences (among other single-copy genes; Table S4 in Supplemental Material) corresponding to the Nanoarchaeota were observed in assembled sequence from sulfidic sediments (NL_2, JCHS_4) and Fe-oxide mats (OSP_8), although these are only distantly related to *N. equitans* (~82–84% similarity, Table S1 in Supplemental Material). Given that *Ignicoccus hospitalis* is not an important member of these archaeal communities, either other hosts are important to these nanoarchaea, or they may be free-living. Insufficient coverage of these novel nanoarchaea does not allow a thorough genomic evaluation; however, nearly 100 kb of assembled sequence was obtained for nanoarchaea present in the Fe-oxide mat samples (OSP_8 and 14). The average G + C content of the nanoarchaeal sequence is ~27%, considerably lower than observed for *N.*
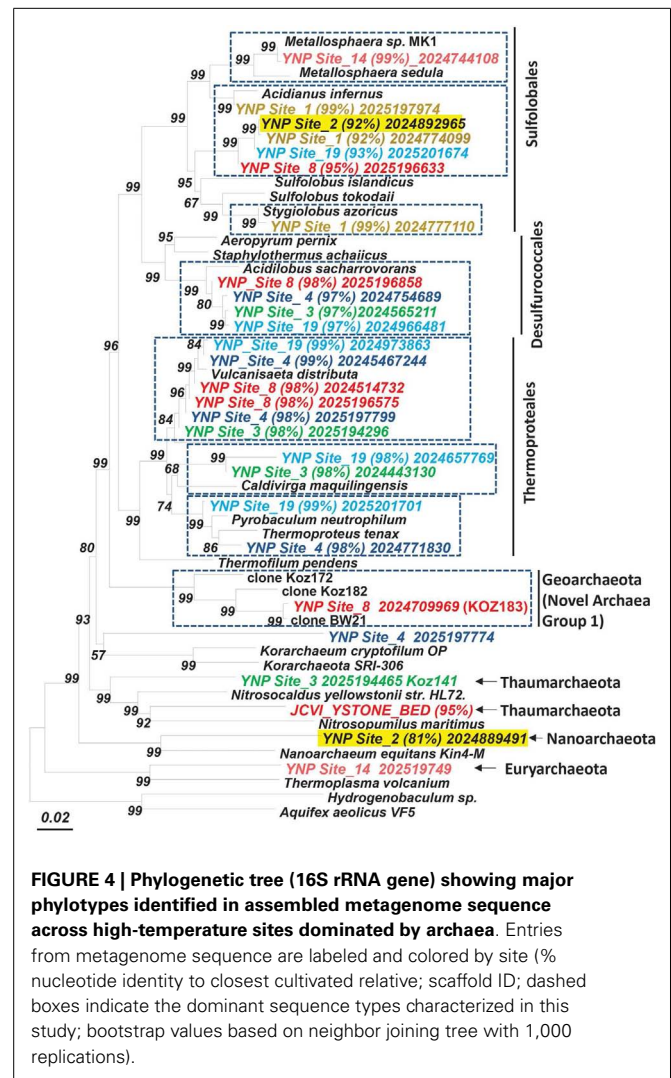
FIGURE 4 | Phylogenetic tree (16S rRNA gene) showing major phylotypes identified in assembled metagenome sequence across high-temperature sites dominated by archaea. Entries from metagenome sequence are labeled and colored by site (% nucleotide identity to closest cultivated relative; scaffold ID; dashed boxes indicate the dominant sequence types characterized in this study; bootstrap values based on neighbor joining tree with 1,000 replications).

equitans (31.6%). Further work will be necessary to fully appreciate the diversity of nanoarchaea in thermal systems of YNP and determine whether the extensive distribution of different nanoarchaeal sequences (Hohn et al., 2002; Casanueva et al., 2008; Clingenpeel et al., 2011) implies a corresponding diversity of host species.

Assembly of viral genomes

In total, 10 scaffolds from the archaeal-dominated samples were classified as "viral," based on phylogenetic analysis of known viruses (**Table 2**; **Figure 5**). Although the similarity of these scaffolds to known viruses varied considerably, the Thermoproteus spherical-like viruses found in sites NL_2, JCHS_4, and CIS_19 are highly similar to and nearly the same length as known isolates. Others such as scf_6649105 and scf_5653402 had very weak matches to predicted viral proteins. The viruses found in these samples are related (if only distantly) to other archaeal viruses and thus consistent with expectation. Two sets of viral scaffolds (Group A and B) were found in more than one sample (**Table 2**), and in both cases the sequences were highly similar (92%+ nucleotide identity). CRISPR regions including both spacer regions and direct

Table 2 | Scaffolds with similarity to known viruses.

	Viral Scaffold Characteristics				**CRISPR Information**		**Similarity Information**			
Sample	**Scaffold ID[1]**	**%G+C**	**Scaffold Length**	**Number of N's**	**Spacer Regions**	**CRISPR Site**	**Similarity Group**	**Matching Scaffolds**	**Matching Viral Genome**	**Genome Length**
YNP_02	scf_6656549	56.3	18010	0	1	2	A	scf_5653343 scf_6606346	Thermoproteus tenax spherical virus 1	20993
YNP_03	scf_6648705	41.4	12461	20	4	3[*]	B	scf_5653210	Sulfolobus islandicus rod-shaped virus 1	32308
YNP_03	scf_6649099	38.2	34046	0	0	-	-		Acidianus filamentous virus 9	41172
YNP_03	scf_6649105	40.5	16038	0	1	3	-		Sulfolobus islandicus filamentous virus	40900
YNP_04	scf_6606346	56.5	15839	0	0	-	A	scf_5653343 scf_6656549	Thermoproteus tenax spherical virus 1	20993
YNP_04	scf_6606202	58.5	27771	653	9	4	-		Acidianus rod-shaped virus 1	24655
YNP_14	scf_6615532	27.6	34310	668	5	14[*]	-		Sulfolobus islandicus rod-shaped virus 1	32308
YNP_19	scf_5653343	56.1	18048	20	1	19	A	scf_6656549 scf_6606346	Thermoproteus tenax spherical virus 1	20993
YNP_19	scf_5653210	42.7	13852	567	5	19	B	scf_6648705	Sulfolobus islandicus rod-shaped virus 2	35450
YNP_19	scf_5653402	50	16801	0	1	19	-		Acidianus filamentous virus 2	31787

[1]Complete scaffold ID (_ = 111868) retained as an identifier in sequence assemblies deposited with IMG/M.

The source of the assembled reads, %G+C, scaffold length, and the number of ambiguous bases (N's) in "viral-related" scaffolds are shown. Scaffolds showing nucleotide sequence similarity to each other are identified by the "Similarity Groups" (A, B), where the similar scaffolds are listed. The best match to known viral genomes (BLASTX to the NRAA database) is shown along with the length of that genome. Sequence from eight of these scaffolds contained matches to CRISPR spacer regions (three mismatches allowed).

FIGURE 5 | Nucleotide word frequency PCA scatter plots of non-viral (black) versus viral (red) scaffolds identified within the metagenome assemblies of archaeal-dominated sites (Table 2 provides additional details on the characteristics of viral scaffolds).
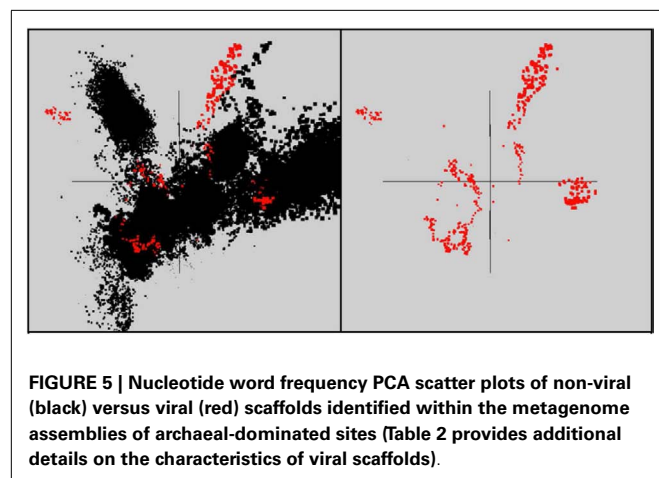
repeats (DR) (Grissa et al., 2007; Makarova et al., 2011) were predicted from these assemblies as well as assemblies generated from the same habitat types sampled ~1 year prior to the current study (Inskeep et al., 2010). Near perfect alignments were found between CRISPR spacer regions and 8 of the 10 viral-like scaffolds (**Table 2**). CRISPR spacer regions from the prior project matched two of the scaffolds identified here implying some continuity between the viral and microbial populations. A total of 5,435 CRISPR spacers were identified from the archaeal samples, and only 16 of these matched the scaffolds annotated as viral (one mismatch allowed). Even relaxing the alignment parameters (three mismatches allowed) only increased this number modestly to 26. A total of 382

spacers have matches to sequences not annotated as viral. Those with multiple matches were examined, but could not be verified as viral due to their short length. Consequently, although the majority of spacer regions identified within CRISPR elements were not recognized as viral, this may be due to our inability to recognize novel and potentially dynamic viral sequence.

We found similar patterns in the diversity and distribution of CRISPR DR across the archaeal-dominated sites (Table S3 in Supplementary Material) corresponding to the dominant phylotypes present (e.g., **Figure 4**). For example, DR sequences classified as most similar to reference Sulfolobales sequence were found distributed across the archaeal sites containing these organisms (CH_1, NL_2, MG_3, OSP_8, and CIS_19). Moreover, sites dominated by a particular phylotype such as the Sulfolobales (e.g., CH_1, NL_2) only contained Sulfolobus-like DR sequences. Sites containing greater crenarchaeal diversity such as CIS_19 and JCHS_4 reveal a significant number of DRs contributed from *Sulfolobus*, *Caldivirga*, and *Pyrobaculum*-like populations, consistent with the dominant populations identified using standard phylogenetic markers. DRs were not found in any of the four replicate Desulfurococcales populations observed from pH 3 to 6 (sites OSP_8, MG_3, JCHS_4, CIS_19). The different "*Thermofilum*-like" DR sequences observed in OSP_8 and WS_18 (Table S3 in Supplementary Material) are each unique to the study; those from site OSP_8 are contributed from the NAG1 population (Candidate phylum Geoarchaeota; Kozubal et al., 2013).

Phylogenetic summary
Manually curated scaffolds/contigs corresponding to seven major archaeal phylotypes (or ~15 *de novo* assemblies including

replicates) were obtained from the metagenome sequence (Table S2 in Supplementary Material). Replicate *de novo* assemblies of similar phylotypes were obtained for *Metallosphaera*-like populations (OSP_8 and 14), *Caldivirga/Vulcanisaeta* types (MG_3, CIS_19, JCHS_4, and OSP_8), *Acidilobus*-like organisms (MG_3, CIS_19, JCHS_4, and OSP_8), *Sulfolobus and Stygiolobus*-like populations (CH_1, NL_2, and CIS_19), and *Pyrobaculum* types (CIS_19, JCHS_4, and WS_18). The amount of assembled sequence obtained for many of these indigenous archaeal populations was greater than 1 Mbp (Table S2 in Supplementary Material), and may represent close to expected genome sizes based on sequenced archaeal relatives. Each contig within each sequence cluster (NWF analysis) was carefully screened using G + C content (%) combined with BLAST scores and functional relevance. Fragment recruitment plots, coverage estimates and evaluation of single-copy genes suggests that the assembled sequence represents near complete (>90%) genomic sequence for several (6–8) of these phylotypes. A modest survey of single-copy genes corresponding to the predominant populations present in these archaeal sites provides a summary of the possible completeness represented in the assembled sequence (Table S4 in Supplementary Material) and also provides insight regarding the possible variation existing within closely related populations. Given that our analysis was limited to near-full length genes, additional single-copy genes corresponding to these populations may be present in small contigs or individual sequences that did not assemble well, and these may also be extremely useful for understanding more variable regions among individuals comprising these phylotypes.

PROTEIN FAMILY ANALYSIS OF ARCHAEAL COMMUNITIES

One of the primary aims of the study was to identify specific metabolic attributes of individual archaeal populations found distributed across chemotrophic habitats, and determine if functional attributes of these communities correlated with specific geochemical properties. Moreover, a thorough evaluation of metabolic capability provides a direct understanding of which oxidation-reduction reactions may be driving productivity in these chemotrophic habitats, and how the functional capabilities of different and/or very similar organisms may vary in response to specific environmental parameters. The abundances of all proteins identified in the assembled metagenome sequence data were evaluated using PCA and hierarchical clustering to compare relative differences and/or similarities among sites. PCA of relative gene abundances across all TIGRFAMS grouped into functional categories showed strong similarity between individual sites with similar phyla (**Figure 6A**). Factor 1 (accounting for ∼74% of the relative TIGRFAM variation across sites) separated sites based roughly on the relative abundance of Sulfolobales (which also tracks with site pH), where CH_1 and NL_2 were dominated by only two population types of Sulfolobales and WS_18 contained little to no Sulfolobales. Sites that contained a greater abundance of Desulfurococcales and Thermoproteales (mildly acidic sulfur sediments, MG_3, CIS_19, JCHS_4; Macur et al., 2013) also clustered together. Site OSP_8 (oxic, no sulfide) contained multiple archaeal populations (including the new Geoarchaeota) and plotted separate from all other sites. Principal components factors 2 and 3 were less important in describing

functional variation across the archaeal sites (only 13 and 7%, respectively); however, PC3 results in separation of OSP_8 relative to all other archaeal sites. OSP_8 was the only oxic habitat included in this study and was the only site that contained the NAG1 population (Candidate phylum Geoarchaeota, Kozubal et al., 2013).

The TIGRFAM categories responsible for observed differences across these archaeal sites were evaluated in more detail using hierarchical clustering (**Figure 7**). The site clustering is consistent with PCA separation (**Figure 6A**), and correlates with environmental factors including pH and dissolved sulfide/oxygen. Examples of TIGRFAM categories most different across sites include processes related to sulfur metabolism, peptide secretion, electron transport, fermentation, biosynthesis of cofactors, and routine cellular processes including cell division, sporulation, and motility (**Figure 7**). Given the substantial phylogenetic differences across sites, it is not surprising that the relative differences within and across TIGRFAM categories retained these signatures. However, it can be difficult to appreciate specific functional differences using broad TIGRFAM categories, and each protein identified in the metagenome sequence must be studied independently to verify putative functional assignment and understand aspects related to gene neighborhood and pathway context.

More detailed comparisons among sites were made using protein hits confined to specific TIGRFAM categories. For example, the relative abundances of TIGRFAM assignments within the category "Electron Transport" resulted in consistent site separation as observed using broad categories (**Figure 6B**), however, the specific TIGRFAMs in this category provide greater insights into microbial processes that are influenced directly by geochemical conditions such as the presence of sulfide versus oxygen. Indeed, numerous "Electron Transport" TIGRFAMs with the greatest relative differences across sites relate to the types of cytochromes, oxygen reductases, sulfur reductases, hydrogenases, or other respiratory proteins present in the metagenome sequence. For example, the ubiquity of heme Cu oxidases (Type 1 HCO; García-Horsman et al., 1994; Kozubal et al., 2011) in *M. yellowstonensis* [an aerobic Fe(II)-oxidizing Sulfolobales] and the NAG1 population (Candidate phylum Geoarchaeota) present in OSP_8, was in contrast to the notable absence of these respiratory proteins in the Desulfurococcales and Thermoproteales populations that dominate hypoxic, mildly acidic sulfur sediments (**Figure A2** in Appendix). Comparison of "Electron Transport" TIGRFAMs emphasized differences in OSP_8 versus WS_18 compared to all other sites, due in large part to the oxic nature of Fe(III)-oxide microbial mats, and to the considerably higher abundance of bacterial pathways in WS_18 (Thermodesulfobacteria, *Sulfurihydrogenibium*, and higher G + C bacteria represent a significant proportion of the total sequences from WS_18; **Figure A1** in Appendix). Respiratory processes in these bacteria are considerably different than the dominant archaea present in other sites, which contained very few bacterial reads (less than 10% on average). Consequently, although comparison of relative TIGRFAM abundances across sites represents functional differences of individual phylotypes, a detailed functional analysis of these populations provides clarification regarding observed functional differences across sites.
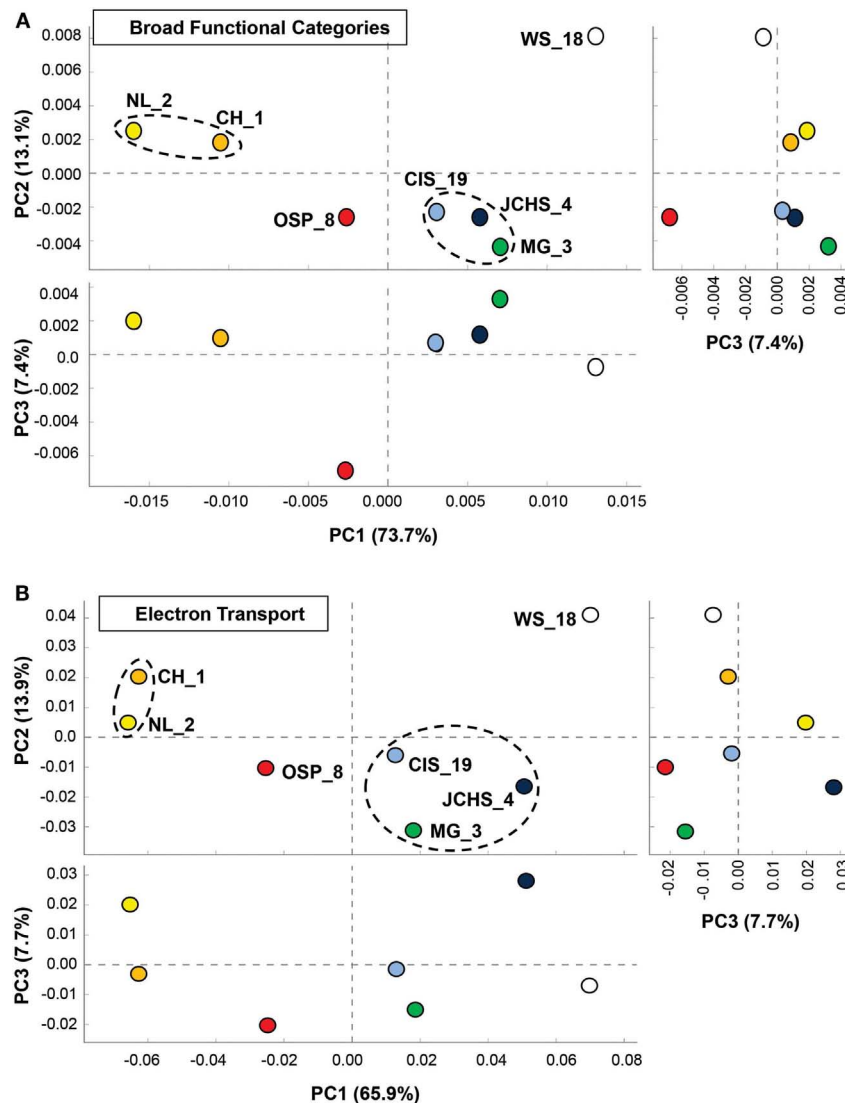
FIGURE 6 | Principal components analysis (PCA) of relative gene abundances across seven high-temperature archaeal-dominated chemotrophic communities. The three panels show pairwise plots of the first three principal components (PC1 and PC2 account for 76% of the variation across sites, while PC3 only represents ∼6%). **(A)** All TIGRFAMs grouped into functional categories, and **(B)**. Only those TIGRFAMs associated with the role category "Electron Transport." Sites are colored as before: *Crater Hills* (CH_1) = gold; *Nymph Lake* (NL_2) = yellow; *Monarch Geyser* (MG_3) = green; *Joseph's Coat Hot Springs* (JCHS_4) = dark-blue; *Cistern Spring* (CIS_19) = light-blue; *One Hundred Spring Plain* (OSP_8) = red; *Washburn Spring* (WS_18) = open.

FUNCTIONAL ANALYSIS OF PREDOMINANT ARCHAEAL PHYLOTYPES

The archaeal populations detected within these high-temperature sites exhibit extensive differences in energy conservation and CO_2 fixation pathways. To obtain more detailed information on specific functional genes present across sites, an extensive list of query genes coding for putative proteins important in CO_2 fixation pathways, electron transport and trace-element detoxification (As, Hg, superoxide) was compared to the assembled metagenomes (**Table 3**). All positive sequence hits were then compared to reference databases (using "blastp"), and analyzed individually (e.g., homology scores, phylogenetic trees) prior to confident assignment (**Table 3**). The detailed inventory of specific functional genes

is consistent with TIGRFAM protein family assignments, but is focused on specific pathways/proteins associated with geochemical processes.

Oxidation-reduction

The primary electron donors and acceptors that support metabolism of these archaeal phylotypes establish a critical link to geochemical processes. The oxidation of H_2 is highly exergonic, and this is no exception in thermal habitats where concentrations of dissolved $H_2(aq)$ are considerable (**Table 1**) and represent an available energy source (Amend and Shock, 2001; Inskeep et al., 2005; Spear et al., 2005). However, evidence for Group I
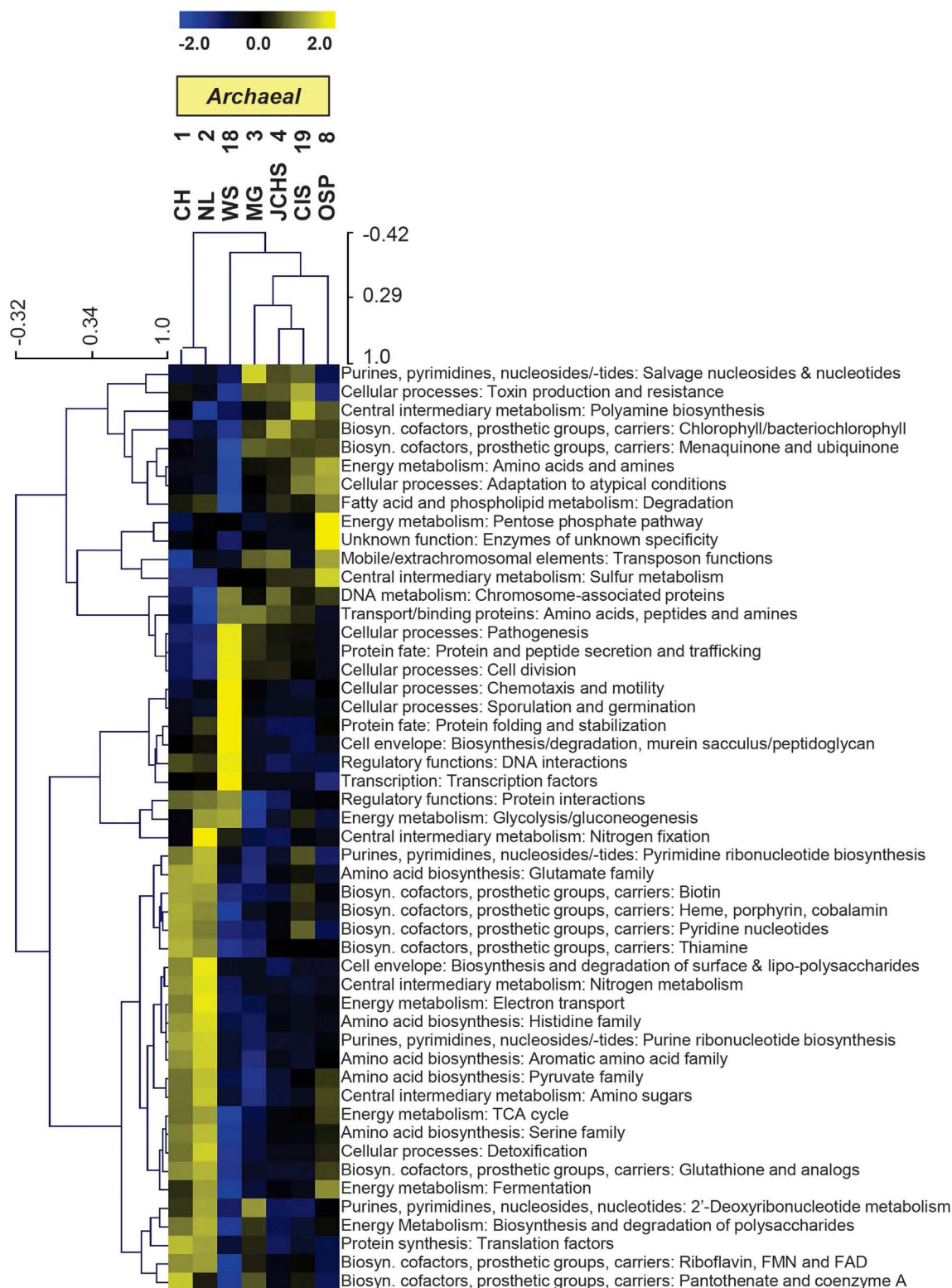
FIGURE 7 | Hierarchical cluster analysis of relative gene abundances across seven archaeal-dominated communities using all TIGRFAMs grouped into functional categories. Broad TIGRFAM categories include all cellular processes such as regulatory functions, energy metabolism, central C metabolism, mobile elements, transcription, cofactors, and transporters. Data was standardized by functional category before clustering to avoid biasing analysis by a few categories with high gene abundance. Pearson correlation was used as the distance measure for average linkage agglomerative clustering. TIGRFAMs from WS_18 and OSP_8 form separate functional clades consistent with the phylogenetic uniqueness of these sites.

Table 3 | Summary of key metabolic genes identified in sequence assemblies of the predominant archaeal populations present across sites, which exhibited a wide range in pH, dissolved sulfide, and dissolved oxygen (see Table 1).

Process/Pathway	Marker Genes[1]	Crater Hills (CH)	Nymph Lake (NL)	100 Spr. Plain (OSP)	Monarch Gysr. (MG)	Cistern Spr. (CIS)	J. Coat Hot Spr. (JCHS)	Washburn Spr. (WS)
CO2 fixation								
4-hydroxybutyryl-CoA dehydratase	Type1 4hcd	S1, S2	S1	My	A[2], S	A[2], S1, T2	T2	
Acetyl-CoA/Propionyl-CoA carboxylase	accA	S1, S2	S1	My		S1, S2	S1	
Electron Donors-Acceptors								
Hydrogen Oxidation (Group 1 Ni-Fe)	hynSL	S1, S2	S1		S	S	S, T2	T2, T3
Thiosulfate Oxidation	tqoA,B	S1, S2	S1	My, S	S	S	S	
Sulfur Oxidation	sor	S2						
Elemental S Oxidation	hdrABC	S1, S2	S1	My, S	S	S	S	K
Sulfide Oxidation	sqr	S1, S2	S1	My, S		S	S	Sfy
Reduced S Oxidation (S-transferase)	rdh	S1, S2	S1	NAG1, My, T1, Thaum	T1	T1	T1, S	Sfy
Sulfite Oxidation (Mo-pterin)	soxBC	S1, S2	S1	A, T1, T2, S	A, T1, T2, S	A, T1, T2, S	A, T1, T2, S	Sfy, T2
Oxygen Reduction, Heme Cu Oxidases[3]	e.g. doxB, foxA	S1 doxB S2 doxB	S1 doxB	NAG1-soxM, My-foxA/doxB, S-soxB/soxM			S-doxB, S-soxB	Sfy-coxA
Oxygen Reduction, cytochrome bd-ubiquinol type	cytAA'	S2	S1	NAG1, A, T1	A, T1	A, T1, T2	A, T1, T2	Sfy, K, T2, T3
Dissimilatory Reduction of Sulfur[4] (DMSO Mo-pterin family)	psrA/sreA[4]	S1, S2	S1	My		T2[4]	T2[4]	Sfy[4],T2[4]
Dissimilatory Reduction of Sulfur (NADPH Elemental S reductase)	nsr	S1, S2	S1	My, A, T1	A, T1	A, T1, S2	A, T1, T2	
Dissimilatory Reduction of Sulfate	dsrAB			T1, T2	T1, T2	T1, T2	T1, T2	T2, Tds
Reduction of Nitric oxide	norB			T1		T1, T2	T1, T2	T2
Detoxification-Stress Response	arsB	S1, S2	S1	NAG1, My, A, T1	A, T2	A, T1, T2, S	A, T1, T2	Sfy, K, T2, T3
	arsC			NAG1				Shy, Dt
	merA	S2	S1	NAG1, S		S	S	T2
	sodA	S1, S2	S1	NAG1, My, A, T1, S	A, T1, S	A, T1, T2, S	A, T1, T2, S	Sfy,T2, Mth, Tds

Population types: S1 and S2, Sulfolobales Types 1 and 2, respectively; My, M. yellowstonensis; A, Desulfurococcales (Acidilobus sp.); T1, T2, T3, Thermoproteales. Types 1, 2, and 3, respectively; NAG1, "Novel Archaea Group I" (Candidate phylum Geoarchaeota; Kozubal et al., 2013); K, Korarchaeota; Sfy, Sulfurihydrogenibium-like; Dt, Dictyoglomus-like; Mth, Methanosarcinales; Tds, Thermodesulfobacteria.

[1] "High confidence" sequence matches to marker genes that code for proteins with high specificity for possible pathway. Query sequences used to search for specific functional genes given in Table S3 in Supplementary Material (Inskeep et al., 2013). Genes not present in these sites = amoA, nifH, mcrA, nirK, nirS, nosZ.

[2] Specific heme Cu oxidases (subunit 1) are listed for each population type.

[3] Uncharacterized DMSO-Mo-pterins belonging to the Thermoproteales (T2) and Sulfurihydrogenibium (Sfy) populations are listed here.

Ni–Fe hydrogenases (responsible for H_2 uptake and oxidation to H^+, Viginais and Billoud, 2007) across these sites is limited to Sulfolobales populations, with the exception of one full-length (large and small subunit) Ni–Fe hydrogenase (on same contig) belonging to one of the Thermoproteales populations (Pyrobaculum-like) from JCHS_4 (**Table 3**). Complete Ni–Fe hydrogenases were present in the Sulfolobales populations at CH_1 and NL_2, but were absent in the M. yellowstonensis-like population present in OSP_8 (or OSP_14), as well as all other predominant archaeal populations present in these sites. Moreover, no evidence was found for key marker genes associated with methanogenesis (mcrA) (Ferry, 1999; Dhillon et al., 2005) or the oxidation of methane (pmoA), arsenite (aroA/asoA/aioA) (Hamamura et al., 2009), or ammonium (amoA), despite the fact that (i) substrates for these enzymes were present in high concentrations, (ii) oxidation reactions with these potential donors are highly exergonic in these thermal environments (Amend and Shock, 2001; Inskeep et al., 2005), and (iii) other cultivated thaumarchaea have been shown to oxidize ammonia (e.g., Walker et al., 2010).

The oxidation of reduced sulfur species (i.e., sulfide, S, thiosulfate, sulfite) is also highly exergonic (Amend and Shock, 2001; Ghosh and Dam, 2009). Genes (or gene complexes) known to code for proteins that catalyze the oxidation of reduced forms of S were found in several of the predominant archaeal populations present in these chemotrophic sites, especially members of the Sulfolobales (**Table 3**). All of the Sulfolobales populations (Types 1, 2, and M. yellowstonensis-like) contained gene complexes that are highly syntenous to and homologous to the heterodisulfide reductase (HDR) gene complex (**Figure A3** in Appendix) found in other Sulfolobales and bacterial genomes (Auernik and Kelly, 2008; Quatrini et al., 2009). The HDR complex is comprised of several heterodisulfide proteins and accessory components, which have been proposed to oxidize elemental sulfur to sulfite, followed by electron transfer to the quinol pool and ultimately to either a bd-ubiquinol oxygen reductase or a terminal oxidase

complex (i.e., heme Cu oxidase). Genes coding for rhodanese domain proteins are also linked with the HDR complex, consistent with their putative role as sulfuryl-transferases (Hedderich et al., 2005; Quatrini et al., 2009). These populations also contain sulfide:quinone reductase genes (*sqr*) shown to code for proteins involved in the oxidation of H_2S, HS^-, and S^{2-} to S^0 or polysulfide chains, followed by electron transfer to the quinone pool through a flavin adenine dinucleotide (FAD) cofactor (Cherney et al., 2010).

Moreover, genes known to code for proteins important in the oxidation of thiosulfate (TqoAB) were present in each of the major Sulfolobales types observed across these habitats (**Table 3**). A separate sulfur:oxygen reductase (*sor*) gene was also found in the Sulfolobales Type 2 population (*Stygiolobus*-like), and is the only archaeal population among these sites that appears to contain this gene, which has been implicated in the oxidation of elemental S in *Acidianus* spp. (Kletzin et al., 2004; Li et al., 2008). Nearly all archaea in these sites contained a sulfite oxidase molybdopterin, responsible for the oxidation of sulfite to sulfate (**Table 3**). However, the potential for oxidation of reduced S species (i.e., sulfide, elemental S, thiosulfate) was prevalent in the Sulfolobales populations, and these same genes were notably absent in members of the Desulfurococcales, Thermoproteales, and other archaeal populations present in these sites.

Respiratory pathways

The assembled metagenomes were also searched intensively for genes associated with aerobic and anaerobic respiration (**Table 3**). The presence of heme Cu oxidases (HCO; subunit I of terminal oxidase complexes) is an excellent indication of the potential to respire on O_2 (García-Horsman et al., 1994). The majority of heme Cu oxidases were associated with the three Sulfolobales populations (S1, S2, and My; **Table 3**) as well as the Geoarchaeota population (NAG1) from OSP_8 (Kozubal et al., 2013). The Thermoproteales (T1, T2), Desulfurococcales (A), and Korarchaeota (K) present across these sites showed no evidence of HCO-type oxygen reductases; however, these assemblies contained either *cyt*AA' or *cyt*AB type *bd*-ubiquinol oxidases (**Table 3**) that may represent high-affinity oxygen reductases important under hypoxic conditions (as observed in MG_3, CIS_19, JCHS_4, and WS_18; **Table 1**), or as O_2 scavenging proteins (García-Horsman et al., 1994; Das et al., 2005).

The predominant Thermoproteales populations present in sites MG_3, CIS_19, JCHS_4, WS_18, and OSP_8 (i.e., Type I *Caldivirga/Vulcanisaeta* and Type 2 *Pyrobaculum/Thermoproteus*-like organisms) were the only archaea in this study to contain putative dissimilatory sulfite/sulfate reductases (DsrAB). Moreover, these populations contain the only *nor*B (nitric oxide reductase) genes found among these sites. No evidence of *nar*G, *nir*K, or *nir*S homologs were found, so it is unclear how these organisms might reduce nitrate to nitric oxide, which would be required prior to reduction of NO to N_2O if using nitrate or nitrite as an electron acceptor (González et al., 2006). The NorB heme Cu oxidases could also play a role in reducing O_2 (Flock et al., 2005), detoxifying NO (Watmough et al., 2009), or possible dismutation of NO (Ettwig et al., 2012).

Other possible electron acceptors used by these archaeal populations include elemental sulfur (S^0) or polysulfides (**Table 3**). DMSO-molybdopterin sulfur reductase genes (*sre*A; Laska et al., 2003; Schut et al., 2007) were found in the Sulfolobales populations (S1, S2, My); however, the only other DMSO-molybdopterin genes found in these assemblies included formate dehydrogenases (*fdh*) and putative phenylacetyl CoA:acceptor oxidoreductases, which have been shown to be important in the oxidation of phenylacetic acids without using molecular oxygen (Rhee and Fuchs, 1999). These and several other novel DMSO-molybdopterin genes were consistently observed in the Thermoproteales T2 populations in MG_3, JCHS_4, CIS_19, and WS_18. Deduced proteins coded by these novel DMSO Mo-pterin genes do not cluster with currently known ArrA or SreA/PsrA proteins; consequently, it is not clear whether they play an important role in energy conservation for the Thermoproteales populations (Jay et al., 2011). The presence of a NAD(P)H elemental sulfur reductase (*nsr*) similar to that described in *Pyrococcus furiosus* (Blumentals et al., 1990; Schut et al., 2007) may represent an additional pathway for reduction of elemental sulfur (or polysulfides) in archaea, and copies of this gene were present in all of the major archaeal populations in these sites, with the exception of the Geoarchaeota (NAG1) population from OSP_8 (**Table 3**).

Carbon dioxide fixation

Evidence for CO_2 fixation in the archaea has focused primarily on the recently discovered 3-hydroxypropionate/4-hydroxybutyrate (3HP–4HB) pathway, originally reported in *M. sedula* (Berg et al., 2007). The marker genes for this complex, 16-step pathway are (i) 4-hydroxybutyryl-CoA dehydratase (*4hbd*), which catalyzes the conversion of 4-hydroxybutyryl-CoA to crotonyl-CoA, and (ii) the bifunctional acetyl-CoA/propionyl-CoA carboxylase (AccB/AccC/PccB) (Hügler et al., 2003; Berg et al., 2007, 2010a,b). Excellent matches to these genes were found distributed throughout the Sulfolobales populations (Types S1, S2, and My; **Table 3**), and it is likely that these organisms are capable of fixing CO_2. However, no *accA*-like genes were found in other archaea besides the Sulfolobales. Copies of a *4hbd* gene were observed in the *Acidilobus*-like (A) and *Pyrobaculum*-like (T2) populations in several sites (**Table 3**); however, despite the presence of other common metabolic genes in the "dicarboxylate" branch of the "dicarboxylate/4*hbd*" pathway, it is not clear from the metagenome assemblies that these organisms have the necessary genes for a complete CO_2 fixation pathway, in part due to the lack of the required enzyme, phosphoenolpyruvate (PEP) carboxylase, which appears to be missing in these populations. The pathways responsible for CO_2 fixation in these phylogenetic groups are still the subject of considerable research (Huber et al., 2006, 2008; Jahn et al., 2007; Ramos-Vera et al., 2009, 2011; Berg et al., 2010a,b), and numerous members of the Desulfurococcales and Thermoproteales cannot grow solely on CO_2 and require complex sources of C for growth (Huber et al., 2006; Boyd et al., 2007; Macur et al., 2013). No Type 1 *4hbd* genes were present in the NAG1 population (Geoarchaeota) from OSP_8, or the *Caldivirga/Vulcanisaeta*-like (T1) organisms in any of these sites (**Table 3**). Consequently, if these organisms

are capable of fixing CO_2, it is likely occurring via a different mechanism.

DISCUSSION

Prior to the current study, the structure and function of hyperthermophilic microbial communities in YNP has been inferred primarily from results of PCR using universal bacterial or archaeal primers, and inferred physiology from cultured relatives. Moreover, it has been difficult to gain any definitive information regarding the relative importance of bacteria versus archaea in high-temperature systems of YNP. The sites described here were chosen to represent several common chemotrophic community types in YNP, and included elemental sulfur systems ranging in pH from 2.5 to 6.5, as well as acidic Fe-oxide mats. Random shotgun sequencing (i.e., metagenomics) showed that these sites were dominated by archaeal populations, with the exception of *Washburn Springs*. Although other sites contained evidence of subdominant bacterial populations, sediment from WS_18 (pH = 6.4, T = 80°C) contained a significant number of sequences corresponding to *Sulfurihydrogenibium* (Aquificales), Thermodesulfobacteria and *Dictyoglomus*-like organisms. Given the circumneutral pH and the high-sulfide concentrations at WS_18, one would expect *Sulfurihydrogenibium* rather than *Thermocrinis* (higher pH) or *Hydrogenobaculum* (lower-pH) (Takacs-Vesbach et al., 2013).

The distribution of different archaeal populations as a function of environmental factors was consistent with the major habitat types identified in decision-tree format (Inskeep et al., 2013). Briefly, archaeal-dominated sites were separated based primarily on pH and the presence of dissolved sulfide and/or elemental sulfur. Temperature was not a major variable in this study since these sites were all between 72 and 85°C, and five of the seven sites were between 78 and 82°C. The abiotic consumption of oxygen by reduced sulfur species contributes to the hypoxic conditions observed in CH_1, NL_2, MG_3, JCHS_4, CIS_19, and WS_18. Consequently, six sulfidic sites were analyzed ranging in pH from 2.5 to 6.5. Our results showed that pH was a major factor controlling the distribution of Sulfolobales versus Thermoproteales and Desulfurococcales, as well as other novel archaeal groups found under limited conditions (e.g., Korarchaeota in WS_18). A combination of low pH, reduced sulfur and high-temperature severely constrained microbial community diversity and two sites with these properties were dominated by only two major Sulfolobales populations. However, the sulfidic (hypoxic) sediments at pH 6.4 (WS_18) were more diverse, and contained a significant number of *Sulfurihydrogenibium* (~15%) and Thermodesulfobacteria-like (~10%) sequence reads. Also, the presence of several Thermoproteales populations in WS_18 was consistent with the increased abundance of these phylotypes with increasing pH [e.g., CIS_19 (pH 4.8) and JCHS_4 (pH 6.1)]. *Washburn Springs* (WS_18) was the only habitat (out of 20 reported in the entire study, Inskeep et al., 2013) to contain a significant korarchaeotal population, and is consistent with recent studies on the distribution of korarchaeotal sequences in Kamchatka and YNP, which showed that these organisms have a limited pH range from ~5 to 7 (Auchtung et al., 2011; Miller-Coleman et al., 2012). Although pH undoubtedly plays an important role in establishing

the realized niche of thermophilic Korarchaeota (Miller-Coleman et al., 2012), sequence data suggest that this population may contain *hdrABC* gene complexes associated with sulfur oxidation via a *bd*-ubiquinol oxidase (**Table 3**). This population appears to have no known CO_2 fixation pathways and may benefit from considerable levels of dissolved organic carbon (~500 μM DOC) present in WS_18 (Table S2 in Supplementary Material, Inskeep et al., 2013).

Hydrodynamic context is also a critical modifying factor that influences the rate of equilibration with atmospheric O_2, and is especially evident within the primary outflow channels of geothermal springs. Moderately acidic habitats (pH ~ 3–3.5) containing Fe(II) and oxygen (i.e., OSP_8 and 14) showed an increase in archaeal diversity relative to the lower-pH habitats containing reduced sulfur. The Fe-oxide mat (OSP_8) contained three to four major lineages within the Crenarchaeota [e.g., Fe-oxidizing Sulfolobales (*M. yellowstonensis*), *Acidilobus*-like, *Vulcanisaeta*-like], and several undescribed lineages within the Thaumarchaeota, Euryarchaeota, and newly proposed Geoarchaeota (Kozubal et al., 2013). The different types of archaea and the corresponding diversity of heme Cu oxidases found in Fe mats (e.g., OSP_8, **Table 3**) is consistent with the fact that these are the most oxic environments included in the study. The Fe-mat also contained members of the Aquificales (*Hydrogenobaculum*-like), but these bacteria were more pronounced in filamentous "streamer" communities (site OSP_14; Takacs-Vesbach et al., 2013).

Archaea are adapted to numerous extreme environments and their respective functional attributes are equally diverse. The TIGRFAMs identified in the current study significantly expand the diversity of proteins reported from metagenome sequence currently in public databases. This is due primarily to the abundant and diverse archaea distributed across these sites and the fact that few metagenomes from high-temperature systems have been reported. The presence of different functional genes among high-temperature chemotrophic communities is defined by the distribution of predominant archaeal phylotypes and provides a foundation for understanding metabolic linkages to environmental constituents such as O_2, S, and Fe, as well as the evolutionary history of these phyla. Likewise, the lack of genes known to code for the oxidation of ammonium (*amoA*) and/or methane (*pmoA*) suggests that these reactions, although exergonic, do not support the metabolism of dominant populations in these sites. The archaeal sites sampled here do not contain significant numbers of methanogens with the exception of WS_18, where the higher pH (6.4), dissolved CO_2 and dissolved H_2 (**Table 1**) appear to support subdominant populations (<1% of total sequences) related to Methanococci and/or Methanosarcinales. Other poorly characterized archaea identified across these high-temperature systems included members of the Nanoarchaeota, Euryarchaeota, and novel phylum-level lineages (e.g., Geoarchaeota from OSP_8). Moreover, the presence of viral sequence in the community metagenomes as well as the identification of unique CRISPR regions in numerous archaeal phylotypes provides genomic evidence that new viruses have yet to be identified and characterized in these habitats. Although there are numerous high-temperature habitats yet to be studied, the sites included here provide an excellent foundation

for understanding both phylogenetic and functional variation within the archaea as a function of major geochemical parameters including pH, reduced sulfur, dissolved oxygen, and ferrous Fe.

MATERIALS AND METHODS

SITE SELECTION, SAMPLE COLLECTION, AND PROCESSING

Seven high-temperature sediment and/or mat samples rich in archaea (**Figure 1**) were sampled from geothermal environments in 2007–2008. The sites were chosen to obtain a range in pH across hypoxic sulfur sediments (2.5–6.4), as well as to contrast reduced sulfur environments with oxic flow channels containing Fe(III)-oxides. The research sites chosen for study have been the subject of significant prior characterization and include: *Crater Hills* (CH_1, *Alice Springs*), *Nymph Lake* (NL_2), *Monarch Geyser* (MG_3), *Cistern Spring* (CIS_19), *Joseph's Coat Hot Spring* (JCHS_4, also known as "JC3" and *Scorodite Spring*), *Washburn Springs* (WS_18), and *One Hundred Springs Plain* (OSP_8). Each microbial community and associated solid phase was sampled aseptically, stored in 50 mL sterile "falcon" tubes on dry ice, and transported to a −80°C freezer (MSU) until DNA extraction.

Parallel samples of the bulk aqueous phase (<0.2 μm) associated with each microbial community were obtained simultaneously and analyzed using a combination of field and laboratory methods. As described in more detail in other reports (Inskeep et al., 2004; Macur et al., 2004), pH, temperature, and other redox sensitive species [Fe^{II}/Fe^{III}; As^{III}/As^{V}; total dissolved sulfide (DS); dissolved O_2 (DO)] were determined using field methods. Major cations and other trace elements (i.e., Na, K, Ca, Mg, Fe, Al, Mn, Cu, Sr, Ba, Li, Zn, Cu, Pb, Si, B, P, As, Sb, S, Se) were determined using inductively coupled plasma (ICP) spectrometry, and major anions (F, Cl, SO_4, S_2O_3, AsO_4, PO_4, NO_3) were determined using ion chromatography (Dionex, Sunnyvale, CA, USA). Ammonium concentrations were determined using colorimetry (autoanalyzer). Dissolved gases (CO_2, H_2, and CH_4) were determined using closed head-space gas chromatography (Inskeep et al., 2005) of sealed serum-bottle samples obtained in the field. The majority of these sites have been sampled many times with excellent replication (Langner et al., 2001; Rice et al., 2001; Inskeep et al., 2005, 2010; Young et al., 2005; Kozubal et al., 2012, 2013). The location and primary physicochemical characteristics obtained during sampling are provided here (**Table 1**); additional geochemical data are included in supplemental information (see Table S2 in Supplemental Material, Inskeep et al., 2013).

DNA EXTRACTION AND LIBRARY CONSTRUCTION

Although a standard DNA extraction protocol (Inskeep et al., 2013) was attempted for all samples, several of the archaeal-dominated sediment samples required extraction kits (MoBio) to obtain sufficient DNA for analysis. Our main emphasis was to obtain representative, unbiased community DNA as template for construction of small insert libraries. Small insert (puc13) libraries were constructed and transformed, then sequenced (Sanger ∼800 bp reads) to generate ∼35–50 Mb per site, for a total sequence for these sites of ∼250 Mb. Sites NL_2, JCHS_4, and

MG_3 also received a half-plate of 454 pyrosequencing. For consistency, this manuscript focuses on the Sanger data across each of the seven sites; pyrosequence data did not result in assemblies containing significantly greater non-redundant sequence, although it did improve the coverage of dominant community members.

ANALYSIS OF INDIVIDUAL SEQUENCE READS

Analysis of individual sequence reads using MEGAN assignments from "blastx" results and G + C content distribution provided a quick and useful phylogenetic summary of predominant community members in each site (**Figure 2**), and indicated where to expect major assemblies. Whole genome-level comparative analysis was accomplished using fragment (read) recruitment of environmental sequence data to reference microbial genomes (Rusch et al., 2007). At the time of writing, the database contained microbial genomes for ∼1,500 bacteria and 100 archaea. Currently, only a handful of microbial genomes served as appropriate references for the indigenous organisms within these chemotrophic communities, consequently, many of the assignments were given at family or domain level.

ANALYSIS OF PREDOMINANT SEQUENCE ASSEMBLIES

Random shotgun DNA sequence (∼35–50 Mbp Sanger per site) was obtained and assembled as described in this issue (Inskeep et al., 2013). Assembled metagenome sequence data was analyzed using PCA of NWF (Teeling et al., 2004; Inskeep et al., 2010) calculated for all contigs/scaffolds greater than 3–4 kb. The sequence clusters were also viewed with a simultaneous phylogenetic classification based on the (APIS; Badger et al., 2006), or a "blast"-based classification (Rusch et al., 2007). Briefly, APIS is a system for automatic creation and summarizing of phylogenetic trees for each protein encoded by a genome or metagenomic dataset.

CRISPR ANALYSIS

CRISPRs were identified using the CRISPRFinder software (Grissa et al., 2007). DR were counted and individual repeats with less than 10 instances in the assembly were excluded from further analysis. The CRISPR spacer and DR were searched (NCBI BLASTN with −e 100 −U T −F "m L" −b 10000; Altschul et al., 1990) against the scaffolds and HSPs with up to three or fewer mismatches were identified. Scaffolds identified by the CRISPR spacers were searched against NRAA (BLASTX −e 1e-4 −U T −F "m L") to determine their similarity to known viruses. Spacer identified scaffolds were also searched against themselves (BLASTN −U T −F "m L" −X 150 −q −5 −r 4).

TIGRFAM PROTEIN FAMILY ABUNDANCE IN ASSEMBLED METAGENOME SEQUENCE

Assembled sequence from each of the archaeal sites was annotated as described in Inskeep et al. (2010) and predicted proteins from the scaffolds were assigned TIGRFAM protein families (Selengut et al., 2007) using HMMER 3 (Eddy, 2011) with E-value cut-off of 1e-6. PCA and statistical analysis of site group differences was performed using the STAMP v2.0 software (Parks and Beiko, 2010). The White's non-parametric t-test and ANOVA tests were used to test for differences between two site groups and multiple site groups respectively. Two-way clustering was done using

row-standardized (across sites) average TIGRFAM category abundance data using the Euclidean distance metric and complete-linkage hierarchical clustering in MeV 4.8 (Saeed et al., 2003) software. Other details are as described above (Inskeep et al., 2013).

FUNCTIONAL ANALYSIS OF ARCHAEA IN YNP

The assembled sequence data was screened for specific functional genes corresponding to known and/or putative pathways in carbon metabolism and electron (energy) transfer. We were specifically interested in assessing metabolic potential for chemolithoautotrophy (CO_2 fixation and electron transfer) in high-temperature geothermal systems. Query DNA sequences known to code for proteins important in the oxidation of reduced chemical constituents or the reduction of a terminal acceptor were used to search the environmental sequence data. Environmental sequence fragments exhibiting homology (E-values $<10^{-10}$) to query sequences (Table S3 in Supplementary Material, Inskeep et al., 2013) were then reanalyzed using "blastn," and carefully assessed individually using phylogenetic analysis of deduced protein sequences against known relatives, as well as fragment length relative to query length. False positives were eliminated by this screening process and included (i) sequences matching the correct protein family, but not the exact query sequence (e.g., Mo-pterin oxidoreductases versus a specific protein within this family), (ii) sequences that matched a query gene due to homologous regions, but were clearly associated with a gene cluster of different function, and (iii) sequences that returned mis-annotated "blastn" relatives. It is also possible that our inventory of metabolic potential has missed sequences related to a specific query gene. For example, some homologous genes found in the metagenome data were of insufficient length relative to known query sequences to make a definitive assignment. Clearly, the metagenomes obtained here do not represent complete sequence for all subdominant populations in these sites.

SEQUENCE AVAILABILITY

All annotated metagenome sequence assemblies (Celera/PGA) discussed in the current manuscript are available through the DOE-JGI IMG/M (Markowitz et al., 2012) website (http://img.jgi.doe.gov/m) under IMG taxon OID numbers as follows: YNPSite01 (2022920009/2014031002), Site02 (2022920014/2015219001,

2016842002), Site03 (2022920002/2014031003, 20162001), Site19 (2022920017/2015219000), Site04 (2022920008/2013843003), Site18 (2022920019/2016842004), Site08 (2022920005/2013515001) and Site14 (2022920007/2013954001). Scaffold ID numbers are preserved in the annotated Celera sequence files, and serve as an appropriate mechanism of referencing assembled sequence data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Microbial_Physiology_and_Metabolism/10.3389/fmicb.2013.00095/abstract

Table S1 | Summary of 16S rRNA gene sequences observed in assembled metagenome sequence data from high-temperature, archaeal-dominated chemotrophic sites in Yellowstone National Park (also, see Figure 4 for phylogenetic tree).

Table S2 | Description of predominant sequence assemblies in archaeal-dominated sites, including largest scaffold (kbp), number of scaffolds in cluster, total consensus sequence (Mbp), average G + C content (%), and closest cultured relative of the 16S rRNA gene found within the assembled data.

Table S3 | Distribution of direct repeats (DR) in archaeal-dominated sites.

Table S4 | Survey of single-copy genes corresponding to the predominant archaeal populations present in high-temperature geothermal microbial communities of YNP.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.

Amend, J. P., and Shock, E. L. (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiol. Rev.* 25, 175–243.

Auchtung, T. A., Shyndriayeva, G., and Cavanaugh, C. M. (2011). 16S rRNA phylogenetic analysis and quantification of Korarchaeota indigenous to the hot springs of Kamchatka, Russia. *Extremophiles* 15, 105–116.

Auernik, K. S., and Kelly, R. M. (2008). Identification of components of electron transport chains in the extremely thermoacidophilic crenarchaeon *Metallosphaera sedula* through iron and sulfur compound oxidation transcriptomes. *Appl. Environ. Microbiol.* 74, 7723–7732.

Auguet, J. C., Barberan, A., and Casamayor, E. O. (2009). Global ecological patterns in uncultured Archaea. *ISME J.* 4, 182–190.

Badger, J. H., Hoover, T. R., Brun, Y. V., Weiner, R. M., Laub, M. T., Alexandre, G., et al. (2006). Comparative genomic evidence for a close

relationship betwin the dimorphic prosthecate bacteria *Hyphomonas neptunium* and *Caulobacter crescentus*. *J. Bacteriol.* 188, 6841–6850.

Ball, J. W., Nordstrom, D. K., Cunningham, K. M., Schoonen, M. A. A., Xu, Y., and DeMonge, J. M. (1998). *Water-Chemistry and Onsite Sulfur Speciation Data for Selected Springs in Yellowstone National Park, Wyoming, 1994–1995*. U.S. Geological Survey Open-File Report 98-574, Boulder, CO.

Ball, J. W., McCleskey, R. B., Nordstrom, D. K., Holloway, J. M.,

and Verplanck, P. L. (2002). *Water-Chemistry Data for Selected Springs, Geysers, and Streams in Yellowstone National Park, Wyoming 1999–2000*. U.S. Geological Survey Open File Report 02-382, Boulder, CO.

Barnes, S. M., Fundyga, R. E., Jeffries, M. W., and Pace, N. R. (1994). Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc. Natl. Acad. Sci. U.S.A.* 91, 1609–1613.

Bathe, S., and Norris, P. R. (2007). Ferrous iron- and sulfur-induced genes in *Sulfolobus metallicus*. *Appl. Environ. Microbiol.* 73, 2491–2497.

Beam, J. P., Jay, Z. J., Kozubal, M. A., and Inskeep, W.P. (2011). "Distribution and ecology of thaumarchaea in geothermal environments," in *The 11th International Conference on Thermophiles Research*, September 11–16 (Big Sky: Thermophiles). [Abstr. PA-009].

Berg, I. A., Kockelkorn, D., Buckel, W., and Fuchs, G. (2007). A 3-hydroxy-propionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* 318, 1782–1786.

Berg, I. A., Ramos-Vera, W. H., Petri, A., Huber, H., and Fuchs, G. (2010a). Study of the distribution of autotrophic CO_2 fixation cycles in Crenarchaeota. *Microbiology* 156, 256–269.

Berg, I. A., Kockelkorn, D., Ramos-Vera, W. H., Say, R. F., Zarzycki, J., Hügler, M., et al. (2010b). Autotrophic carbon fixation in archaea. *Nat. Rev. Microbiol.* 8, 447–460.

Blumentals, I. I., Itoh, M., Olson, G. J., and Kelly, R. M. (1990). Role of polysulfides in reduction of elemental sulfur by the hyperthermophilic archaebacterium *Pyrococcus furiosus*. *Appl. Environ. Microbiol.* 56, 1255–1262.

Boyd, E. S., Jackson, R. A., Encarnacion, G., Zahn, J. A., Beard, T., Leavitt, W. D., et al. (2007). Isolation, characterization, and ecology of sulfur-respiring Crenarchaea inhabiting acid-sulfate-chloride geothermal springs in Yellowstone National Park. *Appl. Environ. Microbiol.* 73, 6669–6677.

Brierley, C. L., and Brierley, J. A. (1973). A chemoautotrophic and thermophilic microorganism isolated from an acid hot spring. *Can. J. Microbiol.* 19, 183–188.

Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008). Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245–252.

Brock, T. D., Brock, K. M., Belly, R. T., and Weiss, R. L. (1972). Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch. Mikrobiol.* 84, 54–68.

Casanueva, A., Galada, N., Baker, G. C., Grant, W. D., Heaphy, S., Jones, B., et al. (2008). Nanoarchaeal 16S rRNA gene sequences are widely dispersed in hyperthermophilic and mesophilic halophilic environments. *Extremophiles* 12, 651–656.

Chaban, B., Ng, S. Y., and Jarrell, K. F. (2006). Archaeal habitats – from the extreme to the ordinary. *Can. J. Microbiol.* 52, 73–116.

Cherney, M. M., Zhang, Y., Solomonsona, M., Weinera, J. H., and James, N. G. (2010). Crystal Structure of Sulfide:quinone oxidoreductase from *Acidithiobacillus ferrooxidans*: insights into sulfidotrophic respiration and detoxification. *J. Mol. Biol.* 398, 292–305.

Clingenpeel, S., Kan, J., Macur, R. E., Lovalvo, D., Varley, J., Inskeep, W. P., et al. (2011). "Yellowstone lake reveals evidence of non-extreme Nanoarchaeota and ecotype diversity," in *The 11th International Conference on Thermophiles Research*, September 11–16 (Big Sky: Thermophiles). [Abstr. PA-079].

Das, A., Silaghi-Dumitrescu, R., Ljungdahl, L. G., and Kurtz, D. M. (2005). Cytochrome bd oxidase, oxidative stress, and dioxygen tolerance of the strictly anaerobic bacterium *Moorella thermoacetica*. *J. Bacteriol.* 187, 2020–2029.

de la Torre, J. R., Walker, C. B., Ingalls, A. E., Könneke, M., and Stahl, D. A. (2008). Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ. Microbiol.* 10, 810–818.

DeLong, E. F. (2005). Microbial community genomics in the ocean. *Nat. Rev. Microbiol.* 3, 459–469.

Dhillon, A., Lever, M., Lloyd, K. G., Albert, D. B., Sogin, M. L., and Teske, A. (2005). Methanogen diversity evidenced by molecular characterization of methyl coenzyme M reductase A (*mcrA*) genes in hydrothermal sediments of the Guaymas Basin. *Appl. Environ. Microbiol.* 71, 4592–4601.

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi:10.1371/journal.pcbi.1002195

Elkins, J. G., Kunin, V., Anderson, I., Barry, K., Goltsman, E., Podar, M., et al. (2008). A korarchaeal genome reveals new insights into the evolution of the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8102–8107.

Ettwig, K. F., Speth, D. R., Reimann, J., Wu, M. L., Jetten, M. S. M., and Keltjens, J. T. (2012). Bacterial oxygen production in the dark. *Front. Microbiol.* 3:273. doi:10.3389/fmicb.2012.00273

Falkowski, P. G., Fenchel, T., and Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science* 23, 1034–1039.

Ferry, J. G. (1999). Enzymology of one-carbon metabolism in methanogenic pathways. *FEMS Microbiol. Rev.* 23, 13–38.

Fitz-Gibbon, S. T., Ladner, H., Kim, U. J., Stetter, K. O., Simon, M. I., and Miller, J. H. (2002). Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*. *Proc. Natl. Acad. Sci. U.S.A.* 99, 984–989.

Flock, U., Watmough, N. J., and Ädelroth, P. (2005). Electron/proton coupling in bacterial nitric oxide reductase during reduction of oxygen. *Biochemistry* 44, 10711–10719.

Friedrich, C. G., Bardischewsky, F., Rother, D., Quentmeier, A., and Fischer, J. (2005). Prokaryotic sulfur oxidation. *Curr. Opin. Microbiol.* 8, 253–259.

García-Horsman, J. A., Barquera, B., Rumbley, J., Ma, J., and Gennis, R. B. (1994). The superfamily of heme-copper respiratory oxidases. *J. Bacteriol.* 176, 5587–5600.

Ghosh, W., and Dam, B. (2009). Biochemistry and molecular biology of lithotrophic sulfur oxidation by taxonomically and ecologically diverse bacteria and archaea. *FEMS Microbiol. Rev.* 33, 999–1043.

González, P. J., Correia, C., Moura, I., Brondino, C. D., and Moura, J. J. (2006). Bacterial nitrate reductases: molecular and biological aspects of nitrate reduction. *J. Inorg. Biochem.* 100, 1015–1023.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35(Suppl. 2), W52–W57.

Hamamura, N., Macur, R. E., Korf, S., Ackerman, G., Taylor, W. P., Kozubal, M. A., et al., (2009). Linking microbial oxidation of arsenic with detection and phylogenetic analysis of arsenite-oxidase genes in diverse geothermal environments. *Environ. Microbiol.* 11, 421–431.

Hatzenpichler, R., Lebedeva, E. V., Spieck, E., Stoecker, K., Richter, A., Daims, H., et al. (2008). A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2134–2139.

Hedderich, R., Hamann, N., and Bennati, M. (2005). Heterodisulfide reductase from methanogenic archaea: a new catalytic role for an iron-sulfur cluster. *Biol. Chem.* 386, 961–970.

Hohn, M. J., Hedlund, B. P., and Huber, H. (2002). Detection of 16S rDNA sequences representing the novel phylum "Nanoarchaeota": indication for a world-wide distribution in high temperature biotopes. *Syst. Appl. Microbiol.* 25, 551–554.

Huber, H., Gallenberger, M., Jahn, U., Eylert, E., Berg, I. A., Kockelkorn, D., et al. (2008). A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum *Ignicoccus hospitalis*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7851–7856.

Huber, H., Hohn, M. J., Rachel, R., Fuchs, T., Wimmer, V. C., and Stetter, K. O. (2002). A new phylum of Archaea represented by a nano-sized hyperthermophilic symbiont. *Nature* 417, 63–67.

Huber, H., Huber, R., and Stetter, K. O. (2006). "Thermoproteales," in *The Prokaryotes*, eds M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt (New York: Springer), 10–22.

Hügler, M., Krieger, R. S., Jahn, M., and Fuchs, G. (2003). Characterization of acetyl-CoA/propionyl-CoA carboxylase in *Metallosphaera sedula*: carboxylating enzyme in the 3-hydroxypropionate cycle for autotrophic carbon fixation. *Eur. J. Biochem.* 270, 736–744.

Inskeep, W. P., Ackerman, G. G., Taylor, W. P., Kozubal, M. A., Korf, S., and Macur, R. E. (2005). On the energetics of chemolithotrophy: case studies of geothermal springs in Yellowstone National Park. *Geobiology* 3, 297–317.

Inskeep, W. P., Jay, Z. J., Tringe, S. G., Herrgard, M., Rusch, D. B., et al. (2013). The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem. *Front. Microbiol.* 4:67. doi:10.3389/fmicb.2013.00067

Inskeep, W. P., Macur, R. E., Harrison, G., Bostick, B. C., and Fendorf, S. (2004). Biomineralization of As(V)-hydrous ferric oxyhydroxide mats in an acid-sulfate-chloride geothermal spring of Norris Geyser Basin, Yellowstone National Park. *Geochim. Cosmochim. Acta* 68, 3141–3155.

Inskeep, W. P., Rusch, D. B., Jay, Z. J., Herrgard, M. J., Kozubal, M. A., Richardson, T. H., et al. (2010). Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. *PLoS ONE* 5:e9773. doi:10.1371/journal.pone.0009773

Itoh, T., Suzuki, K., and Nakase, T. (2002). *Vulcanisaeta distributa* gen. nov., sp nov., and *Vulcanisaeta souniana* sp nov., novel hyperthermophilic rod-shaped crenarchaeotes isolated from hot springs in Japan. *Int. J. Syst. Evol. Microbiol.* 52, 1097–1104.

Itoh, T., Suzuki, K., Sanchez, P. C., and Nakase, T. (1999). *Caldivirga maquilingensis* gen. nov., sp nov., a new genus of rod-shaped crenarchaeote isolated from a hot spring in the Philippines. *Int. J. Syst. Bacteriol.* 49, 1157–1163.

Jahn, U., Huber, H., Eisenreich, W., Hugler, M., and Fuchs, L. (2007). Insights into the autotrophic CO_2 fixation pathway of the archaeon *Ignicoccus hospitalis*: comprehensive analysis of the central carbon metabolism. *J. Bacteriol.* 189, 4108–4119.

Jay, Z. J., Planer-Friedrich, B., Rusch, D. B., and Inskeep, W. P. (2011). "Linking geochemistry to microbial community structure and function in sulfidic geothermal systems of Yellowstone National Park," in *Goldschmidt Conference*, August 2011, Prague, CR. *Mineralogical Magazine* 75, 1104.

Kletzin, A., Urich, T., Müller, F., Bandeiras, T. M., and Gomes, C. M. (2004). Dissimilatory oxidation and reduction of elemental sulfur in thermophilic archaea. *J. Bioenerg. Biomembr.* 36, 77–91.

Kozubal, M., Macur, R. E., Korf, S., Taylor, W. P., Ackerman, G. G., Nagy, A., et al. (2008). Isolation and distribution of a novel iron-oxidizing crenarchaeon from acidic geothermal springs in Yellowstone National Park. *Appl. Environ. Microbiol.* 74, 942–949.

Kozubal, M. A., Dem Jennings, R., Jay, Z., Beam, J. P., Tringe, S. G., Rusch, D. B., et al. (2013). Geoarchaeota: a new Candidate phylum in the *Archaea* from high-temperature acidic iron mats in Yellowstone National Park. *ISME J.* 7, 622–634.

Kozubal, M. A., Dlakic, M., Macur, R. E., and Inskeep, W. P. (2011). Terminal oxidase diversity and function in "*Metallosphaera yellowstonensis*": gene expression and protein modeling suggest mechanisms of Fe(II) oxidation in the sulfolobales. *Appl. Environ. Microbiol.* 77, 1844–1853.

Kozubal, M. A., Macur, R. E., Jay, Z. J., Beam, J. P., Malfatti, S. A., Tringe, S. G., 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. *Front. Microbiol.* 3:109. doi:10.3389/fmicb.2012.00109

Langner, H. W., Jackson, C. R., McDermott, T. R., and Inskeep, W. P. (2001). Rapid oxidation of arsenite in a hot spring ecosystem, Yellowstone National Park. *Environ. Sci. Technol.* 35, 3302–3309.

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

Li, M., Chen, Z., Zhang, P., Pan, X., Jiang, C., An, X., et al. (2008). Crystal structure studies on sulfur oxygenase reductase from *Acidianus tengchongensis*. *Biochem. Biophys. Res. Commun.* 369, 919–923.

Macur, R. E., Jay, Z. J., Taylor, W. P., Kozubal, M. A., Kocar, B. D., and Inskeep, W. P. (2013). Microbial community structure and sulfur biogeochemistry in mildly-acidic sulfidic geothermal springs in Yellowstone National Park. *Geobiology* 11, 86–99.

Macur, R. E., Langner, H. W., Kocar, B. D., and Inskeep, W. P. (2004). Linking geochemical processes with microbial community analysis: successional dynamics in an arsenic-rich, acid-sulfate-chloride geothermal spring. *Geobiology* 2, 163–177.

Makarova, K. S., Aravind, L., Wolf, Y. I., and Koonin, E. V. (2011). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct* 6, 38.

Mardanov, A. V., Svetlitchnyi, V. A., Beletsky, A. V., Prokofeva, M. I., Bonch-Osmolovskaya, E. A., Ravin, N. V., et al. (2010). The genome sequence of the crenarchaeon *Acidilobus saccharovorans* supports a new order, Acidilobales, and suggests an important ecological role in terrestrial acidic hot springs. *Appl. Environ. Microbiol.* 76, 5652–5657.

Markowitz, V. M., Chen, I.-M. A., Chu, K., Szeto, E., Palaniappan, K., Grechkin, Y., et al. (2012). IMG/M: the integrated metagenome data management and comparative analysis system. *Nucleic Acids Res.* 40, D123–D129.

McCleskey, R. B., Ball, J. W., Nordstrom, D. K., Holloway, J. M., and Taylor, H. E. (2005). *Water-Chemistry Data for Selected Hot Springs, Geysers, and Streams in*

Yellowstone National Park, Wyoming 2001–2002. United States Geological Survey Open File Report 2004-1316, Boulder, CO.

Meyer-Dombard, D., Shock, E. L., and Amend, J. P. (2005). Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. *Geobiology* 3, 211–227.

Miller-Coleman, R. L., Dodsworth, J. A., Ross, C. A., Shock, E. L., Williams, A. J., Hartnett, H. E., et al. (2012). Korarchaeota diversity, biogeography, and abundance in Yellowstone and Great Basin Hot Springs and ecological niche modeling based on machine learning. *PLoS ONE* 7:e35964. doi:10.1371/journal.pone.0035964

Nordstrom, D. K., Ball, J. W., and McClesky, R. B. (2005). "Geothermal water to surface water: chemistry of thermal outflows in Yellowstone National Park," in *Geothermal Biology and Geochemistry in Yellowstone National Park*, eds W. P. Inskeep, and T. R. McDermott (Bozeman, MT: Thermal Biology Institute), 73–94.

Parks, D. H., and Beiko, R. G. (2010). Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 26, 715–721.

Pester, M., Schleper, C., and Wagner, M. (2011). The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr. Opin. Microbiol.* 14, 300–306.

Prokofeva, M. I., Kostrikina, N. A., Kolganova, T. V., Tourova, T. P., Lysenko, A. M., Lebedinsky, A. V., et al. (2009). Isolation of the anaerobic thermoacidophilic crenarchaeote *Acidilobus saccharovorans* sp. nov. and proposal of *Acidilobales* ord. nov., including *Acidilobaceae* fam. nov. and *Caldisphaeraceae* fam. nov. *Int. J. Syst. Evol. Microbiol.* 59, 3116–3122.

Quatrini, R., Appia-Ayme, C., Denis, Y., Jedlicki, E., Holmes, D. S., and Bonnefoy, V. (2009). Extending the models for iron and sulfur oxidation in the extreme acidophile *Acidithiobacillus ferrooxidans*. *BMC Genomics* 10:394. doi:10.1186/1471-2164-10-394

Ramos-Vera, W. H., Berg, I. A., and Fuchs, G. (2009). Autotrophic carbon dioxide assimilation in Thermoproteales revisited. *J. Bacteriol.* 191, 4286–4297.

Ramos-Vera, W. H., Weiss, M., Strittmatter, E., Kockelkorn, D., and Fuchs, G. (2011). Identification

of missing genes and enzymes for autotrophic carbon fixation in Crenarchaeota. *J. Bact.* 193, 1201–1211.

Rhee, S.-K., and Fuchs, G. (1999). Phenylacetyl-CoA:acceptor oxidoreductase, a membrane-bound molybdenum-iron-sulfur enzyme involved in anaerobic metabolism of phenylalanine in the denitrifying bacterium *Thauera aromatic*. *Eur. J. Biochem.* 262, 507–515.

Rice, G., Stedman, K., Snyder, J., Wiedenheft, B., Willits, D., Brumfield, S., et al. (2001). Viruses from extreme thermal environments. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13341–13345.

Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., et al. (2007). The sorcerer II global ocean sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* 5:e77. doi:10.1371/journal.pbio.0050077

Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., et al. (2003). A free, open-source system for microarray data management and analysis. *BioTechniques* 34, 374–378.

Schleper, C., Jurgens, G., and Jonuscheit, M. (2005). Genomic studies of uncultivated archaea. *Nat. Rev. Microbiol.* 3, 479–488.

Schut, G. J., Bridger, S. L., and Adams, M. W. W. (2007). Insights into the metabolism of elemental sulfur by the hyperthermophilic archaeon *Pyrococcus furiosus*: characterization of a coenzyme A-dependent NAD(P)H sulfur-oxidoreductase. *J. Bacteriol.* 189, 4431–4441.

Segerer, A., Langworthy, T. A., and Stetter, K. O. (1988). *Thermoplasma acidophilum* and *Thermoplasma volcanium* sp-nov from solfatara fields. *Syst. Appl. Microbiol.* 10, 161–171.

Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W. C., et al. (2007). TIGRFAMs and genome properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Res.* 35, D260–D264.

Spang, A., Hatzenpichler, R., Brochier-Armanet, C., Rattei, T., Tischler, P., Spieck, E., et al. (2010). Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends Microbiol.* 18, 331–340.

Spear, J. R., Walker, J. J., McCollom, T. M., and Pace, N. R. (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem.

Proc. Natl. Acad. Sci. U.S.A. 102, 2555–2560.

Stetter, K. O. (2006). History of discovery of the first hyperthermophiles. *Extremophiles* 10, 357–362.

Takacs-Vesbach, C., Inskeep, W. P., Jay, Z. J., Herrgard, M., Rusch, D. B., Tringe, S. G., et al. (2013). Metagenome sequence analysis of filamentous microbial communities obtained from geochemically distinct geothermal channels reveals specialization of three Aquificales lineages. *Front. Microbiol.* 4:84. doi:10.3389/fmicb.2013.0084

Teeling, H., Meyerdierks, A., Bauer, M., Amann, R., and Glöckner, F. O. (2004). Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environ. Microbiol.* 6, 938–947.

Viginais, P. M., and Billoud, B. (2007). Occurrence, classification and biological function of hydrogenases. *Chem. Rev.* 107, 4206–4272.

Volkl, P., Huber, R., Drobner, E., Rachel, R., Burggraf, S., Trincone, A., et al. (1993). *Pyrobaculum aerophilum* sp. nov., a novel nitrate-reducing hyperthermophilic archaeum. *Appl. Environ. Microbiol.* 59, 2918–2926.

Walker, C. B., de la Torre, J. R., Klotz, M. G., Urakawa, H., Pinel, N., Arp, D. J., et al. (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8818–8823.

Waters, E., Hohn, M. J., Ahel, I., Graham, D. E., Adams, M. D., Barnstead, M., et al. (2003). The genome of *Nanoarchaeum equitans:* insights into early archaeal evolution and derived parasitism. *Proc. Natl. Acad. Sci. U.S.A.* 100, 12984–12988.

Watmough, N. J., Field, S. J., Hughes, R. J., and Richardson, D. J. (2009). The bacterial respiratory nitric oxide reductase. *Biochem. Soc. Trans.* 37, 392–399.

Woese, C. R., and Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domains: the primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5088–5090.

Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U.S.A.* 87, 4576–4579.

Xu, Y., Schoonen, M. A. A., Nordstrom, D. K., Cunningham, K. M., and Ball, J. W. (1998). Sulfur geochemistry of hydrothermal waters in Yellowstone National Park: I. The origin of thiosulfate in hot spring waters. *Geochim. Cosmochim. Acta* 62, 3729–3743.

Xu, Y., Schoonen, M. A. A., Nordstrom, D. K., Cunningham, K. M., and Ball, J. W. (2000). Sulfur geochemistry of hydrothermal waters in Yellowstone National Park, Wyoming, USA. II. Formation and decomposition of thiosulfate and polythionate in Cinder Pool. *J. Volcanol. Geotherm. Res.* 97, 407–423.

Young, M., Wiedenheft, B., Snyder, J., Spuhler, J., Roberto, F., and Douglas, T. (2005). "Archaeal viruses from Yellowstone's high-temperature environments," in *Geothermal Biology and Geochemistry in YNP*, eds W. P. Inskeep, and T. R. McDermott (Bozeman, MT: Montana State University Publications), 289–304.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

FIGURE A1 | G + C content (%) frequency plot of random shotgun sequence reads (Sanger) obtained from Washburn Springs (WS_18) sediments rich in elemental sulfur and pyrite. Phylogenetic classification of each sequence read (∼800 bp) was performed using MEGAN ("blastx" analysis) (gray = total reads; orange = domain *Bacteria*; light-pink = domain *Archaea*; dark-green = Aquificales (primarily *Sulfurihydrogenibium*-like); light-green/yellow = class Thermodesulfobacteria; bronze = phylum Korarchaeota; white = class Thermoprotei; dark-blue = *Pyrobaculum*-like; light-blue = *Thermofilum*-like; other minor assignments with <100 total reads are not shown).
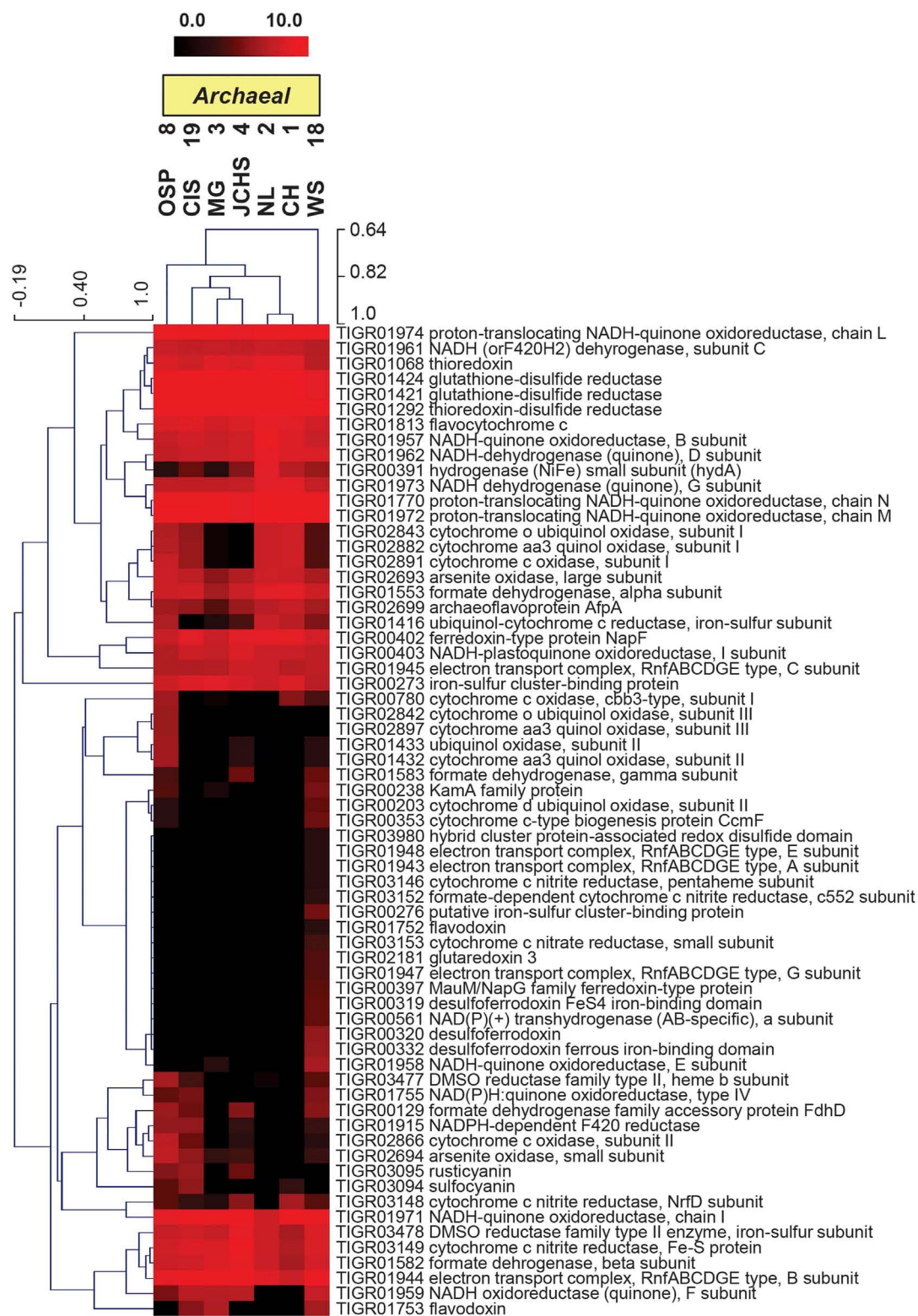
FIGURE A2 | Hierarchical cluster analysis of relative gene abundances in the TIGRFAM role category "Electron Transport" across seven archaeal-dominated geothermal communities. TIGRFAM gene families with low variation across the sites were removed before the clustering to retain ~50 of the most variable families. Subunits of the protein complexes were only represented by one representative TIGRFAM family. Pearson correlation was used as the distance measure for average linkage agglomerative clustering.
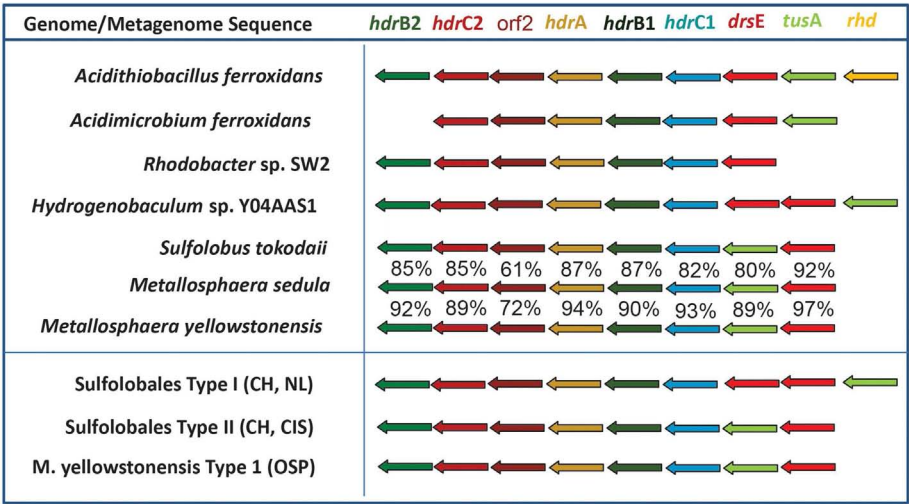
FIGURE A3 | Conservation of open reading frames found in the heterodisulfide reductase (HDR) gene complex of several acidophilic bacteria and archaea known to be utilized during the oxidation of reduced sulfur. Highly syntenous components of the HDR-gene complex are conserved in several of the thermophilic archaea found in YNP, as well as bacteria within the order Aquificales (Takacs-Vesbach et al., 2013). Percent amino acid similarities of the deduced HDR proteins are shown for *M. sedula* and *M. yellowstonensis* relative to *S. tokodaii*. The Sulfolobales Type I and *Hydrogenobaculum* sp. Y04AAS1 have two copies of *drs*E [gene names: *hdr*B1/*hdr*B2 = heterodisulfide reductase, subunit B (COG2048); *hdr*C1/*hdr*C2 = heterodisulfide reductase, subunit C (COG1150); *orf*2 = hypothetical conserved protein; *hdr*A = heterodisulfide reductase, subunit A (COG1148); *dsr*E = peroxiredoxin family protein (COG2210); *tus*A = *sir*A family of regulatory proteins; *rhd* = rhodanese-related sulfurtransferase (COG0607)].

Geomicrobiology of sublacustrine thermal vents in Yellowstone Lake: geochemical controls on microbial community structure and function

William P. Inskeep[1,2]*, Zackary J. Jay[2], Richard E. Macur[3], Scott Clingenpeel[4], Aaron Tenney[5], David Lovalvo[6], Jacob P. Beam[2†], Mark A. Kozubal[2], W. C. Shanks[7], Lisa A. Morgan[7], Jinjun Kan[8†], Yuri Gorby[8], Shibu Yooseph[5] and Kenneth Nealson[8]

[1] Thermal Biology Institute, Montana State University, Bozeman, MT, USA, [2] Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA, [3] Center for Biofilm Engineering, Montana State University, Bozeman, MT, USA, [4] DOE Joint Genome Institute, Walnut Creek, CA, USA, [5] J. Craig Venter Institute, La Jolla, CA, USA, [6] Eastern Oceanics, West Redding, CT, USA, [7] US Geological Survey, Denver, CO, USA, [8] Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA

Yellowstone Lake (Yellowstone National Park, WY, USA) is a large high-altitude (2200 m), fresh-water lake, which straddles an extensive caldera and is the center of significant geothermal activity. The primary goal of this interdisciplinary study was to evaluate the microbial populations inhabiting thermal vent communities in Yellowstone Lake using 16S rRNA gene and random metagenome sequencing, and to determine how geochemical attributes of vent waters influence the distribution of specific microorganisms and their metabolic potential. Thermal vent waters and associated microbial biomass were sampled during two field seasons (2007–2008) using a remotely operated vehicle (ROV). Sublacustrine thermal vent waters (circa 50–90°C) contained elevated concentrations of numerous constituents associated with geothermal activity including dissolved hydrogen, sulfide, methane and carbon dioxide. Microorganisms associated with sulfur-rich filamentous "streamer" communities of Inflated Plain and West Thumb (pH range 5–6) were dominated by bacteria from the Aquificales, but also contained thermophilic archaea from the Crenarchaeota and Euryarchaeota. Novel groups of methanogens and members of the Korarchaeota were observed in vents from West Thumb and Elliot's Crater (pH 5–6). Conversely, metagenome sequence from Mary Bay vent sediments did not yield large assemblies, and contained diverse thermophilic and nonthermophilic bacterial relatives. Analysis of functional genes associated with the major vent populations indicated a direct linkage to high concentrations of carbon dioxide, reduced sulfur (sulfide and/or elemental S), hydrogen and methane in the deep thermal ecosystems. Our observations show that sublacustrine thermal vents in Yellowstone Lake support novel thermophilic communities, which contain microorganisms with functional attributes not found to date in terrestrial geothermal systems of YNP.

Keywords: metagenome, Aquificales, *Archaea*, hydrogen, sulfide, methane, thermophiles, methanotrophs

Introduction

Submarine and sublacustrine thermal vents are found throughout the world and support an enormous diversity of life. Hydrothermal vent fluids often contain high concentrations of reduced constituents such as iron, sulfide, hydrogen, methane, arsenic, and/or ammonia that provide numerous possibilities for chemolithotrophic metabolism (Reysenbach et al., 2000; Amend and Shock, 2001; Coumou et al., 2008), as well as carbon dioxide important for supporting autotrophic organisms (Lovalvo et al., 2010). Hydrothermal discharge creates complex and dynamic temperature and geochemical gradients upon mixing with colder waters; the microorganisms that colonize different niches surrounding hydrothermal vents are of considerable interest in marine biology (e.g., Van Dover et al., 2001, 2007; Harmer et al., 2008), in part due to the potential microbial linkages with element cycling as well as the evolutionary implications of thermophilic organisms in marine settings (Reysenbach et al., 2000). The presence of eukaryotic mutualists adjacent to hydrothermal vents is often made possible by microbial symbionts capable of chemolithotrophic metabolism using reduced constituents present in vent fluids (Harmer et al., 2008; Setoguchi et al., 2014). Consequently, thermal vent microorganisms often conduct redox transformations and/or provide a source of nutrients important in the evolution of eukaryotes.

Prior mapping and detailed geophysical analysis of Yellowstone Lake has provided critical information on the volcanology, geologic history and current location of major thermal activity on the lake floor (Morgan et al., 2003; Morgan and Shanks, 2005; Shanks et al., 2005). Prior sampling of hydrothermal vents in Yellowstone Lake provided important background information regarding the location and characteristics of different vent types (Johnson et al., 2003; Morgan et al., 2003, 2007; Morgan and Shanks, 2005; Shanks et al., 2005). The northern region of Yellowstone Lake is one of the most seismically active areas in Yellowstone Park and supports high geothermal heat fluxes of 500–2000 mW m^{-2} (**Figure 1**). Mary Bay itself was created as a result of an explosion crater that occurred approximately 0.2 Ma (Wold et al., 1977), and numerous other smaller features in this region attest to a dynamic and recent volcanic history (Morgan et al., 2009). The isotopic and geochemical composition of Yellowstone lake waters, vent waters and tributaries have shown that elevated levels of numerous trace elements (As, Se, B, Li, Cs, Ga) in Yellowstone Lake are due to hydrothermal inputs that represent ~10% of the total chloride flux from all of the geothermal features in YNP (Shanks et al., 2005, 2007; Balistrieri et al., 2007). Moreover, Cl$^-$ vs. ^{2}H$_2$O plots place submerged vents in Yellowstone Lake on a mixing line between lake bottom-water and thermal fluids, which have an approximate temperature of 220°C (Shanks et al., 2005). High levels of trace elements, major nutrients, and/or energy sources near vent discharge have been shown to influence the diversity and productivity of biological communities in Yellowstone Lake (Lovalvo et al., 2010; Clingenpeel et al., 2011, 2013; Kan et al., 2011; Yang et al., 2011).
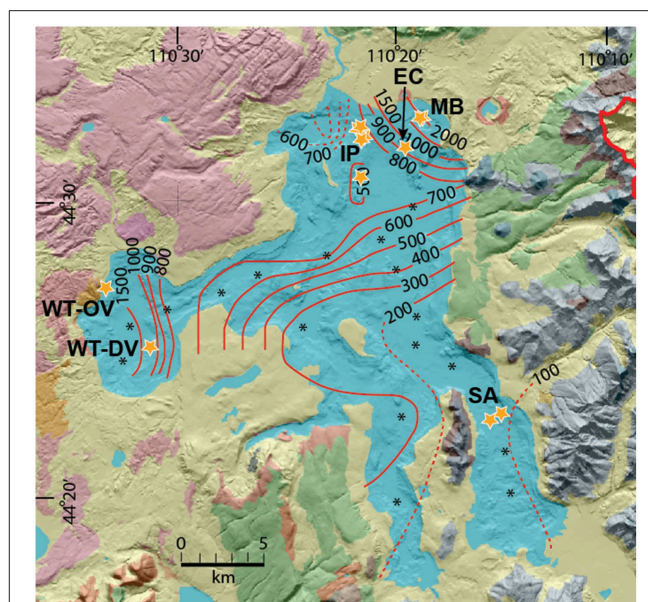
FIGURE 1 | Bathymetric map (Morgan and Shanks, 2005) of Yellowstone Lake showing heat flux iso-lines (mW/m^2) (Morgan et al., 1977) and sampling locations of thermal vents (Table 1) discussed in the current study (IP, Inflated Plain; WT-DV, West Thumb Deep Vents; WT-OV, West Thumb Otter Vent; EC, Elliott's Crater; MB, Mary Bay; SA, Southeast Arm; see Table S1 for GPS coordinates).

Efforts to characterize microbial communities from several vent sites in Yellowstone Lake using modest bacterial 16S rRNA gene surveys have shown that thermophilic bacteria from the order Aquificales were important in sulfidic habitats (Yang et al., 2011). Sulfur oxidizing Proteobacteria were also important in several vent sites, including organisms related to *Thiovirga* spp., *Thiobacillus* spp., and *Sulfuricurvum* spp. Geochemical analyses of the higher-temperature (i.e., >50°C), deeper (>49 m) vent sites (3) confirmed high levels of sulfide and other reduced sulfur species, which upon mixing with oxygenated lake water, provide habitats suitable for sulfur-oxidizing microbial communities, and which support significant rates of dark CO$_2$ fixation (Yang et al., 2011).

The prior geochemical work on Yellowstone Lake thermal vents (Shanks et al., 2005; Balistrieri et al., 2007), as well as efforts to characterize microorganisms present in these communities (Yang et al., 2011), or in filtered vent fluids (Clingenpeel et al., 2011, 2013; Kan et al., 2011), suggested that thermal vents in Yellowstone Lake contain thermophilic communities whose functional attributes can be correlated with pronounced chemosynthetic gradients. Moreover, several sublacustrine vents in Yellowstone Lake exhibit unique chemical signatures that support novel assemblages of both *Bacteria* and *Archaea*. Here we report an integrated study of hydrothermal vent geochemistry, and associated molecular and microscopic analysis of microbial communities from several of the major vent types in Yellowstone Lake (YNP, USA). The primary objectives of the study were to (i) determine the geochemical composition of hydrothermal vent fluids and predominant solid

phases associated with hydrothermal vents in Yellowstone Lake, (ii) identify predominant thermophilic microbial populations inhabiting major vent types in Yellowstone Lake using both 16S rRNA gene and random shotgun sequencing, and (iii) compare differences in functional genes observed in metagenome sequence obtained from vents exhibiting different geochemical signatures. Geochemical analysis indicated that thermal vents in Yellowstone Lake contain high concentrations of dissolved gases including H_2S, H_2, CH_4, and CO_2, as well as various trace elements and hydrogen ions (pH values ranged from 5 to 6.4 in deep vents, compared to bulk lake water pH = 7.0). Our results showed a definitive linkage between vent chemistry, microbial community structure, and associated metabolic attributes of microorganisms supported by high-temperature systems in Yellowstone Lake.

Results and Discussion

Geochemical Analysis of Sublacustrine Thermal Vents in YNP
Aqueous Samples

Temperature values measured at the sampling end of the suction arm (**Table 1**) confirmed that all vent waters collected with the ROV (**Figure S1**) had received significant inputs of hydrothermal water, and/or had been heated due to adjacent thermal activity. The large range in vent temperature(s) at a single sampling location was due to the dynamics of mixing with surrounding lake water at temperatures of 8–10°C. In most cases, stable temperatures above 60°C were maintained for extended measurement periods of 1–2 h during fluid collection. The concentrations of many constituents considered signatures

TABLE 1 | Key geochemical characteristics[a], temperature values and sample depths of sublacustrine thermal vent waters (and lake water from the Southeast Arm) obtained from Yellowstone Lake using the remotely operated vehicle (ROV) during September 2007 and 2008.

Sample	Depth	Temp	pH	DIC[a]	CO_2 (aq)[b]	DS[a]	O_2(aq)[b]	CH_4 (aq)[b]	H_2 (aq)[b]	Sample	
Location	m	°C		mM			μM		nM	ID[c]	Date
Inflated Plain	30	92–94	5.6	8.5	8.1	633	bd	21.8	414	329-Sy-1	9/9/2007
	32	70–76	5.6	4.1	3.2	463	25	21.2	4837	330-Sy-1	9/10/2007
	32	40–45	5.2	3.1	3.1	230	bd	20.9	1031	348-Sy-P	9/11/2008
	30	40–60	5.2	3.2	3.2	266	bd	22.5	1430	348-Sy-S	9/11/2008
	33.6	44–52	5.5	1.2	1.2	85	25	6.7	1023	359-VC	9/16/2008
	33.6	41–49	5.7	1.1	1.1	111	bd	5.4	1974	359-Sy-S	9/16/2008
West Thumb	52	60–66	6.2	4.7	1.6	2	113	6.4	41	339-VC	9/18/2007
Deep Vent	52	60–66	6.2	nd[d]	1.8	bd[e]	bd	7.2	63	339-Sy	9/18/2007
	52	60–76	6.2	1.3	0.4	1	bd	14.5	30	341-Sy-1	9/19/2007
	54	66	6.1	2.5	1.0	8	188	4.6	102	343-Sy-1	9/19/2007
	53.2	40–53	6.1	nd	2.1	10	147	5.6	23	369-VC	9/20/2008
	53.2	38	5.9	nd	2.0	10	238	6.8	33	369-Sy-P	9/20/2008
	53.2	62-66	5.9	nd	3.0	13	210	10.5	41	369-Sy-S	9/20/2008
Mary Bay	49.6	30–57	5.1	4.0	3.3	172	bd	13.7	454	335-Sy-S	9/16/2007
	52.3	62–66	5.1	6.4	4.5	385	bd	17.9	511	336-Sy-1	9/16/2007
	52.3	73-74	5.1	5.9	6.0	433	bd	20.4	100	336-Sy-2	9/16/2007
	50.5	65-70	5.0	3.8	3.8	bd	bd	28.1	2984	349-Sy-P	9/12/2008
	50.5	62-70	5.4	1.8	1.8	123	94	12.4	2797	349-Sy-S	9/12/2008
Elliot's Crater	14.1	71	6.4	1.4	0.5	43	202	4.4	688	351-Sy-S	9/13/2008
	14.1	71	6.2	1.6	0.7	54	190	1.9	672	351-Sy-P	9/13/2008
	14.1	63-68	6.4	1.3	0.5	40	120	2.3	660	352-VC	9/14/2008
West Thumb	4.3	65-70	8.4	0.8	0.0	bd	47	2.1	551	332-Sy	9/11/2007
Otter Vent	4.3	63-68	8.4	0.7	0.0	bd	26	0.1	43	333-VC	9/12/2007
Southeast Arm	2.5	10.6	7.0	0.6	0.013	bd	234	0.1	10	344	9/20/2007
	3	11	7.0	0.6	0.014	bd	313	0.2	47	354	9/15/2008
	17	10.5	7.1	0.6	0.019	bd	344	0.1	23	356	9/15/2008

[a]DIC, dissolved inorganic C; DS, dissolved sulfide; other constituents given in **Table S1**.
[b]Dissolved gas species determined using headspace GC, aq, aqueous.
[c]Sy, ROV syringe, P, port side, S, starboard side, VC, vent carboy/peristaltic pump.
[d]nd, not determined.
[e]bd, below detection; detection limit DS = 0.3 μM; O_2 = 3 μM.

of geothermal activity, such as dissolved CO_2, H_2, H_2S, and CH_4 were considerably higher in thermal vent waters relative to background lake water (e.g., Southeast Arm, **Table 1**). The deep thermal vents were all mildly acidic compared to bulk lake water, ranging from pH 5.1 at Mary Bay (MB), 5.2–5.6 at Inflated Plain (IP), 5.9–6.2 at West Thumb (WT), and 6.2–6.4 at Elliot's Crater (EC). A shallow (4.3 m) "alkaline siliceous" thermal vent on the west side of WT (i.e., the Otter Vent) exhibited a pH ∼ 8.2. Lower pH values at MB and IP were correlated with higher concentrations of Fe and Al (**Table S1**), consistent with mineral solubility as a function of pH. Other key indicator constituents of geothermal inputs were observed at concentrations significantly higher than background lake water (>5–10x), and included F, NH_4, As, Sb, W, Mo, Li, Cs, B, and/or Na (**Table S1**). Concentrations of major cations (Ca, Mg, K) and anions (Cl, SO_4) were generally similar in vent vs. lake waters, although vent waters at WT revealed high levels of Cl and SO_4, as well as Na.

Dissolved gas [$H_2S(aq)$, $CO_2(aq)$, $H_2(aq)$, and $CH_4(aq)$] concentrations from thermal vents were one to two orders of magnitude higher than in background lake water (**Table 1**), and were considerably higher than measured in terrestrial sites of YNP (Spear et al., 2005; Inskeep et al., 2013a). Vent waters from MB and IP contained the highest levels of total dissolved sulfide (DS), $H_2(aq)$, and $CH_4(aq)$, and were also the most acidic waters found in the study. Although, the concentrations of dissolved gases varied across different sample types collected for a given vent, the measurements were reasonably stable considering the sampling challenges presented under these circumstances (i.e., rapid mixing with bulk lake water). The large flux of $H_2S(g)$ from the IP vent region resulted in concentrations of DS well-above detection (e.g., 3–5 μM) in several surface (0–10 cm) lake samples obtained within discharge zones at IP.

Microscopy and Solid Phase Analysis

Scanning electron microscopy (FE-SEM) of vent biomass provided considerable insight regarding the characteristics of each sample, and the potential processes responsible for the formation of filamentous structures. Images of the sulfur-rich streamers from IP (**Figure S2**) reveal coccoid, rod-shaped, and filamentous organisms contained in a complex extracellular matrix including rhombohedral crystals of elemental S (**Figure 2A**). Extracellular substances were a dominant feature observed in streamers from IP, and although the exact composition of these materials is not known, the resultant "streamer structures" are very resistant to dispersion and/or disaggregation. West Thumb streamers were notably more complex, and contained diverse cellular structures, less elemental S, and more diatom shells. The vent sediments collected from MB and EC also contained numerous diatom shells intermixed with a complex suite of siliceous minerals, aluminosilicates and organic material (**Figure 2B**). The extracellular matrix evident in the thermophilic IP streamers envelopes bundles of individual filaments and sulfur crystals into dense "liquid-like" structures that exhibit significant cohesion (**Figure S2**).

Microbial Community Structure and Function

Long-fragment (>1000 bp) archaeal and bacterial 16S rRNA gene sequences indicated the major types of thermophilic microorganisms present in vent biomass (**Table 2**, **Figure 3**).
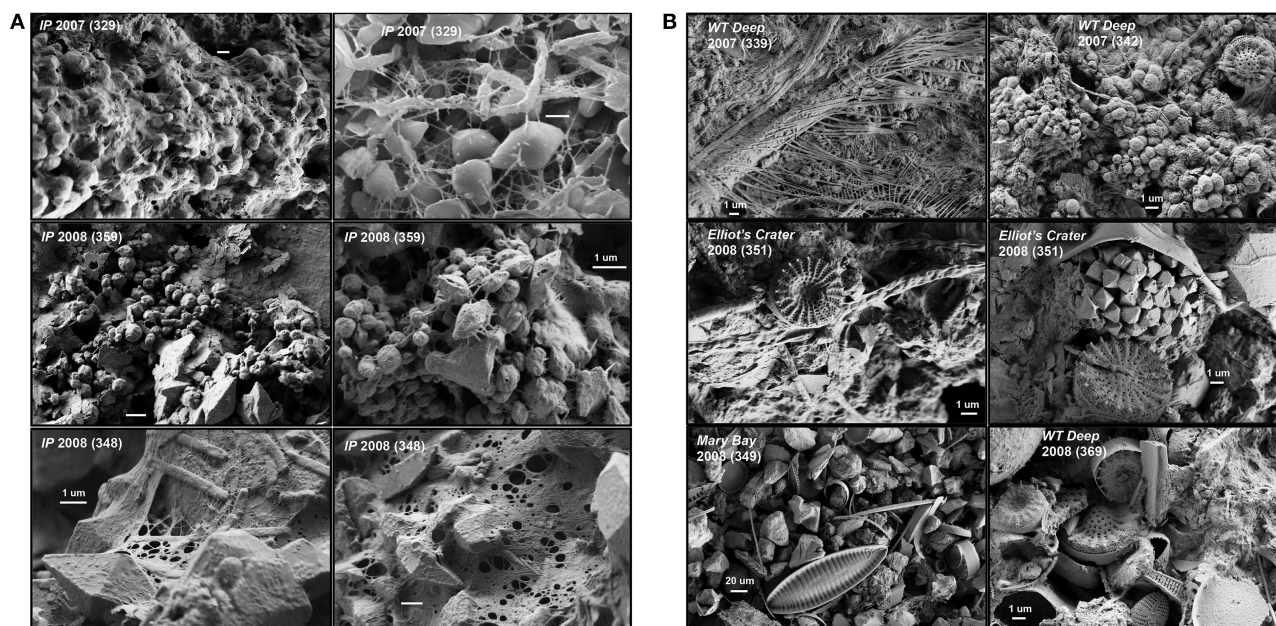
FIGURE 2 | (A) Scanning electron micrographs of thermal streamer communities obtained from 30 to 33 m vents in the Inflated Plain region, Yellowstone Lake (Sample ID). All scale bars = 1 μm. **(B)** Scanning electron micrographs of thermal vent biomass samples obtained from vent sites at West Thumb deep (Sample ID 339, 342; 2007), Elliot's Crater (351; 2008), and Mary Bay (349; 2008). Sediments associated with thermal vents show accumulation of diatom shells (e.g., Mary Bay, 349, lower right), which were also trapped in filamentous streamer communities (e.g., West Thumb, 369, lower left).

TABLE 2 | Summary of predominant long-fragment bacterial 16S rRNA gene sequences (nt > 1200 bp) obtained from thermal vent biomass samples from Yellowstone Lake using clone-library analysis (n ~ 48 per site).

Location year (ID)	Closest cultivated relative[a]	Taxonomic group[b]	Percent library	Clone ID	IDY (%)	Relative NCBI No.
Inflated Plain	*Sulfurhydrogenibium* sp. Y04ACS1	Aquificales	83	329-10-20	~99	AM259493.1
	Geothermobacterium ferrireducens	Thermodesulfobacteria	5	329-5	98.6	AF411013.1
2007 (329S)	*Caldisericum exile* AZM16c01	Candid. Div. OP5	5	329-21	98.3	NR_075015.1
	Thiovirga sulfuroxydans	Chromatiales	5	329-22	97.3	NR_040986.1
2008 (348S)	*Sulfurhydrogenibium* sp. Y04ACS1	Aquificales	100	348-5	>99	AM259493.1
2008 (359S)	*Sulfurhydrogenibium* sp. Y04ACS1	Aquificales	100	359-5	>99	AM259493.1
West Thumb	*Sulfurhydrogenibium yellowstonense*	Aquificales	30	339-15,16	>99	JQ346738.1
	Thiovirga sulfuroxydans	Chromatiales	20	339-20, 24	97.3	NR_040986.1
2007 (339)	*Methylothermus thermalis* MYH	Methylococcales	10	339-13	92.8	NR_043209.1
	Caldisericum exile AZM16c01	Candid. Div. OP5	10	339-17	98.5	NR_075015.1
2007 (342)	*Methylothermus thermalis* MYH	Methylococcales	64	342-4	93.0	NR_043209.1
	Sulfurhydrogenibium sp. Y03AOP1	Aquificales	10	342-12	99.4	NR_074557.1
	Thiobacillus denitrificans	Hydrogenophilales	10	342-5	91.0	NR_074417.1
	Curvibacter sp. PL21	Burkholderiales	3	342-7	98.9	KF206393.1
	Thermus aquaticus	Thermales	3	342-9	95.4	NR_025900.1
	Bellilinea sp. clone 96	Nov. Chloroflexi	3	342-8	<95	JQ183076.1
2008 (369S)	*Methylothermus thermalis* MYH	Methylococcales	28	369-20	92.8	NR_043209.1
	Sulfurhydrogenibium sp. Y04ACS1	Aquificales	25	369-21	99.6	AM259493.1
	Thiobacillus dentrificans	Hydrogenophilales	6	369-25	96.1	NR_074417.1
	Geothermobacterium ferrireducens	Thermodesulfobacteria	3	369-19	98.3	AF411013.1
	Fervidobacterium sp. CBS-3	Thermotogales	3	369-18	91.3	EF222230.1
	Syntrophomonas palmitatica MPA	Syntrophobacterales	3	369-22	94.5	NR_041528.1
	Thermodesulforhabdus norvegica	Syntrophobacterales	3	369-23	91.8	NR_025970.1
	Rhodoferax ferrireducens	Burkholderiales	3	369-27	97.5	NR_074760.1
	Thermus aquaticus	Thermales	3	369-24	98.0	NR_025900.1
	Desulfomicrobium thermophilum P6.2	Uncl. Proteobacteria	3	369-16	96.8	NR_042924.1
	Pelosinus sp. UFO1	Firmicutes	3	369-268	94.0	DQ295866.1
Mary Bay	*Caldisericum exile* AZM16c01	Candid. Div. OP5	12	349-20	89.9	NR_075015.1
	Thermanaerothrix daxensis GNS-1	Chloroflexi	12	349-10	92.9	HM596746.1
2008 (349S)	*Cystobacter violaceus* Cbvi34	Myxococcales	12	349-11	93.2	KF267724.1
	Desulfosarcina cetonica	Desulfurobacteriales	12	349-12	92.1	NR_028896.1
	Ornatilinea apprima	Chloroflexi	8	349-15	88.9	NR_109544.1
	Prosthecobacter fluviatilis	Verrucomicrobiales	8	349-21	83.5	NR_041608.1
	Methylobacter psychrophilus Z-0021	Methylococcales	8	349-17	98.5	NR_025016.1
	Geobacter daltonii	Desulfomonadales	4	349-13	84.4	NR_074916.1
	Syntrophus aciditrophicus SB	Syntrophobacterales	4	349-19	83.2	NR_102776.1
	Aquihabitans daechungensis	Acidimicrobiales	4	349-14	93.5	NR_132289.1
Elliot's Crater	*Thiobacillus thioparus* DSM505	Hydrogenophilales	57	351-8	96.8	NR_117864.1
	Bellilinea caldifistulae	Chloroflexi	14	351-1	90.4	NR_041354.1
2008 (351)	*Sulfurhydrogenibium* sp. Y04ACS1	Aquificales	9	351-5	98.9	AM259493.1
	Pelobacter massiliensis DSM6233TN	Desulfuromonadales	4	351-4	85.3	FR749901.1
	Denitratisoma sp. TSA61	Rhodocyclales	4	351-132	94.9	AB542411.1
	Thermomonas hydrothermalis SGM-6	Xanthomonadales	2	351-7	98.0	NR 025265.1
	Caldisericum exile AZM16c01	Candid. Division OP5	2	351-130	97.4	NR 075015.1

(Continued)

TABLE 2 | Continued

Location year (ID)	Closest cultivated relative[a]	Taxonomic group[b]	Percent library	Clone ID	IDY (%)	Relative NCBI No.
Otter Vent	*Fervidobacterium changbaicum*	Thermotogales	27	332-7	94.7	NR 043248.1
	Synechococcus sp. TS-97 B'	Cyanobacteria	11	332-16	99.7	AY884056.1
2007 (332)	*Anaeromyxobacter* sp. K	Myxococcales	11	332-13	86.0	NR 074969.1
	Thermodesulfoba. hydrogeniphilum	Thermodesulfobacteria	11	332-17	79.6	NR 025146.1
	Chloroflexus sp. strain 396-1	Chloroflexi	8	332-5	98.1	AJ308498.1
	Thiobacter subterraneus C55	Burkholderiales	8	332-18	92.1	NR 024834.1
	Acidobacteria bacterium KBS 96	Acidobacteriales	5	332-3	91.8	FJ870384.1
	Fischerella sp. JSC-11	Cyanobacteria	5	332-12	99.9	HM636645.1
	Anaerolinea thermophila UNI-1	Anaerolineales	3	332-4	98.6	AP012029.1

[a]*In some cases, clones are listed due to distant cultivated relatives. All Yellowstone Lake 16S rRNA gene sequences are deposited in GenBank [Accession Numbers KT453543 - KT453636].*
[b]*Major phylum, order, or family.*

Sulfurihydrogenibium spp. (order Aquificales) were a significant fraction of the bacterial populations observed in sulfur streamers from IP and WT, and these organisms are also found in sulfidic geothermal springs of YNP (Nakagawa et al., 2005; Reysenbach et al., 2005; Inskeep et al., 2010; Takacs-Vesbach et al., 2013). Other bacteria observed in streamer communities from IP and WT included *Caldisericum* (Candidate Division OP5), *Geothermobacterium*, *Sulfuricurvum*, *Thiovirga,* and *Thiobacillus* spp. (Proteobacteria), all of which are often found in sulfidic environments (Inskeep et al., 2005; Ito et al., 2005; Mori et al., 2009; Han et al., 2012). Deep (∼50 m) vents at WT were the only samples to exhibit relatives of *Methylothermus thermalis* (Methylococcales), and these sequences comprised ∼10, 28, and 64% of the bacteria observed in 3 independent vents from this region (**Table 2**).

The sulfur streamers from WT contained 16S rRNA gene sequences representing 6 major lineages in the *Archaea* (**Figure 3**), including members of the Korarchaeota and Euryarchaeota, which were notably absent in replicate (temporal and spatial) streamer samples from IP. Archaea present in the sulfur streamers from IP were dominated by members of the Crenarchaeota (including the Desulfurococcales and Thermoproteales), as well as a novel group of Euryarchaeota (related to the Thermoplasmatales), which are also observed in sulfur sediments of terrestrial YNP springs (Inskeep et al., 2013b). The MB sediments also contained undescribed archaeal populations including members of the Aigarchaeota, Thaumarchaeota, and Euryarchaeota (primarily relatives of methanogens), although no Crenarchaeota were observed.

Compared to IP streamers, larger contributions of non-thermophilic bacteria were detected in samples from WT, MB, and EC. Bacterial sequences from MB sediments revealed an extensive diversity of different Proteobacteria, many of which are more closely related to moderate thermophiles and/or mesophiles often found in extreme sulfur and/or iron-rich habitats (e.g., Ito et al., 2005). The greater number of different bacterial sequence types observed in MB and EC sediments (**Table 2**) was consistent with sampling constraints at these locations, which resulted in collection of a significant amount of sediment adjacent to the vent exit walls. Bacterial sequences

from the shallow phototrophic communities (pH 8.2) at the WT-Otter Vent (OV) corresponded to two major cyanobacterial groups (*Synechococccus* and *Fisherella* spp.), different members of the Chloroflexi, as well as major contributions (∼27% of the clone library) from a novel Thermotogales population (*Fervidobacterium* spp.) (**Table 2**).

Pyro-tag Sequencing

Four vent biomass samples were subjected to more intensive 16S rRNA gene sequencing as well as random shotgun sequencing. The majority of phylotypes observed using pyro-tag sequencing (**Table 3**) of IP streamers (2 sites), WT streamers, and MB sediments were also found using long-fragment sequence analysis, and provided corroborative evidence of the major taxonomic groups present. Aquificales-like sequences (i.e., *Sulfurihydrogenibium* sp.) dominated the bacterial 16S rRNA gene libraries (74–84%) obtained from two IP sulfur streamers (**Figure 4**). Conversely, the WT streamers exhibited significantly greater bacterial diversity and contained only 10% Aquificales (**Table 3**), which is consistent with lower DS and $H_2(aq)$ relative to the vents at IP (**Table 1**). Mary Bay vent sediments contained very few Aquificales sequences, consistent with the lack of any notable streamers at this site, and the significant contribution from mesophilic organisms. Populations related to *Caldisericum exile* (Mori et al., 2009; candidate phylum OP5) were observed in all samples, but especially in association with the sulfur streamers at IP (**Figure 4**).

Other major groups of *Bacteria* varied with vent sites, but included members of the Bacteroidetes, Proteobacteria (the Epsilon group was more important in IP whereas Beta and Delta groups were more important in WT and MB), Thermotogae and Deinococcus-Thermus (7.2 and 4.5% in WT streamers), Acidobacteria (4% in *MB* sediments), Actinobacteria (5.4% in MB sediments), Thermodesulfobacteria (2.4% in WT), Planctomycetes ∼2% in WT and MB), as well as members of the Chloroflexi (∼6–8% in WT and MB sediments) (**Table 3**). It is unlikely that Chloroflexi-like sequences are contributed from organisms conducting photosynthesis at these depths; phylogenetic placement of long-fragment 16S rRNA sequences that were highly related to the shorter pyro-tag reads suggest that
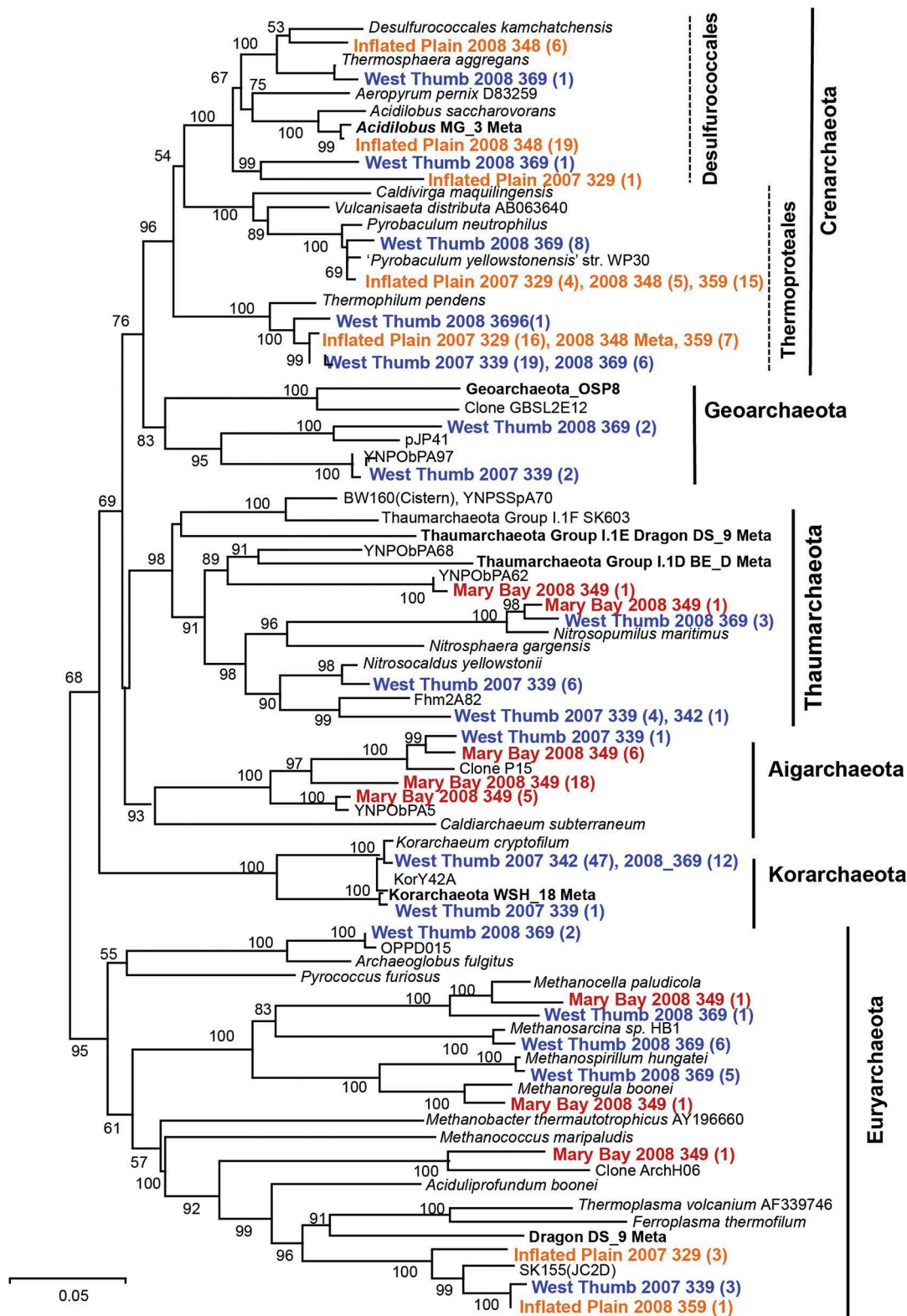
FIGURE 3 | Phylogenetic tree (16S rRNA gene sequences) of the domain *Archaea* including long-fragment sequences observed in thermal vent microbial communities from Yellowstone Lake (neighbor-joining tree; bootstrap values reported based on 1000reps Log Det.). All long-fragment 16S rRNA gene sequences from Yellowstone Lake are deposited in GenBank (KT453543-KT453636).

TABLE 3 | Major taxonomic groups (fraction of total bacterial or archaeal sequences) in vent-associated microbial communities determined using pyro-tag 16S rRNA gene sequencing of amplicons generated with universal bacterial (top) and archaeal (bottom) primer sets (Clingenpeel et al., 2011, 2013; Kan et al., 2011).

Taxonomic groups[a]	Site name			
Bacteria	IP 348S	IP 359S	WT 369S	MB 349S
Aquificae	83.8	73.9	10.3	0.4
Caldiserica	5.8	4.9	0.7	2.6
Bacteroidetes	1.0	2.8	12.7	23.9
Proteobacteria	4.5	13.1	25.3	13.6
Acidobacteria	2.2	0.5	1.4	3.6
Chloroflexi	0.1	0.0	6.4	7.4
Thermotogae	0.0	0.1	7.1	0.1
Actinobacteria	0.3	0.6	1.3	5.0
Deinococcus-Thermus	0.0	0.0	4.5	0.0
Firmicutes	0.1	0.3	3.3	2.1
Cyanobacteria	0.1	0.0	1.0	12.7
Thermodesulfobacteria	0.8	0.1	2.4	0.1
Chlorobi	0.0	0.0	2.2	1.1
Planctomycetes	0.0	0.1	1.3	1.1
Gemmatimonadetes	0.0	0.0	0.5	1.6
Nitrospira	0.0	0.0	1.0	0.3
Unclassified Bacteria	1.0	2.7	15.6	21.5
Total[b]	99.7	99.1	97	97.1
n	27540	32471	27338	25753
Archaea	IP	IP	WT	MB
CRENARCHAEOTA				
Desulfurococcales	48.8	12.5	17.4	17.0
Thermoproteales	49.2	86.2	21.9	14.2
Sulfolobales	0.2	0.2	0.0	0.1
Other "Crenarchaeota"	1.8	0.9	2.4	22.8
EURYARCHAEOTA				
Methanomicrobiales	0.0	0.0	0.3	0.6
Thermoplasmatales	0.0	0.0	0.5	0.9
Novel Euryarchaeota	0.0	0.1	41.6	12.2
Korarchaeota	0.0	0.0	9.7	23.1
Unclassified Archaea	0.0	0.0	5.4	8.7
Total[b]	100.0	99.9	99.2	99.6
n	24191	13945	13495	26382

[a]RDP training set 9, RDP Naive Bayesian rRNA Classifier version 2.5, May 2012 Classifications performed March 18, 2013. Novel Crenarchaeota include what was referred to as "Marine Crenarchaeota," now established within the Candidate phylum Thaumarchaeota.
[b]Total = percent of total sequences (n).

many of the Chloroflexi sequences were contributed by relatives of anaerobic, heterotrophic strains (Yamada et al., 2007; Klatt et al., 2013) (**Table 3**).

Different types of *Archaea* were observed across sites (**Table 3**), and the major groups identified using pyro-tag sequencing were also observed in long-fragment clone libraries (e.g., **Table 2**, **Figure 3**). The highly sulfidic and H$_2$-rich IP streamers (pH ∼ 5.2–5.6) exhibited a consistent signature of Crenarchaeota (>99% of archaeal reads), including members of the Thermoproteales (*Pyrobaculum* and *Thermofilum*-like populations) and Desulfurococcales (*Desulfurococcus* and *Acidilobus*-like sequences; Jay et al., 2014). Very few Sulfolobales sequences were observed, which is expected given the pH range of these vent communities (pH 5–6) (Macur et al., 2013; Jay et al., 2014). Members of the Korarchaeota were found primarily in the less sulfidic and higher pH streamers from WT, as well as in sediments from MB (**Table 3**). A significant number of novel euryarchaeotal sequences were observed in WT and MB, and represent several novel methanogens, an undescribed group related to the order Thermoplasmatales (∼85% nt identity, **Figure 3**), as well as members of the Thaumarchaeota and Aigarchaeota (Brochier-Armanet et al., 2008; Nunoura et al., 2011). Long-fragment clone libraries also indicated the presence of different types of Euryarchaeota and Thaumarchaeota in WT streamers and MB sediments (**Figure 3**), including relatives of both low-temperature thaumarchaea (Hatzenpichler, 2012) as well as thermophilic clades (Beam et al., 2014). The korarchaeotal sequences observed using pyro-tag analysis (∼10–23% of WT and MB pyro-tag sequences) corresponded to long-fragment 16S rRNA gene sequences, which were observed at several WT vent sites in both 2007 and 2008 (**Figure 3**).

Metagenome Sequence Analysis

Random shotgun sequence (average read length ∼400 bp) obtained from four vent sites (IP, WT, and MB) was analyzed using Blastx (NCBI) and G + C content (%) to examine the predominant populations present in each site (**Figure 5**). The random sequence data indicated a lower abundance of archaea relative to bacteria in all vents sampled, representing from less than 5% of the total sequences in three of the four vent sites up to nearly 30% in one of the sulfur streamers from IP (348S). The major phylotypes identified with random sequence were also consistent with those observed using amplification techniques. For example, random sequence reads from two different streamer communities from IP were dominated by sequences related to *Sulfurihydrogenibium*, *Caldisericum*, and other Proteobacteria (**Figure 5**). The streamer communities from WT were dominated by sequences related to members of the Bacteroidetes, Aquificales, and Proteobacteria, and the sediments from MB contained a diverse assemblage of distant relatives of the Bacteroidetes (lower G + C), Proteobacteria (higher G + C), Chlamydiae/Verucomicrobia, and Actinobacteria. Much of the random shotgun sequence from MB (and to a lesser extent in WT) was not sufficiently similar to reference organisms (NCBI) to assign individual sequence reads to specific genera.

The amount of assembled genome sequence (**Table S2**) obtained from the four vent sites was inversely correlated with the number of dominant sequence types observed using 16S rRNA gene inventories. For example, the higher percent of reads assembled from IP streamers (348S and 359S) resulted in larger contigs with higher sequence coverage (**Table S2**). Phylogenetic assignment of 16S rRNA genes obtained from
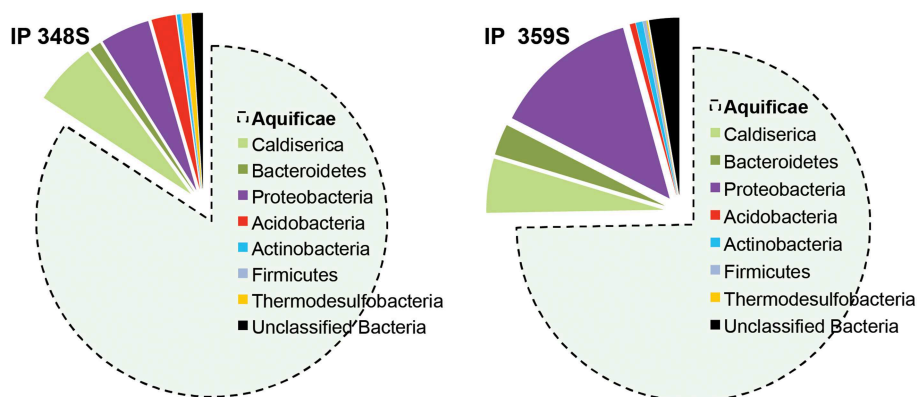
FIGURE 4 | Phylogenetic classification of short fragment bacterial 16S rRNA gene sequences from two different sulfur streamer communities from Inflated Plain (depth ∼ 30–33 m; pH ∼ 5.6) obtained using pyro-tag sequencing (sequences classified using RDP Naïve Bayesian rRNA Classifier version 2.5; also see Table 3).

assembled sequence (**Table S3**) was consistent with populations observed using 16S rRNA gene-only approaches (**Tables 2**, **3**, **Figures 3**, **4**). Consequently, the sequence assemblies from IP and WT represent an excellent opportunity for linking specific metabolic genes with known phylotypes.

Functional Gene Analysis

The predominant energy cycling reactions mediated by microorganisms present in vent communities was investigated using specific query (marker) genes that code for proteins known to mediate the assimilation of inorganic C, electron transfer, and/or stress response (**Table 4**). Nearly, all phylotypes identified using different functional genes were consistent with those determined using phylogenetic analysis of 16S rRNA genes (e.g., **Figures 3**, **4**, **Tables 2**, **3**). Consequently, a consistent picture emerges regarding the functional attributes of major population types identified in IP and WT (**Table 4**). The low fraction of assembled sequence obtained from the MB vent sediments precluded confident assignment, and the majority of genes identified were less than 25–30% of their full length (not shown).

Metabolic evidence for the fixation of carbon dioxide (CO_2) via the reductive TCA cycle (e.g., ATP citrate lyase; Takacs-Vesbach et al., 2013) was identified in all streamer communities, and was especially evident in *Sulfurihydrogenibium* (Aquificales) populations (**Table 4**). Copies of acetyl-CoA carboxylase (*accA*) were noted in several bacterial phylotypes as well as a Thermoproteales population in IP 348S. In bacteria, acetyl-coA carboxylase is required for the synthesis of fatty acids. Consequently, the phylogenetic identity of these genes is essentially consistent with the major bacterial phylotypes present across the 3 streamer communities. In the *Archaea*, acetyl-CoA carboxylase is involved in the 4-hydroxybutyrate/3-hydroxyproprionate CO_2 fixation cycle (or decarboxylase version) (Berg et al., 2007); however, this gene was only observed in the *Thermofilum pendens*-like population present in one of the IP streamer communities (348S), and other key marker genes for the 4-HB/3-HP pathway were not observed (Berg et al., 2007, 2010). Consequently, the sequence

data suggest that the primary mechanism of CO_2 fixation in these communities occurs via the reductive-TCA cycle (Beh et al., 1993; Hügler et al., 2007), and supports measurements of dark CO_2 fixation rates obtained in a prior study (Yang et al., 2011).

Genes coding for proteins known to be important in the oxidation of reduced sulfur species were observed in these communities, and were most-closely related to the dominant bacterial populations present including *Sulfurihydrogenibium*, *Sulfuricurvum*, *Thiovirga*, *Thiobacillus,* and *Caldisericum* spp. (**Table 4**). Specifically, *hdrAB* genes indicative of a S oxidation pathway (Friedrich et al., 2005) were found in *Sulfurihydrogenibium* sequences, as has been observed in terrestrial sites of YNP (Takacs-Vesbach et al., 2013). Several other key marker genes and pathways for S oxidation (*sqr*, *sox*) were identified as *Sulfuricurvum*, *Thiobacillus,* and *Thiovirga* spp., as well as *Sulfurihydrogenibium*-like (**Table 4**).

Group I Ni-Fe hydrogenases, indicative of H_2 uptake and oxidation (Viginais and Billoud, 2007), were found in *Sulfurihydrogenibium*, *Thermofilum*, *Thermoproteus*, and *Thiobacillus*-like assemblies (**Table 4**). The hydrogenases present in the *Sulfurihydrogenibium*-like sequence assemblies are most closely related to other Aquificales genera, because the only known *Sulfurihydrogenibium* sp. to contain a Group 1 Ni-Fe hydrogenase is *S. azoricus* (Aguiar et al., 2004; Reysenbach et al., 2009). To date, the *Sulfurihydrogenibium*-like populations characterized in terrestrial sites of YNP do not contain Group I Ni-Fe hydrogenases (Inskeep et al., 2010; Hamamura et al., 2013; Takacs-Vesbach et al., 2013). The higher concentrations of H_2(aq) ($>4 \mu M$) at IP vent sites correlates with the presence of hydrogenases in *Sulfurihydrogenibium*-like sequences found in two replicate streamer communities (348S, 359S).

A near-complete methane oxidation pathway (particulate methane monooxygenase subunits ABC) was identified in the streamers from WT (369S) (with the exception of formaldehyde dehydrogenase). The *pmoABC* genes were most closely related to genes from the gamma-proteobacterium *Methylothermus subterraneus* (95% nt identity for *pmoA*) (Tsubota et al., 2005; Hirayama et al., 2011). Bacterial populations (similar to *M.*
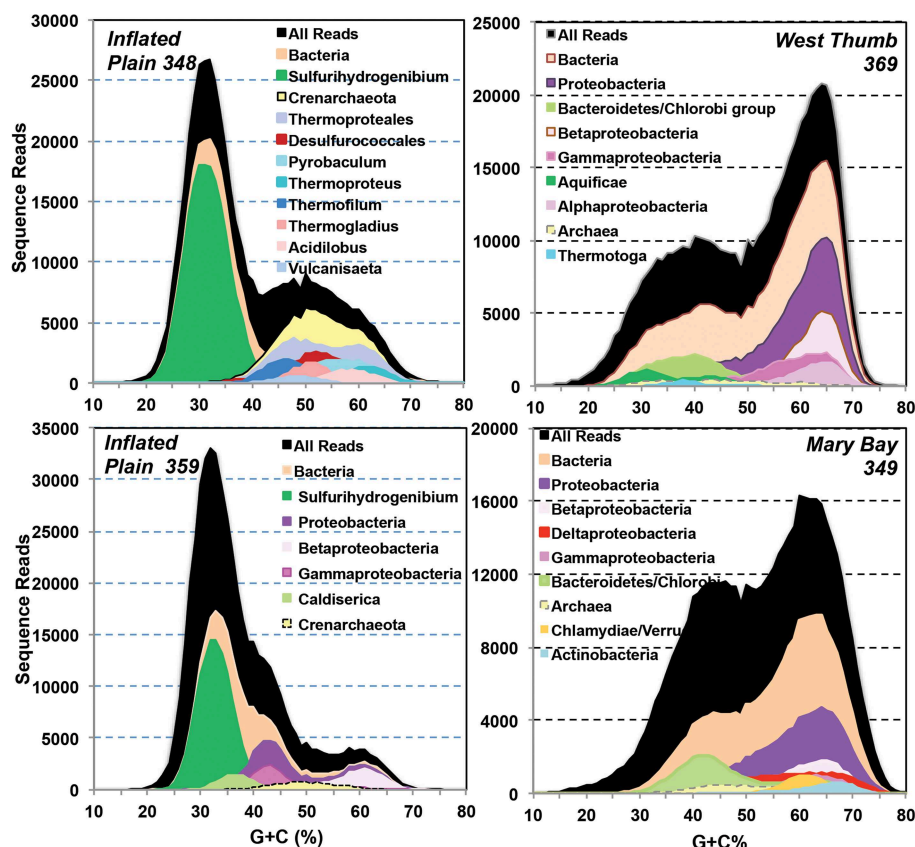
FIGURE 5 | Random shotgun sequence reads from four Yellowstone Lake thermal vent microbial communities plotted as a function of G + C content (%) and subjected to phylogenetic analysis using blast (90% identity). A significant number of sequence reads were not related to bacteria or archaea in current public databases.

subterraneus and *M. thermalis*) were identified as major taxa in WT streamers (pH $\sim$ 6.1) in both 2007 and 2008 ($n =$ 3) (**Table 2**). Moreover, no *pmoABC* genes were identified in other vent sites. The *pmoCAB* operon architecture (Ward et al., 2004) was not recovered from the metagenome assembly, and a definitive pathway of CO_2 fixation via formaldehyde assimilation could not be determined, as the key gene for the ribulose monophosphate pathway (3-hexulose-6-phosphate synthase) was not identified. To date, *pmoABC* genes have not been observed in metagenomes from numerous terrestrial sites in YNP (Inskeep et al., 2010, 2013a; Swingley et al., 2012). Moreover, methanotrophs and/or methylotrophs have not been observed as dominant population types in terrestrial thermal habitats characterized to date, despite fairly high concentrations of $CH_4(aq)$ in some locations (e.g., 1–2 μM; Inskeep et al., 2005, 2013a). Concentrations of $CH_4(aq)$ measured in vent sites at IP, WT, and MB ranged from 5 to 30 μM (**Table 1**); however, WT (i.e., pH $\sim$ 6; T $\sim$ 60°C, lower sulfide) was the only site to exhibit abundant methanotrophic population(s). The lower pH values and higher sulfide of vents at IP and MB (**Table 1**) may preclude methanotrophic populations, as these conditions are not optimum for the oxidation of CH_4 using O_2 as an electron acceptor (Tsubota et al., 2005).

Oxygen is an important electron acceptor in thermal vent communities of Yellowstone Lake as evinced by the presence of Type C (cbb3) heme Cu oxidases in many of the dominant bacterial population types, including *Sulfurihydrogenibium*, *Sulfuricurvum*, *Rhodoferax*, *Thiomonas*, and *Thiobacillus*-like populations (**Table 4**). These types of heme Cu oxidases have been shown to exhibit low K_m values for O_2, and are often found in hypoxic environments (García-Horsman et al., 1994; Jünemann, 1997; Borisov et al., 2011). Ubiquinol oxidases (e.g., *cydA*) were also observed in several archaeal populations present in the highly sulfidic sites at IP (e.g., Desulfurococcales and Thermoproteales; Jay et al., 2014, 2015) as well as in the Thermodesulfobacteria at WT. These oxidases are common in hypoxic environments and may function in respiration or as O_2 scavenging proteins (Borisov et al., 2011).

Other electron acceptors important for specific members of these communities may include elemental S, arsenate, nitrate, and sulfate (**Table 4**). Novel DMSO molybdopterins (tabulated as *psrA/sreA*) related to *Sulfurihydrogenibium* and Thermoproteales populations in the IP streamers may play a role in the reduction of elemental sulfur and/or arsenate (Jay et al., 2015), and these metabolisms would be expected within the S-rich streamer fabric (**Figure S2**). The only evidence of dissimilatory nitrate reduction

TABLE 4 | Summary of functional genes[a] (and their phylogenetic identity[b]) related to key geochemical processes, which were identified in assembled metagenome sequence of three thermal vent microbial communities from Yellowstone Lake, WY.

Process/Pathway	Marker genes[a]	Inflated Plain (IP-348S)	Inflated Plain (IP-359S)	West Thumb (WT-369S)
FIXATION Of CO$_2$				
ATP citrate lyase	aclB	Sk, Sy	Sk, Sy	
Citryl coA lyase	ccl	Sy	Sy	Sy
Citryl coA synthetase	ccsA	Tu	Ce, Sk, Sy	At, Sy, Td
Acetyl-coA carboxylase	accA	Ce, Sy, Tp	Acid, Ce, Sk, Sy, Thio	Ma, Td
OXIDATION REDUCTION				
Hydrogen oxidation	hynS, hynL	Sy, Tp, Tu,	Sy, Tp	Td
Elemental S oxidation	hdrA/hdrB	Sy	Sy	Sy, Td
Sulfide oxidation	sqr	Sy,	Sk, Sy, Thio	Td,
Sulfur transferase	rdh	Sk, Sy	Sy, Sk	Nl, Sy, Td
Sulfur oxidation	soxBCDY	Sy,	Ce, Sy, Thio	Sy, Td
Methane oxidation	pmoABC			Ms
Formate oxidation	fdh	As, Ce	Acid, Ce	Ma
Oxygen reduction, Heme Cu Oxidases	cbb3	Sy	Rf, Sk, Sy, Thio	Sy, Td
Oxygen reduction, bd-ubiquinol type	cydA	As, Py, Sy, Tp, Tu	Py, Tp	Thdes
Sulfur reduction	psrA/sreA[c]	Py, Sy, Tu	Py	
Sulfate reduction	dsrAB			Des, Td
Nitric oxide reduction	norB	Py, Tu	Py	
Nitrate reduction	narG		Sk	
STRESS RESPONSE				
Arsenite efflux-detoxification	arsB	As, Py, Sy, Tu	Sk, Sy	Sy
Arsenate reduction-detoxification	arsC	Ce, Sy	Ce, Sy	
Mercuric reductase	merA		Thio	
Heavy-metal ATPases	znta		Sy, Thio	Ma, Sy
Superoxide dismutase	sodA	As, Py, Tu	Acid, Py, Thio	Mp
Hydroperoxide reductase (peroxiredoxin)	perox	As, Ce, Sy, Tp, Tu	As, Ce, Sk, Sy, Tp	Ma, Sy, Ta, Td, Thio
Desulfoferrodoxin (superoxide reductase)	sorA	Ce, Tp	Ce, Tp	
Rubredoxin	rub	Ce, Sy	Acid, Ce, Sy	Thdes
Motility	flaB	Sy, Thio	Sy, Thio	Fp, Sy

[a]*Functional genes that code for proteins with high specificity for possible pathway; no genes were found for nitrification (amoA), denitrification (e.g., nirK, nirS, nosZ), methanogenesis (mcrA), thiosulfate oxidase (tqoAB), or arsenite oxidation (aroA = aioA); a ferric reductase from an Acidovorax sp. population was observed in WT.*

[b]*Population Types (closest relatives): Acid, Acidovorax sp.; As, Acidilobus saccharovorans; At, Anaerolinea thermophila; Ce, Caldisericum exile; Des, Desulfobacterium sp.; Fp, Fervidobacterium pennivorans; Ma, Methylomicrobium alcaliphilum; Ms, Methylothermus subterraneus; Mp, Mucilaginibacter paludis; Py, Pyrobaculum sp.; Tu, Thermoproteus uzoniensis; Thio, Thiovirga sulfuryoxidans; Sy, Sulfurihydrogenibium sp.; Sk, Sulfuricurvum kujiense; Tp, Thermofilum pendens; Td, Thiobacillus denitrificans; Ta, Thermocrinis albus; Nl, Nitrosoarchaeum limnia; Thiom, Thiomonas sp.; Thdes, Thermodesulfobacteria.*

[c]*includes unclassified DMSO proteins that may be related to sulfur and/or arsenic reduction in the Thermoproteales (Jay et al., 2015).*

(*narG*) was found in the *Sulfuricurvum* population from IP. The role of *norB* genes present in several Thermoproteales populations is not fully understood (NorB may also exhibit activity as an oxygen reductase), in part because no evidence for a complete denitrification pathway has been documented in this group of organisms (Jay et al., 2015). The only evidence of sulfate reduction (*dsrAB*) was associated with nonthermophilic populations at WT (i.e., *Thiobacillus*, *Desulfobacteria*).

Summary

Subaerial thermal vents in Yellowstone Lake make a significant contribution to the total chloride flux from the Yellowstone hot spot, and exhibit high concentrations of electron donors (e.g., H$_2$S, H$_2$, CH$_4$) capable of supporting active thermal

microbial communities. The high concentrations of CO$_2$, H$_2$S, H$_2$, and CH$_4$ in thermal vents of Yellowstone Lake (**Table 1**) are approximately an order of magnitude higher than many terrestrial systems of YNP (Inskeep et al., 2005, 2010, 2013a; Spear et al., 2005). Consequently, geochemical attributes of Yellowstone Lake thermal vents make them unique for geomicrobiological investigation.

The streamer communities from IP were comprised primarily (>80–85%) of *Sulfurihydrogenibium* spp., and these habitats appear to be highly similar to those observed in terrestrial sites where these filamentous bacteria grow in turbulent, sulfidic channels ranging from pH = 6–8 and T = 65–85°C (Reysenbach et al., 2005; Fouke, 2011; Takacs-Vesbach et al., 2013). However, Ni-Fe hydrogenases were identified in *Sulfurihydrogenibium* populations from replicate IP vent communities, and these

genes have not been found in terrestrial *Sulfurihydrogenibium* assemblies from MHS where the concentrations of $H_2(aq)$ are an order of magnitude lower than those measured in subaerial thermal vents. Streamer communities in the higher pH (~6.1), lower sulfide (<10 μM) habitats of WT contained less *Sulfurihydrogenibium*, which is consistent with the distribution of this organism as a function of sulfide and hydrogen.

All deep thermal vents contained high levels of dissolved CH_4; however, *Methylothermus* populations and associated *pmoABC* genes were found only in vent biomass from WT (pH 6; lower sulfide). Cultivated *Methylothermus* spp. utilize CH_4 as an energy source under aerobic and/or microaerobic conditions at optimum temperature and pH values ranging from ~55 to 60°C and 6–7, respectively (Tsubota et al., 2005; Hirayama et al., 2011). This is the first CH_4 oxidation pathway identified from a thermophile in YNP, and correlates with high levels of $CH_4(aq)$, low sulfide concentrations, circumneutral pH values, and temperatures near 60°C (**Table 1**). Other microbial populations observed in WT streamers are consistent with the higher pH of these habitats (relative to IP), and included considerably greater numbers of novel organisms more closely related to moderately thermophilic and/or mesophilic Proteobacteria, Thermotoga, Chloroflexi, Bacteroidetes, and other Aquificales (i.e., *Thermocrinis*-like), as well as a small contribution (~10%) from Archaea.

Although bacteria were more abundant in subaerial vent samples obtained in this study, thermophilic archaea were also observed and included several novel groups. Members of the Thermoproteales and Desulfurococcales were the most numerous archaea in sulfidic habitats from IP, and their occurrence in elemental S streamers is consistent with observations from other circumneutral geothermal environments distributed globally. In contrast, the deep vents in WT and MB contained greater numbers of Euryarchaeota, Korarchaeota, Aigarchaeota, and Thaumarchaeota. The presence of archaea potentially involved in methanogenesis (e.g., *Methanosarcina*, *Methanospirillum* spp.) may be supported by the high concentrations of CO_2 and H_2 in these vent waters. Members of the Korarchaeota were observed primarily in vents from WT, and represented one of the important archaeal groups amplified from vent biomass in both 2007 and 2008. The distribution of korarchaea has been limited to habitats ranging from ~pH 6 to 8 (Auchtung et al., 2011; Miller-Coleman et al., 2012; Inskeep et al., 2013b); consequently, this may be one factor explaining why members of this phylum were not found in other vent samples. Our results document the presence of novel populations not found hitherto in geothermal habitats of YNP. Moreover, these populations exhibit functional attributes consistent with the geochemistry of thermal vent habitats, such as high concentrations of dissolved CO_2, H_2S, H_2, and/or CH_4.

Methods

Sampling

At least 22 sublacustrine hydrothermal vents were sampled during 2007 and 2008 in the Inflated Plain (IP), Mary Bay (MB), and West Thumb (WT) regions of Yellowstone Lake,

Yellowstone National Park (YNP) (**Figure 1**, **Table 1**). Lake water from Southeast Arm (SA) was also sampled for comparison to thermal vent samples, because no thermal vents are found in this region of the lake, and this location is nearly 5 km from major vent sites and total heat flux in the northern region. Several prior studies (1996 and 1999) on thermal vents from these locations (Balistrieri et al., 2007) provided important background information regarding the general location and properties of vent fluids. However, given the number of vents within these active regions, coupled with the difficulty of locating and sampling vents, the vents sampled here are not necessarily identical to those sampled in prior studies. Sublacustrine vent fluids and solid phase samples were collected in September 2007 and September 2008 using a remotely operated vehicle (ROV) tethered to the *Cutthroat* (**Figure S1**). Thermal vent waters were obtained to minimize mixing with lake water by using Norprene™ tubing attached a small-diameter suction arm inserted directly into vent discharge sites after positioning the remotely operated vehicle (ROV, **Figure S1**). Vent waters were collected using retractable polycarbonate piston syringes (2) onboard the ROV, or by a peristaltic pump located onboard the *Cutthroat*. A thermocouple located on the end of the sampling arm was used to continuously record temperature during sample collection. Vent biomass and associated sediments were obtained using the side-mounted syringes (port and starboard), or a separate sampling can (2008 samples). Video of several sampling sites show visible "shimmering" caused by hot water discharge, as well as associated filamentous streamer communities; **Figures S3–S7**).

Geochemical analyses were performed immediately on the *Cutthroat* for time-sensitive constituents (e.g., dissolved oxygen, sulfide, pH); other sample types were either preserved and stored for further characterization at a temporary field laboratory established at Lake Village or at Montana State University (Bozeman). Glutaraldehyde (1% final concentration) was used to preserve samples for field-emission scanning electron microscopy (FE-SEM). All molecular samples were frozen on dry ice, then transferred to a −80°C freezer.

Aqueous Geochemistry

Several chemical species were analyzed onboard the *Cutthroat*, and included Fe^{II} and Fe^{III} (Ferrozine method; To et al., 1999), total dissolved sulfide (DS) (amine sulfuric acid method; APHA, 1998), pH, and dissolved oxygen (Winkler method; APHA, 1998). Aqueous pH values were obtained using a Fisher Accumet AP-71 m and AP-55 probe equipped with temperature compensation. Additional aqueous samples were filtered (0.2 μm) directly into sterile 50 mL Falcon tubes and refrigerated at 4°C. Two samples were preserved with trace metal grade HNO_3 (1%) and HCl (0.5%) for analysis using inductively coupled plasma (ICP)-optical emission spectrometry (OES) (Perkin Elmer) and ICP-mass spectrometry (MS) (Aligent Model 7500)] for total dissolved elements including Ag, Al, As, Ba, Be, Bi, B, Ca, Cd, Ce, Co, Cr, Cs, Cu, Dy, Er, Eu, Fe, Ga, Gd, Ge, Hf, Ho, In, K, La, Li, Lu, Mg, Mn, Mo, Na, Nb, Nd, Ni, P, Pb, Pr, Re, Rb, Sb, Sc, Se, Si, Sm, Sn, Sr, Ta, Te, Tb, Th, Ti, Tl, Tm, U, V, W, Y, Yb, Zn, and Zr. One unacidified sample was analyzed for predominant inorganic anions (F^-, Cl^-, SO_4^{2-}, NO_3^-, $S_2O_3^{2-}$, AsO_4^{3-}) using

anion exchange chromatography (Dionex DX 500; AS16-4 mm column), and aqueous NH_4^+ using the phenolate colorimetric ($A_{630\ nm}$) procedure on a flow injection analyzer (APHA, 1998). Dissolved inorganic carbon (DIC) and dissolved organic C (DOC) were determined on separate samples taken in closed headspace, baked (500°C) serum bottles using a Shimatzu Model TOC-VCSH total C analyzer. Aqueous samples were collected using either the ROV syringe or the peristaltic pump mentioned above to pump vent fluids through a 140 mm diameter filter (0.4 μm) into closed 160 mL serum bottles. The concentrations of dissolved gases (H_2, CH_4, and CO_2) were determined in the laboratory using headspace gas chromatography with a dual-channel Varian gas chromatograph (Model CP2900) equipped with thermal conductivity detection (Inskeep et al., 2005). Aqueous geochemical modeling was performed using temperature corrected thermodynamic constants (Allison et al., 1991; Inskeep et al., 2005).

Characterization of Vent Biomass

Vent biomass samples (streamers and/or sediments) were analyzed using a field emission scanning electron microscope (FE-SEM) coupled with energy dispersive analysis of x-rays (EDAX). Aliquots of glutaraldehyde (1%) stored samples were aseptically transferred to 10 mm diameter (0.2 μm) filters, washed with nano-pure water, and then placed on Al stubs for sputter-coating with Ir. Imaging was performed at low voltage (1 kV) and small working distances (∼4 mm), whereas elemental analysis was performed at 15 kV and 15 mm working distance.

Microbial Community Analysis: DNA Extraction, Amplification and Sequencing

Microbial mat samples were analyzed to assess the predominant 16S rRNA gene sequences distributed across different thermal vents. Streamer and/or sediment samples collected using the ROV were immediately placed on dry ice, and stored within 24 h at −80°C. Total DNA was extracted from the samples using the FastDNA SPIN Kit for Soil (Q-Biogene, Irvine, CA). The primers used for near full-length amplification of 16S rRNA genes included the *Bacteria*-specific Bac8f (5′-AGA GTTTGATCCTGGCTCAG-3′) and the *Archaea*-specific Arc2f (5′-TTCCGGTTGATCCYGCCGGA-3′) primers, each coupled with universal primer Univ1392r (5′-ACGGGCGGTGTGTAC-3′). Purified PCR products were cloned using the pGEM-T Vector System from Promega Corp. (Madison, WI), and the inserts were sequenced using T7 and SP6 primers (TGEN, Phoenix, AZ). Resultant sequences were edited, and checked for chimeras.

Metagenome and Pyro-tag Sequencing

Four biomass/sediment samples from three thermal vent locations (Inflated Plain, West Thumb, Mary Bay) were subjected to random 454 pyrosequencing and pyro-tag analysis of 16S rDNA amplicons obtained using two different primer sets focused on *Bacteria* and *Archaea*. DNA extractions (as described above) were used to provide starting material for random 454 pyrosequencing and amplification steps necessary to generate short-fragment 16S rDNA amplicons for pyro-tag analysis (Clingenpeel et al., 2011; Kan et al., 2011).

Four vent samples (348S, 349S, 359S, 369S) received one-half plate of random sequencing (median trimmed read length = 360–400 nucleotides). A split DNA sample of IP 359S was used to generate a paired-end library, which generated a greater number of longer contigs (**Table S2**). Assemblies of all five random sequence libraries were generated using both Celera (Version 4.0) and Newbler assemblers (Celera assembly parameters: doOverlapTrimming = 0, doFragmentCorrection = 0, globalErrorRate = 12, utgErrorRate = 150, utgBubblePopping = 1, and useBogUnitig = 0). Newbler assemblies resulted in a larger number of contigs (as well as longer) than those generated with Celera, and were used for subsequent phylogenetic and functional analysis.

Phylogenetic and Functional Analysis

Phylogenetic analysis of long-fragment 16S rRNA gene sequences (1200–1450 bp) was accomplished using blastn to identify closest neighbors in Genbank, and by construction of phylogenetic trees compared to known reference organisms. Phylogenetic trees of long-fragment 16S sequences from the domain *Archaea* were prepared using neighbor-joining and maximum parsimony methods (1000 bootstraps). Phylogenetic trees of bacterial 16S sequences are not shown, due primarily to the extensive diversity of different phylotypes present in sediment samples, and the predominance of *Sulfurihydrogenibium*-like organisms in "streamer" samples from Inflated Plain (IP), which are ∼99% identical (16S rRNA gene) to populations also found in terrestrial sites of YNP (Reysenbach et al., 2005; Fouke, 2011; Takacs-Vesbach et al., 2013). Classification of short-fragment 16S rRNA pyro-tag sequences (n ranged from ∼14,000 to 32,000 sequences per site) was performed (March 18, 2013) using the Ribosomal Database Project (RDP) Naïve Bayesian rRNA Classifier (Version 2.5, Bayesian RDP training set 9, May 2012).

Random metagenome sequence reads (∼360–400 bp) obtained from four vent communities were classified using blastx and sorted by G + C (%) content. Assembled environmental sequence data was also classified using blastx and screened for specific functional genes corresponding to known pathways in material and energy transfer. Query DNA sequences known to code for proteins important in the oxidation of reduced chemical constituents or the reduction of a terminal acceptor were used to search (WU-tblastn) the assembled metagenome sequence (gene list identified in Inskeep et al., 2010, 2013a). Environmental sequence fragments exhibiting homology ($E < 10^{-10}$) to query sequences were then reanalyzed using NCBI-blastp against the nr database. Positive functional gene hits were considered when (i) the gene fragment length relative to query length was >0.5, (ii) the phylogenetic identity was confirmed to match the primary population types observed using other sequencing protocols, and (iii) the genes have been described previously in similar phylotypes with closed genomes. There is no guarantee that all functional genes are ancestral; however, our criteria reports only those genes which are on contigs with phylogenetic consistency. Two 'streamer' samples from Inflated Plain (348S, 359S) yielded significant consensus sequence of the predominant *Sulfurihydrogenibium* populations present; this phylotype(s) represented >70% of the sequence reads in these samples and generated significant assembled sequence (∼25 × coverage).

Two of the four vent biomass samples did not produce sufficient assemblies to generate adequate coverage of all major phylotypes (especially Mary Bay sediments).

Sequence Data

Individual sequence reads and assembled contigs from four random metagenome datasets of vent communities (348S and 359S from IP; 349S from MB; 369S from WT) are available under the National Center for Biotechnology Information (NCBI) BioProject PRJNA60433. Long-fragment 16S rRNA gene sequences are deposited with GenBank (NCBI) under Accession Numbers KT453543 – KT453636 (file SUB1068923).

Acknowledgments

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.01044

Table S1 | Concentrations of all major geochemical constituents measured in sublacustrine thermal vent waters (and lake water from the Southeast Arm) obtained from Yellowstone Lake during September 2007 and 2008.

Table S2 | Number of sequence reads, read lengths, and assembly (Newbler) statistics of random shotgun sequencing performed on four thermal vent biomass samples collected in 2008.

Table S3 | Summary of 16S rRNA gene sequences observed in assembled metagenome sequence data from four sublacustrine thermal vent samples in Yellowstone Lake. The two samples from Inflated Plain were sampled approximately 300 m apart and exhibited highly similar geochemical signatures (**Table 1**).

Figure S1 | The remotely operated vehicle (ROV) deployed from the *Cutthroat.*

Figure S2 | Biomass and elemental sulfur collected from a thermal vent streamer community at Inflated Plain (359S) using the ROV-mounted sampling chamber (September 2008).

Figure S3–S7 | Video clips of several thermal vent sites sampled in Yellowstone Lake (2007-08) using a remotely operated vehicle (ROV). Includes 6 files (*.wmv): Figure S3 = IP 329 9_9 _2007; Figure S4 = IP 348 9_11_2008; Figure S5 = WT OV 333 9_12_2007; Figure S6 = WT Deep 339 9_18_2007; Figure S7 = MB 349 9_12_2008.

References

Aguiar, P., Beveridge, T. J., and Reysenbach, A.-L. (2004). Sulfurihydrogenibium azorense, sp. nov., a thermophilic hydrogen-oxidizing microaerophile from terrestrial hot springs in the Azores*Int. J. Syst. Evol. Microbiol.* 54, 33–39. doi: 10.1099/ijs.0.02790-0

Allison, J. D., Brown, D. S., and Novo-Gradac, K. J. (1991). *MINTEQ A2/PRODEFA2, a Geochemical Assessment Model for Environmental Systems: Version 3.0 User's Manual.* Athens: Environmental Research Laboratory, Office of Research and Development, USEPA.

Amend, J. P., and Shock, E. L. (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic *Archaea* and *Bacteria. FEMS Microbiol. Rev.* 25, 175–243. doi: 10.1111/j.1574-6976.2001.tb00576.x

APHA (1998). *Standard Methods for the Examination of Water and Wastewater.* Washington, DC: American Public Health Association.

Auchtung, T. A., Shyndriayeva, G., and Cavanaugh, C. M. (2011). 16S rRNA phylogenetic analysis and quantification of Korarchaeota indigenous to the hot springs of Kamchatka, Russia. *Extremophiles* 15, 105–116. doi: 10.1007/s00792-010-0340-5

Balistrieri, L. I., Shanks, W. C., Cuhel, R. L., Aguilar, C., and Klump, J. V. (2007). *The Influence of Sublacustrine Hydrothermal Vent Fluids on the Geochemistry of Yellowstone Lake,* ed L. A. Morgan (Boulder, CO: United States Geological Survey).

Beam, J. P., Jay, Z. J., Kozubal, M. A., and Inskeep, W. P. (2014). Novel thermoacidophilic Thaumarchaeota inhabit iron oxide and elemental sulfur sediment habitats of Yellowstone National Park. *ISME J.* 8, 938–951. doi: 10.1038/ismej.2013.193

Beh, M., Strauss, G., Huber, R., Stetter, K.-O., and Fuchs, G. (1993). Enzymes of the reductive citric acid cycle in the autotrophic eubacterium *Aquifex pyrophilus* and in the archaebacterium *Thermoproteus neutrophilus. Arch. Microbiol.* 160, 306–311. doi: 10.1007/BF00292082

Berg, I. A., Kockelkorn, D., Buckel, W., and Fuchs, G. (2007). A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* 318, 1782–1786. doi: 10.1126/science.1149976

Berg, I. A., Ramos-Vera, W. H., Petri, A., Huber, H., and Fuchs, G. (2010). Study of the distribution of autotrophic CO_2 fixation cycles in Crenarchaeota. *Microbiology* 156:256–269. doi: 10.1099/mic.0.034298-0

Borisov, V. B., Gennis, R. B., Hemp, J., and Verkhovsky, M. I. (2011). The cytochrome *bd* respiratory oxygen reductases. *Biochim. Biophys. Acta* 1807, 1398–1413. doi: 10.1016/j.bbabio.2011.06.016

Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008). Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245–252. doi: 10.1038/nrmicro1852

Clingenpeel, S., Kan, J., Macur, R. E., Woyke, T., Lovalvo, D., Varley, J., et al. (2013). Yellowstone Lake Nanoarchaeota. *Front. Microbiol.* 4:274. doi: 10.3389/fmicb.2013.00274

Clingenpeel, S., Macur, R. E., Kan, J., Inskeep, W. P., Lovalvo, D., Varley, J., et al. (2011). Yellowstone Lake: high-energy geochemistry and rich bacterial diversity. *Environ. Microbiol.* 13, 2172–2185. doi: 10.1111/j.1462-2920.2011.02466.x

Coumou, D., Driesner, T., and Heinrich, C. A. (2008). The structure and dynamics of mid-ocean ridge hydrothermal systems. *Science* 321, 1825–1828. doi: 10.1126/science.1159582

Fouke, B. W. (2011). Hot-spring systems geobiology: abiotic and biotic influences on travertine formation at Mammoth Hot Springs, Yellowstone National Park, USA. *Sedimentology* 58, 170–219. doi: 10.1111/j.1365-3091.2010.01209

Friedrich, C. G., Bardischewsky, F., Rother, D., Quentmeier, A., and Fischer, J. (2005). Prokaryotic sulfur oxidation. *Curr. Opin. Microbiol.* 8, 253–259. doi: 10.1016/j.mib.2005.04.005

García-Horsman, J. A., Barquera, B., Rumbley, J., Ma, J., and Gennis, R. B. (1994). The superfamily of heme-copper respiratory oxidases. *J. Bacteriol.* 176, 5587.

Hamamura, N., Meneghin, J., and Reysenbach, A.-L. (2013).. Comparative community gene expression analysis of Aquificales-dominated geothermal springs. *Environ. Microbiol.* 15, 1226–1237. doi: 10.1111/1462-2920.12061

Han, C. O., Kotsyurbenko, O., Chertkov, B., Held, A., Lapidus, M., Nolan, M., et al. (2012). Complete genome sequence of the sulfur compounds

oxidizing chemolithoautotroph *Sulfuricurvum kujiense* type strain (YK-1T). *Stand. Genomic Sci.* 6, 94–103. doi: 10.4056/sigs.2456004

Harmer, T. L., Rotjan, R. D., Nussbaumer, A. D., Bright, M., Ng, A. W., DeChaine, E. G., et al. (2008). Free-living tube worm endosymbionts found at deep-sea vents. *Appl. Environ. Microbiol.* 74, 3895–3898. doi: 10.1128/AEM.02470-07

Hatzenpichler, R. (2012). Diversity, physiology, and niche differentiation of ammonia-oxidizing Archaea. *Appl. Environ. Microbiol.* 78, 7501–7510. doi: 10.1128/AEM.01960-12

Hirayama, H., Yohey, S., Abe, M., Miyazaki, M., Makita, H., Inagaki, F., et al. (2011). Methylothermus subterraneus sp. nov., a moderately thermophilic methanotroph isolated from a terrestrial subsurface hot aquifer. *Int. J. Syst. Evol. Microbiol.* 61, 2646–2653. doi: 10.1099/ijs.0.028092-0

Hügler, M., Huber, H., Molyneaux, S. J., Vetriani, C., and Sievert, S. M. (2007). Autotrophic CO_2 fixation via the reductive tricarboxylic acid cycle in different lineages within the phylum Aquificae: evidence for two ways of citrate cleavage. *Environ. Microbiol.* 9, 81–92. doi: 10.1111/j.1462-2920.2006.01118.x

Inskeep, W. P., Ackerman, G. G., Taylor, W. P., Korf, S., Kozubal, M. A., and Macur, R. E. (2005). On the energetics of chemolithotropy in nonequilibrium systems: case studies of geothermal springs in Yellowstone National Park, *Geobiology* 3, 297–317. doi: 10.1111/j.1472-4669.2006.00059.x

Inskeep, W. P., Jay, Z. J., Herrgard, M. J., Kozubal, M. A., Rusch, D. B., Tringe, S. G., et al. (2013b). Phylogenetic and functional analysis of metagenome sequence from high-temperature archaeal habitats demonstrate linkages between metabolic potential and geochemistry. *Front. Microbiol.* 4:95. doi: 10.3389/fmicb.2013.00095

Inskeep, W. P., Jay, Z. J., Tringe, S. G., Herrgård, M., Rusch, D. B., YNP Metagenome Project Steering Committee and Working Group Members. (2013a). The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem. *Front. Microbiol.* 4:67. doi: 10.3389/fmicb.2013.00067

Inskeep, W. P., Rusch, D. B., Jay, Z. J., Herrgard, M. J., Kozubal, M. A., Richardson, T. H., et al. (2010). Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. *PLoS ONE* 5:e9773. doi: 10.1371/journal.pone.0009773

Ito, T., Sugita, K., Yumoto, I., Nodasaka, Y., and Okabe, S. (2005). Thiovirga sulfuroxydans gen. nov., sp. nov., a chemolithooautotrophic sulfur-oxidizing bacterium isolated from a microaerobic waste-water biofilm. *Int. J. Syst. Evol. Microbiol.* 55, 59–64. doi: 10.1099/ijs.0.63467-0

Jay, Z. J., Beam, J. P., Dohnalkova, A., Lohmayer, R., Bodle, B., Planer-Friedrich, B., et al. (2015). Pyrobaculum yellowstonensis strain WP30 respires on elemental sulfur and/or arsenate in circumneutral sulfidic geothermal sediments of Yellowstone National Park. *Appl. Environ. Microb.* 81, 5907–5916. doi: 10.1128/AEM.01095-15

Jay, Z. J., Rusch, D. B., Tringe, S. G., Bailey, C., Jennings, R. M., and Inskeep, W. P. (2014). Metabolic potential of Acidilobus-like populations from different geothermal systems in Yellowstone National Park reveal a common functional role. *Appl. Environ. Microbiol.* 80, 294–305. doi: 10.1128/AEM.02860-13

Johnson, S. Y., Stephenson, W. J., Morgan, L. A., Shanks, W. C., III, and Pierce, K. L. (2003). Hydrothermal and tectonic activity in northern Yellowstone Lake, Wyoming, *Geological Society of America Bulletin*, 115, 954–971. doi: 10.1130/B25111.1

Jünemann, S. (1997). Cytochrome bd terminal oxidase. *Biochim. Biophys. Acta* 1321, 107–127. doi: 10.1016/S0005-2728(97)00046-7

Kan, J., Clingenpeel, S., Macur, R. E., Inskeep, W. P., Lovalvo, D., Varley, J., et al. (2011). Archaea in Yellowstone Lake. *Int. Soc. Microbial. Ecol. J.* 5, 1784–1795. doi: 10.1038/ismej.2011.56

Klatt, C. G., Inskeep, W. P., Herrgard, M. J., Jay, Z. J., Rusch, D. B., Tringe, S. G., et al. (2013). Community structure and function of high-temperature chlorophototrophic microbial mats inhabiting diverse geothermal environments. *Front. Microbiol.* 4:106. doi: 10.3389/fmicb.2013.00106

Lovalvo, D., Clingenpeel, S. R., McGinnis, S., Macur, R. E., Varley, J. D., Inskeep, W. P., et al. (2010). A geothermal-linked biological oasis in Yellowstone Lake, Yellowstone National Park, Wyoming, *Geobiology* 8, 327–336. doi: 10.1111/j.1472-4669.2010.00244.x

Macur, R. E., Jay, Z. J., Taylor, W. P., Kozubal, M. A., Kocar, B. D., and Inskeep, W. P. (2013). Microbial community structure and sulfur biogeochemistry in mildly-acidic geothermal springs in Yellowstone National Park. *Geobiology* 11, 86–99. doi: 10.1111/gbi.12015

Miller-Coleman, R. L., Dodsworth, J. A., Ross, C. A., Shock, E. L., Williams, A. J., Hartnett, H. E., et al. (2012). Korarchaeota diversity, biogeography, and abundance in Yellowstone and Great Basin Hot Springs and ecological niche modeling based on machine learning. *PLoS ONE* 7:e35964. doi: 10.1371/journal. pone.0035964

Morgan, L. A., Shanks, W. C., III, Lovalvo, D. A., Johnson, S. Y., Stephenson, W., Pierce, K. L., et al. (2003). Exploration and discovery in Yellowstone Lake: results from high-resolution sonar imaging, seismic reflection profiling, and submersible studies. *J. Volcanol. Geothermal. Res.* 122, 221–242. doi: 10.1016/S0377-0273(02)00503-6

Morgan, L. A., Shanks, W. C., III, Pierce, K. L., Lovalvo, D. A., Lee, G. K., Webring, M. W., et al. (2007). "The floor of Yellowstone Lake is anything but quiet - new discoveries from high resolution sonar imaging, seismic-reflection profiling, and submersible studies," in *Integrated Geoscience Studies in the Greater Yellowstone Area - Volcanic, Tectonic, and Hydrothermal. Processes in the Yellowstone Geoecosystem*, Vol. 1717, ed L. A. Morgan (Boulder, CO: US Geological Survey Professional Paper), 95–126.

Morgan, L. A., Shanks, W. C. P., III, and Pierce, K. L. (2009). Hydrothermal processes above a large magma chamber: large hydrothermal systems and hydrothermal explosions in yellowstone national park. *Geol. Soc. Am. Spec. Pap.* 459, 1–95. doi: 10.1130/2009.2459

Morgan, L. A., and Shanks, W. C. III. (2005). "Influences of rhyolitic lava flows on hydrothermal processes in Yellowstone Lake and on the Yellowstone Plateau," in *Geothermal Biology and Geochemistry in Yellowstone National Park,* eds W. P. Inskeep and T. R. McDermott (Bozeman, MT: Montana State University Press), 32–52.

Morgan, P., Blackwell, D. D., Spafford, R. E., and Smith, R. B. (1977). Heat flow measurements in Yellowstone Lake and the thermal structure of the Yellowstone Caldera. *J. Geophys. Res.* 82, 3719–3732. doi: 10.1029/JB082i026p03719

Mori, K., Yamaguchi, K., Sakiyama, Y., Urabe, T., and 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. *Int. J. Syst. Evolut. Microbiol.* 59, 2894–2898. doi: 10.1099/ijs.0.010033-0

Nakagawa, S., Shtaih, Z., Banta, A., Beveridge, T. J., Sako, Y., and 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. *Int. J. Syst. Evol. Microbiol.* 55, 2263–2268. doi: 10.1099/ijs.0.63708-0

Nunoura, T., Takaki, Y., Kakuta, J., Nishi, S., Sugahara, J., Kazama, H., et al. (2011). Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res.* 39, 3204–3223. doi: 10.1093/nar/gkq1228

Reysenbach, A. L., Banta, A. B., Boone, D. R., Cary, S. C., and Luther, G. W. (2000). Microbial essentials at hydrothermal vents. *Nature* 404, 835. doi: 10.1038/35009029

Reysenbach, A.-L., Banta, A., Civello, S., Daly, J., Mitchell, K., Lalonde, S., et al. (2005). "The aquificales of Yellowstone National Park," in *Geothermal Biology and Geochemistry in Yellowstone National Park*, eds W. P. Inskeep, and T. R. McDermott (Bozeman, MT: Montana State University Press), 129–142.

Reysenbach, A. L., Hamamura, N., Podar, M., Griffiths, E., Ferreira, S., Hochstein, R., et al. (2009). Complete and draft genome sequences of six members of the Aquificales. *J. Bacteriol.* 191, 1992–1993. doi: 10.1128/JB.01645-08

Setoguchi, Y., Nomaki, H., Kitahashi, T., Watanabe, H., Inoue, K., Ogawa, N. O., et al. (2014). Nematode community composition in hydrothermal vent and adjacent non-vent fields around Myojin Knoll, a seamount on the Izu-Ogasawara Arc in the western North Pacific Ocean. *Mar. Biol.* 161:2460. doi: 10.1007/s00227-014-2460-4

Shanks, W. C. III., Alt, J. C., and Morgan, L. A. (2007). "Geochemistry of sublacustrine hydrothermal deposits in Yellowstone Lake—hydrothermal reactions, stable isotope systematics, sinter deposition, and spire growth," in *Integrated Geoscience Studies in the Greater Yellowstone Area— Volcanic Hydrothermal and Tectonic Processes in the Yellowstone Geoecosystem,*

Vol. 1717, ed L. A. Morgan (Boulder, CO: US Geological Survey Professional Paper), 205–234.

Shanks, W. C. III., Morgan, L. A., Balistrieri, L., and Alt, J. (2005). "Hydrothermal vent fluids, siliceous hydrothermal deposits, and hydrothermally altered sediments in Yellowstone Lake," in *Geothermal Biology and Geochemistry in Yellowstone National Park,* eds W. P. Inskeep and T. R. McDermott (Bozeman, MT, Montana State University Press), 53–72.

Spear, J. R., Walker, J. J., McCollom, T. M., and Pace, N. R. (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2555–2560. doi: 10.1073/pnas.0409574102

Swingley, W. D., Meyer-Dombard, D. R., Shock, E. L., Alsop, E. B., Falenski, H. D., Havig, J. R., et al. (2012). Coordinating environmental genomics and geochemistry reveals metabolic transitions in a hot spring ecosystem. *PLoS ONE* 7:e38108. doi: 10.1371/journal.pone.0038108

Takacs-Vesbach, C., Inskeep, W. P., Jay, Z. J., Herrgard, M. J., Rusch, D. B., Tringe, S. G., et al. (2013). Metagenome sequence analysis of filamentous microbial communities obtained from geochemically distinct geothermal channels reveals specialization of three Aquificales lineages. *Front. Microbiol.* 4:84. doi: 10.3389/fmicb.2013.00084

To, T. B., Nordstrom, D. K., Cunningham, K. M., Ball, J. W., and McCleskey, R. B. (1999). New method for the direct determination of dissolved Fe(III) concentration in acid mine waters. *Environ. Sci. Technol.* 33, 807–813.

Tsubota, J., Eshinimaev, B., Khmelenina, V. N., and Trotsenko, Y. A. (2005). Methylothermus thermalis gen. nov., sp. nov., a novel moderately thermophilic obligate methanotroph from a hot spring in Japan. *Int. J. Syst. Evol. Microbiol.* 55, 1877–1884. doi: 10.1099/ijs.0.63691-0

Van Dover, C. L., Humphris, S. E., Fornari, D., Cavanaugh, C. M., Collier, R., Goffredi, S. K., et al. (2001). Biogeography and ecological setting of Indian Ocean hydrothermal vents. *Science* 294, 818–823. doi: 10.1126/science.1064574

Van Dover, C. L., Ward, M. E., Scott, J. L., Underdown, J., Anderson, B., Gustafson, C., et al. (2007). A fungal epizootic In mussels at a deep-sea hydrothermal vent. *Marine Ecol.* 28, 54–62. doi: 10.1111/j.1439-0485.2006.00121.x

Viginais, P. M., and Billoud, B. (2007). Occurrence, classification and biological function of hydrogenases. *Chem. Rev.* 107, 4206–4272. doi: 10.1021/cr050196r

Ward, N., Larsen, Ø., Sakwa, J., Bruseth, L., Khouri, H., Durkin, A. S., et al. (2004). Genomic insights into methanotrophy: the complete genome sequence of Methylococcus capsulatus(Bath). *PLoS Biol.* 2:e303. doi: 10.1371/journal.pbio.0020303

Wold, R. J., Mayhew, M. A., and Smith, R. B. (1977). Geophysical study of a hydrothermal explosion crater in Mary Bay, Yellowstone Lake, Wyoming, *J. Geophys. Res.* 82, 3733–3738. doi: 10.1029/JB082i026p03733

Yamada, T., Imachi, H., Ohashi, A., Harada, H., Hanada, S., Kamagata, Y., et al. (2007). Bellilinea caldifistulae gen. nov., sp. nov. and Longilinea arvoryzae gen. nov., sp. nov., strictly anaerobic, filamentous bacteria of the phylum Chloroflexi isolated from methanogenic propionate-degrading consortia. *Int. J. System. Evolut. Microbiol.* 57, 2299–2306. doi: 10.1099/ijs.0.65098-0

Yang, T., Lyons, S., Aguilar, C., Cuhel, R., and Teske, A. (2011). Microbial communities and chemosynthesis in Yellowstone Lake sublacustrine hydrothermal vent waters. *Front. Microbiol.* 2:130. doi: 10.3389/fmicb.2011.00130

Community structure and function of high-temperature chlorophototrophic microbial mats inhabiting diverse geothermal environments

Christian G. Klatt[1,2†], **William P. Inskeep**[1,2]*, **Markus J. Herrgard**[3], **Zackary J. Jay**[1,2], **Douglas B. Rusch**[4], **Susannah G. Tringe**[5], **M. Niki Parenteau**[6,7], **David M. Ward**[1,2], **Sarah M. Boomer**[8], **Donald A. Bryant**[9,10] and **Scott R. Miller**[11]

[1] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA
[2] Thermal Biology Institute, Montana State University, Bozeman, MT, USA
[3] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark
[4] Center for Genomics and Bioinformatics, Indiana University, Bloomington, IN, USA
[5] Department of Energy Joint Genome Institute, Walnut Creek, CA, USA
[6] Search for Extraterrestrial Intelligence Institute, Mountain View, CA, USA
[7] National Aeronautics and Space Administration Ames Research Center, Mountain View, CA, USA
[8] Western Oregon University, Monmouth, OR, USA
[9] Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA
[10] Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA
[11] Department of Biological Sciences, University of Montana, Missoula, MT, USA

Six phototrophic microbial mat communities from different geothermal springs (YNP) were studied using metagenome sequencing and geochemical analyses. The primary goals of this work were to determine differences in community composition of high-temperature phototrophic mats distributed across the Yellowstone geothermal ecosystem, and to identify metabolic attributes of predominant organisms present in these communities that may correlate with environmental attributes important in niche differentiation. Random shotgun metagenome sequences from six phototrophic communities (average ~53 Mbp/site) were subjected to multiple taxonomic, phylogenetic, and functional analyses. All methods, including G + C content distribution, MEGAN analyses, and oligonucleotide frequency-based clustering, provided strong support for the dominant community members present in each site. Cyanobacteria were only observed in non-sulfidic sites; *de novo* assemblies were obtained for *Synechococcus*-like populations at *Chocolate Pots* (CP_7) and *Fischerella*-like populations at *White Creek* (WC_6). Chloroflexi-like sequences (esp. *Roseiflexus* and/or *Chloroflexus* spp.) were observed in all six samples and contained genes involved in bacteriochlorophyll biosynthesis and the 3-hydroxypropionate carbon fixation pathway. Other major sequence assemblies were obtained for a Chlorobiales population from CP_7 (proposed family *Thermochlorobacteriaceae*), and an anoxygenic, sulfur-oxidizing *Thermochromatium*-like (Gamma-proteobacteria) population from *Bath Lake Vista Annex* (BLVA_20). Additional sequence coverage is necessary to establish more complete assemblies of other novel bacteria in these sites (e.g., Bacteroidetes and Firmicutes); however, current assemblies suggested that several of these organisms play important roles in heterotrophic and fermentative metabolisms. Definitive linkages were established between several of the dominant phylotypes present in these habitats and important functional processes such as photosynthesis, carbon fixation, sulfur oxidation, and fermentation.

Keywords: microbial mats, microbial interactions, phototrophic bacteria, functional genomics, thermophilic bacteria

INTRODUCTION

Many naturally occurring microorganisms have eluded isolation, due in part to a poor understanding of the chemical, physical, and biotic factors defining their realized niches (Rappé and Giovannoni, 2003). Moreover, much of the sequence diversity revealed by amplification of specific gene targets (e.g., 16S rRNA) is susceptible to biases inherent in primer-design and PCR protocols.

Random shotgun sequencing of environmental DNA provides a direct and potentially less biased view of the composition and functional attributes of microbial communities. For example, three new chlorophototrophic organisms (i.e., organisms capable of (bacterio)chlorophyll-based phototrophy) were discovered in prior metagenome analyses of oxygenic mats in YNP, two of which lie outside the clades of known phototrophic organisms

in the Chlorobiales and Chloroflexi (Klatt et al., 2011). Moreover, the third organism, "*Candidatus* Chloracidobacterium thermophilum" ("*Ca.* C. thermophilum"), represents the only known occurrence of chlorophototrophy in the phylum Acidobacteria (Bryant et al., 2007; Klatt et al., 2011; Garcia Costas et al., 2012). Metagenome sequencing and subsequent bioinformatic analyses provide an opportunity to identify the metabolic attributes of uncultivated organisms that can be used to postulate detailed biochemical linkages among individual community members necessary for the development of computational models describing microbial interaction and community function (Taffs et al., 2009).

High-temperature phototrophic microbial mats have served as models for studying microbial community structure and function. Studies have included investigations of microbial community composition (Miller et al., 2009), the ecophysiology of novel isolates (Pierson and Castenholz, 1974; Bryant et al., 2007; van der Meer et al., 2010), comparative genomics, metagenomics, and metatranscriptomics (Bhaya et al., 2007; Klatt et al., 2007, 2011; Becraft et al., 2011; Liu et al., 2011, 2012; Melendrez et al., 2011), community network modeling (Taffs et al., 2009), phage-host interactions (Heidelberg et al., 2009), as well as theoretical models of evolution (Ward et al., 2008). The high temperature and relative geochemical stability of geothermal phototrophic mats in YNP generally result in communities with several dominant phylotypes and have provided opportunities for understanding environmental factors controlling community composition (Brock, 1978; Cohen and Rosenberg, 1989; Ward and Castenholz, 2000; Ward et al., 2012). Prior investigations have revealed that temperature, pH, and sulfide are among the most important environmental variables dictating differences in phototrophic mat community structure (Castenholz, 1976, 1977; Castenholz and Pierson, 1995; Madigan et al., 2005; Cox et al., 2011; Boyd et al., 2012). The presence of sulfide was used in the current study to separate anoxygenic versus oxygenic communities common in YNP (Inskeep et al., 2013). Oxygenic and/or anoxygenic photoautotrophs are generally the predominant primary producers in geothermal mats at temperatures of ~50–72°C and moderately acidic to alkaline pH (5–9). These mat communities support a diverse array of (photo-)heterotrophic, fermentative, sulfate-respiring, and methanogenic organisms, whose physiological attributes are critical for understanding community function (Zeikus and Wolfe, 1972; Jackson et al., 1973; Henry et al., 1994; Nold and Ward, 1996; Ward et al., 1998; Taffs et al., 2009; Klatt et al., 2011; Liu et al., 2012).

The distribution of different chlorophototrophic bacteria is often controlled by specific geochemical parameters. For example, members of the Cyanobacteria are not generally found in acidic or sulfidic environments (Castenholz, 1976, 1977). However, filamentous anoxygenic phototrophs (FAPs) of the phylum Chloroflexi exhibit a wider habitat range than other chlorophototrophs. Closely related members of the Chloroflexi [>97% nucleotide identity (NT ID) of the 16S rRNA gene] with different phenotypes have been cultured from geothermal environments (Madigan et al., 1974; Madigan and Brock, 1975). FAPs isolated from a high-sulfide (>100 μM) spring in the absence of cyanobacteria (*Chloroflexus* sp. GCF strains) fixed inorganic carbon using sulfide as the electron donor (Giovannoni et al., 1987). However, most other cultured *Chloroflexus* spp. from low-sulfide environments

are photoheterotrophic and do not utilize reduced sulfur for photosynthesis (Madigan et al., 1974; Pierson and Castenholz, 1974). Natural populations of FAPs are known to consume organic compounds produced by cyanobacterial community members (van der Meer et al., 2005); however, genomic and biochemical evidence is needed to improve our understanding of how different populations of Chloroflexi function *in situ*.

The overall goal of this study was to investigate the underlying environmental factors and potential physiological adaptations important in defining the microbial community structure and function of different types of chlorophototrophic mats commonly found in association with certain geothermal features of YNP (Inskeep et al., 2013). The specific objectives of this study were to (i) utilize metagenome sequencing and bioinformatic analyses to determine the community composition of thermal chlorophototrophic mats in YNP, (ii) identify key metabolic attributes of the major chlorophototrophic organisms present in these communities, and (iii) evaluate the predominant environmental and/or geochemical attributes that contribute to niche differentiation of thermophilic chlorophototrophic communities. The habitats sampled in the current study were chosen to focus on several of the major high-temperature phototrophic mat types that are distributed across the YNP geothermal ecosystem.

RESULTS
GEOCHEMICAL AND PHYSICAL CONTEXT

The predominant differences among the six phototrophic microbial mat communities included both geochemical characteristics such as pH and dissolved sulfide (DS), as well as temperature, and the sample depth (**Figure 1**; **Table 1**). Temperature ranged from 40–60°C across these six sites, and is a critical parameter controlling community composition. Four of the geothermal sites contained no measurable DS, while both samples from *Bath Lake Vista Annex Spring* (BLVA_5 and BLVA_20) were collected from hypoxic sulfidic environments (total DS ~117 μM). Although the dissolved oxygen content at the source of *Chocolate Pots* (near sample location CP_7) was below detection (<1 μM), this spring contained no sulfide and high concentrations of Fe (II) (~76 μM) (**Table 1**), which results in the precipitation of Fe(III)-oxides upon discharge and reaction with oxygen (Trouwborst et al., 2007). The phototrophic mat obtained from *White Creek* (WC_6) occurs within an oxygenated, alkaline-siliceous geothermal drainage channel containing no detectable DS (**Table 1**). The site was included in the study to target a population of the heterocyst-forming cyanobacterium *Fischerella* (*Mastigocladus*) *laminosus* that has been the focus of prior work at this location (Miller et al., 2006, 2007, 2009).

Samples from *Mushroom Spring* (MS_15) and *Fairy Geyser* (FG_16) were obtained from laminated phototrophic mats after removal of the top layer (See Materials and Methods). Dissection of these mats was performed to focus on FAPs, which were known to occur in higher abundance at greater depths below a surface layer dominated by cyanobacteria (Boomer et al., 2002; Nübel et al., 2002). The phototrophic mats at FG_16 are referred to as "splash-mats" due to the fact that these communities receive frequent inputs of geothermal water emanating from the main source

FIGURE 1 | Site photographs of phototrophic microbial mats selected for metagenome sequencing. The sites cover a range in geochemical conditions including (i) highly sulfidic environments at *Bath Lake Vista Annex* (BLVA_5, 20), (ii) oxygenic phototrophic communities at *White Creek* (WC_6) and *Chocolate Pots* (CP_7), and (iii) subsurface mat layers at *Mushroom Spring* (MS_15) and *Fairy Geyser* (FG_16) (also oxygenic systems). The anoxygenic phototrophic communities at *Bath Lake Vista Annex* (BLVA) were sampled at two different time points (Table S2 in Inskeep et al., 2013) to compare *Chloroflexus* mats in the absence (BLVA_5) and presence (BLVA_20) of purple-bacteria (Arrows indicate approximate sample locations and types; inset at BLVA_5 shows mat dissection at sampling).

Table 1 | Sample locations and aqueous geochemical parameters[1] of six, high-temperature phototrophic microbial communities sampled in Yellowstone National Park (YNP) and used for metagenome sequencing.

Location	T (°C)	pH	Na^+	Cl^-	SO_4^{2-}	[1]DIC	[1]DS	[1]DO	[1]DOC	As	[2]Fe	[3]NH_4^+	Coordinates
					mM					μM			
Bath Lake Vista Annex-Green (BVLA_5)	57	6.2	3.9	4.4	5.6	15.8	117	<3	104	24	0.7	40	44° 57′ 54.180″ N 110° 42′ 42.228″ W
Bath Lake Vista Annex-Purple (BLVA_20)	54	6.2	5.5	5.7	7.3	24.2	117	<3	75	23	0.7	40	44° 57′ 54.180″ N 110° 42′ 42.228″ W
White Creek (WC_6)	52	8.2	3.6	1.8	0.23	nd	<3	188	nd	5	1.7	1.9	44° 31′ 53.399″ N 110° 47′ 51.799″ W
Chocolate Pots (CP_7)	52	6.2	4.1	0.89	0.23	13.2	<3	<3	38	9	75.5	4.2	44° 42′ 36.288″ N 110° 44′ 28.824″ W
Mushroom Spring (MS_15)	60	8.2	12.6	7.3	0.18	2.1	<3	141	nd	26	<1	4.4	44° 32′ 19.284″ N 110° 47′ 52.692″ W
Fairy Geyser (FG_16)	36–38	9.1	9.4	5.2	0.18	4.8	<3	31	30	13	<1	1.3	44° 32′ 31.812″ N 110° 51′ 40.788″ W
[4]Correlation (r^2)		0.89*		0.93*	0.96*		0.99***				0.72**	0.99**	

[1]DS, total dissolved sulfide; DO, dissolved oxygen; DIC, dissolved inorganic carbon; DOC, dissolved organic carbon.

[2]Mn (total soluble) values were also significant in CP_7 (24 μM) and WC_6 (5 μM), but low in other sites (0.1–0.2 μM, or below detection of 0.1 μM).

[3]Nitrate values ranged from 2.1–6.7 μM across sites.

[4]Correlation significance values: *p < 0.05, **p < 0.01, ***p < 0.001.

pool (85–88°C) (**Figure 1**). The "splash-mats" surrounding FG_16 are reasonably thick (~3–5 cm), and the sample discussed here was collected from a 2–4 mm "red-layer," found within a

temperature range of 35–50°C and a pH approaching 9 (Boomer et al., 2000, 2002). The visual characteristic of the "red-layer" was apparent during sampling and represents a different subsurface

environment than the sample obtained from MS_15. No measurable DS was present in the bulk aqueous phase (**Table 1**) of these mats; however, subsurface mats in these systems (MS_15 and FG_16) have been shown to be less oxic than their respective near-surface layers (Jensen et al., 2011).

ANALYSIS OF METAGENOME SEQUENCES

Individual sequences (average length ∼800 bp) were analyzed using several complementary approaches including alignment-based comparisons to reference databases, and evaluation of the guanine and cytosine content (% G + C) of each sequence read. In addition, comparison of all sequences to the NCBI nr database (blastx) was accomplished using MEGAN (Huson et al., 2007). The most highly represented phyla across all sites included the Chloroflexi (28%), Cyanobacteria (12%), Proteobacteria (8%), Bacteroidetes (6%), and Chlorobi (2%). Many sequence reads (27%) did not match those available in NCBI ("no hits"); this indicated that some members of these communities are not represented in current databases.

Taxonomic assignment of individual sequences was combined with %G + C distribution to obtain a profile of community composition (**Figure 2**). Each site contained populations similar to *Chloroflexus* and/or *Roseiflexus* spp., with average G + C contents of 55 and 61%, respectively. The two sulfidic samples (BLVA_5 and BLVA_20) showed contributions from both *Chloroflexus* and *Roseiflexus*-like populations (**Figure 2**). The oxic community from *White Creek* (WC_6) also contained significant contributions from *Chloroflexus*-like organisms, while CP_7, MS_15 and FG_16 were enriched in *Roseiflexus*-like sequences (**Figure 2**). All sites contain a significant number of sequences contributed from novel Chloroflexi that have not been adequately characterized, and for which appropriate reference organisms have not yet been cultivated or sequenced.

The phototrophic mat communities from WC_6 and CP_7 contained a significant fraction of sequences (23 and 25%,

respectively) contributed from members of the Cyanobacteria. Both sites contained sequences related to *Synechococcus* spp. strains A and B′ (mean G + C content of 60%; Bhaya et al., 2007) (**Figure 2**; **Figure A1** in Appendix), but the WC_6 community yielded a large proportion of Cyanobacteria-like sequences (73%) that could not be classified beyond the phylum-level, and these sequences exhibit a large range in G + C content (40–65%). *Fisherella laminosus* (order Stigonematales) has been shown to be an important community member at WC_6 (Miller et al., 2009), and many of the cyanobacterial sequences from WC_6 showed high sequence identity (95% average NT ID of alignments) to the draft genome of *Fischerella* sp. JSC-11 (average G + C = 41%; **Figure A2** in Appendix), which was the only representative genome available for this group of cyanobacteria (at time of writing). The G + C content frequency plots also revealed major contributions from organisms within the Chlorobi (at sites CP_7 and FG_16), *Thermotoga* (MS_15), and *Thermochromatium* spp. (purple-sulfur bacteria) in BLVA_20 with an average G + C content of 64%. Moreover, all sites contained bacterial sequences that could not be identified beyond the level of Domain *Bacteria* (especially G + C contents ranging from 20–40%, **Figure 2**), in part because appropriate reference genomes are not currently available, and significant assemblies were not obtained for phylotypes present in lower abundance.

ANALYSIS OF METAGENOME ASSEMBLIES

The assembly of individual sequence reads into contigs and scaffolds is a powerful method for linking functional attributes with specific phylotypes. Assembly yielded scaffolds ranging from 1 kb (small contigs) to nearly 126 kb (largest scaffold), and an average scaffold size of 2,330 bp across all six sites. Community structure plays a role in the degree of assembly and the ability to obtain large scaffolds; communities with larger proportions of metagenome sequence originating from fewer, more dominant

FIGURE 2 | Percent G + C content and taxonomic analysis of random shotgun sequence reads obtained from six thermophilic phototrophic mat communities from Yellowstone National Park (YNP). The frequency plot of all sequence reads (black) versus G + C content (%) is shown with corresponding taxonomic analysis (MEGAN-"blastx") as indicated by the color key (right).

FIGURE 3 | Principal components analysis of oligonucleotide frequencies of assembled sequence from *Bath Lake Vista Annex*. BLVA_20 was sampled 8 months after BLVA_5 to capture a bloom of purple-sulfur bacteria shown in prior work to be related to *Thermochromatium tepidum* (Castenholz, 1977; Ward et al., 1989). Both sites contained scaffolds from dominant populations of *Chloroflexus* spp., Firmicutes, Nitrospira, and additional proteobacteria, but only BLVA_20 contained numerous scaffolds corresponding to the population of purple-sulfur bacteria (Gamma-proteobacteria, family *Chromatiaceae*, average G + C ∼64%) that is notably absent in BLVA_5.

organisms resulted in longer assemblies. Diversity metrics of PCR-based 16S rRNA sequences that were produced simultaneously from the same samples indicated that subsurface mat communities from MS_15 and FG_16 exhibited higher Simpson's diversity values (reported as the reciprocal of the Simpson's index, λ^{-1}; **Table A1** in Appendix). The greater degree of species "evenness" in MS_15 and FG_16 yielded considerably smaller assemblies, and only two scaffolds >10 kb were obtained from each of these two sites. Contrastingly, CP_7 exhibited the lowest Simpson's λ^{-1}, and the largest assemblies were obtained from this site, which contributed 42% of the large scaffolds (>10 kb) obtained across all six sites. Large assemblies were also obtained from the anoxygenic mats at BLVA (BLVA_5, _20), and these samples had similarly low values for Simpson's λ^{-1}.

NUCLEOTIDE WORD-FREQUENCY ANALYSIS OF DOMINANT POPULATIONS

Sequence assemblies were examined using principal components analysis (PCA) of nucleotide word frequencies (NWF) (Teeling et al., 2004) in conjunction with a taxonomic classification algorithm of average scaffold identity (APIS; Badger et al., 2006). For example, NWF PCA plots of the sulfidic system at BLVA sampled 8 months apart revealed major differences in community composition associated with a visible bloom of purple-sulfur bacteria in BLVA_20 (**Figures 1** and **3**). The major change in community composition between the two samples was the *Thermochromatium*-like population in BLVA_20, which corresponded with a decrease in *Roseiflexus*-like sequences (**Figure 3**). Both BLVA samples revealed a dominant *Chloroflexus*-like population that corresponded to the G + C peak at 55% (**Figure 2**). Similar NWF PCA analyses of assemblies from CP_7 revealed three predominant community members related to *Roseiflexus*, *Synechococcus*, and "*Candidatus* Thermochlorobacter aerophilum"-like organisms ("*Ca.* T.

aerophilum" represents a novel clade in the order Chlorobiales; Liu et al., 2012). Several other organisms were present in lower abundance and were distantly related to members of the Firmicutes, Bacteroidetes, and Spirochetes (**Figure A3** in Appendix). The large Chlorobi-like assemblies obtained from CP_7 were phylogenetically related (average NT ID = 91%) to "*Ca.* T. aerophilum" assemblies obtained from *Mushroom* and *Octopus Springs* metagenomes (Klatt et al., 2011; Liu et al., 2012). Translated PscD sequences from this newly described lineage of uncultivated Chlorobi are clearly distinct from other previously described phototrophic Chlorobi (PscD sequences from the CP_7 and *Mushroom* populations have 95% amino acid identity (AA ID) (**Figure A4** in Appendix).

A Monte-Carlo approach was also used to compare normalized oligonucleotide frequencies across the six phototrophic sites, which clustered the scaffolds of highly related organisms (e.g., genus/species level). A minimum scaffold length of 10 kbp was used to focus the analysis on dominant assemblies; consequently, smaller scaffolds from subsurface mat communities (MS_15 and FG_16) were not well represented in this analysis. Twelve scaffold clusters (consensus k-means groupings) were observed across sites (**Figure 4**; **Table 2**), and each of these populations corresponded with dominant community members identified using G + C content (%) and BLASTP assignments (**Figure 2**; **Figure A5** in Appendix). Clustering by oligonucleotide frequency afforded greater discrimination among populations that exhibited similar G + C content. For example, *Roseiflexus*-like organisms have similar G + C content (61%) to *Synechococcus* sp. strains A and B′ (**Figure 2**), yet these different genera are clearly separated based on differences in sequence character using oligonucleotide clustering analysis (**Figure 4**).

A sequence cluster corresponding to *Thermochromatium* spp. (Gamma-proteobacteria) contained sequences solely from

FIGURE 4 | Scaffold oligonucleotide frequency similarity network.
Oligonucleotide (tri-, tetra-, penta-, and hexa-nucleotide) counts were
normalized to scaffold length and subjected to k-means clustering ($k = 8$, 100
trials). The scaffolds that group together in $\geq 90\%$ trials are shown, with lines
connecting scaffolds ranging from blue (90%) to red (100%). The sample
origins of scaffolds shown here are indicated by site color (see legend) where
open circles correspond to reference genomes; scaffolds containing
phylogenetic or functional marker genes are indicated by larger nodes.

Table 2 | Properties of scaffold clusters obtained from metagenome assemblies as demarcated with oligonucleotide composition and confirmed using phylogenetic analyses.

Scaffold cluster	Taxonomic affiliation	Sites	No. of scaffolds	Median size (kbase)	G+C (%) ± SD	Total sequence (Mbase)	Depth of coverage (x)
1	*Roseiflexus* spp.	BLVA_5, CP_7, MS_15, FG_16	112	12.5	60.0 ± 1.2	1.55	2.6 ± 0.4
2	*Chloroflexus* spp.	BLVA_5, BLVA_20, WC_6, CP_7	211	13.5	54.3 ± 1.2	3.21	2.9 ± 0.7
3	*Ca*. Thermochlorobacter spp.	CP_7	73	14.8	49.5 ± 0.8	1.13	2.7 ± 0.5
4	*Thermochromatium* spp.	BLVA_20	29	12.5	63.0 ± 1.3	0.37	2.1 ± 0.4
5	*Synechococcus* spp.	WC_6, CP_7	78	26.2	58.7 ± 1.1	2.59	4.0 ± 0.7
6	Cyanobacteria	WC_6, CP_7	26	11.7	49.8 ± 1.2	0.32	2.4 ± 0.5
7	Bacteroidetes	WC_6	30	11.1	37.7 ± 0.9	0.37	2.4 ± 0.4
8	Chloroflexi-like	BLVA_5, MS_15, BLVA_20	37	10.6	63.9 ± 2.3	0.44	2.5 ± 0.5
9	Firmicute-like	BLVA_5, CP_7, BLVA_20	47	14.2	36.0 ± 1.5	0.79	2.7 ± 0.4
10	Firmicute-like	BVLA_5, BLVA_20	11	12.7	29.0 ± 1.4	0.16	2.6 ± 0.6
11	Spirochaetes	CP_7	21	11.8	30.5 ± 1.4	0.25	2.3 ± 0.4

Scaffold clusters 7–11 represent novel bacteria that are not well represented in public databases, and are currently defined at the phylum level.

BLVA_20, which is consistent with visual evidence of this population at the time of sampling (**Figure 1**), as well as further NWF PCA analysis using contigs >20 kb (**Figure A6** in Appendix). Other major sequence clusters identified included the "*Ca.* T. aerophilum"-like population from CP_7 (discussed above). Although relatives of the Bacteroidetes were found to occupy all sites, larger assemblies of several of these community members were obtained from WC_6. Three scaffold clusters with comparatively low G + C content (<40%) were observed, but neither AMPHORA (based on phylogenetic analysis) nor MEGAN ("blastx" alignments) could classify the sequences in these groups. This suggested that they originated from organisms that are currently poorly represented in public databases.

USE OF SINGLE-COPY GENES TO DEMARCATE DOMINANT POPULATIONS

Phylogenetically informative single-copy genes were identified among the metagenome assemblies using AMPHORA (Wu and Eisen, 2008), and provided yet another method for evaluating the predominant taxa represented in the six metagenomes. The distribution of dominant phylotypes predicted using AMPHORA (**Figure 5A**) was similar to that observed using the combined "blastx" and G + C (%) analyses of individual sequences (**Figure 2**), as well as to the taxonomic distribution of PCR-based 16S rRNA gene libraries from these same sites (**Figure 5B**). Moreover, the distribution of predominant populations (e.g., Chloroflexi, Cyanobacteria, Chlorobi, Proteobacteria) across sites was consistent with detailed analysis of major oligonucleotide clusters (e.g., **Figures 3** and **4**). All approaches showed that members of the Chloroflexi were ubiquitous across all sites. The relative contribution of *Chloroflexus* versus *Roseiflexus*-like organisms varied across different sites, and all sites contained novel organisms from undescribed lineages within the Chloroflexi (discussed in greater detail below). Other phototrophs detected in these sites included populations of Alpha-proteobacteria (Family
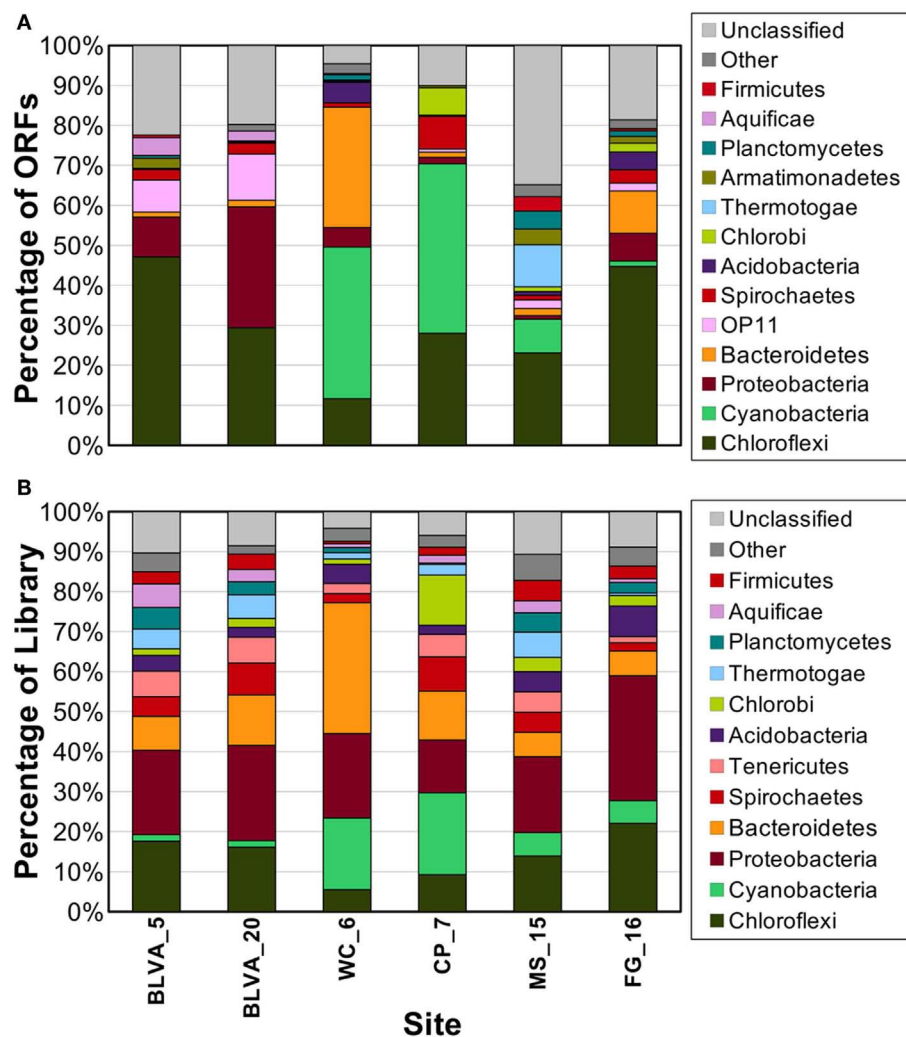

FIGURE 5 | Phylogenetic summary of marker genes from metagenome sequences compared to 16S rRNA gene sequences. Phylogenetic marker genes in the metagenome sequences were **(A)** assigned and classified using AMPHORA, and compared to **(B)** 16S rRNA sequences from ribosomal panels (*n* ∼ 300 per site) classified at the phylum-level against the RDP at a confidence threshold of 80%.

Hyphomicrobiaceae) in FG_16, "*Ca.* C. thermophilum" (phylum Acidobacteria) (Bryant et al., 2007) in WC_6, and "*Ca.* T. aerophilum"-like organisms (order Chlorobiales) in MS_15, FG_16 and especially CP_7 (**Figure 5B**). The MS_15 community contained a *Thermotoga*-like population as well as several low G + C organisms that have not yet been characterized. Although the subsurface mat community from FG_16 contained a novel high G + C proteobacterial population not seen in the other sites (**Figure 2**), these sequences could not be linked unambiguously to the *Hyphomicrobiaceae* 16S rRNA sequences described above, due to inadequate sequence coverage of this population and the lack of a good reference genome that would undoubtedly have assisted in sequence identification.

The distribution of phylogenetically unique Chloroflexi-like 16S rRNA gene sequences across sites was compared to the abundance of Chloroflexi marker genes in the metagenome assemblies identified using AMPHORA (**Figure 6**). The majority of Chloroflexi-like 16S rRNA sequences were most similar to either *Chloroflexus* or *Roseiflexus* spp.; however, many sequences fell outside of the family Chloroflexaceae and grouped with other members of the Chloroflexi that are not known to exhibit phototrophy (**Figure 6**). Additionally, *Roseiflexus*-like populations from MS_15, CP_7, and FG_16 and *Chloroflexus*-like populations

FIGURE 6 | Comparison of Chloroflexi phylogenetic marker genes from metagenomes and Chloroflexi 16S rRNA clones. Phylogenetic marker genes within the metagenome sequences assigned to the phylum Chloroflexi using AMPHORA **(A)** compared to the identity (confidence threshold of 80%) of Chloroflexi-like 16S rRNA genes **(B)** observed in the ribosomal clone library (n ∼ 300 per site). Taxonomic groups of Chloroflexi: red = *Roseiflexus* spp., green = *Chloroflexus* spp., brown shades = other taxa within the order Chloroflexales, and yellow shades = other taxa within phylum Chloroflexi.

from BLVA and WC_6 each formed monophyletic groups that excluded sequences from all other springs (**Figure A7** in Appendix). Other spring-specific clades were observed for sequences from FG_16 within the class *Anaerolineae*, a group of Chloroflexi that was very recently shown to contain phototrophic members (Klatt et al., 2011). The presence of these 16S rRNA gene sequences, combined with observed Chloroflexi-like photosynthesis genes associated with these populations, suggests that these undescribed Chloroflexi may also contribute to phototrophy in these mat communities.

FUNCTIONAL ANALYSIS OF PREDOMINANT SEQUENCE ASSEMBLIES
Carbon fixation
The gene content of major scaffold clusters provides a basis for inferring the possible metabolic functions of dominant populations present in these communities (**Table 3**). For example, genes encoding key enzymes involved in the 3-hydroxypropionate (3-HP) pathway of inorganic carbon fixation were present in the metagenomes from all six sites, and were associated with the predominant *Chloroflexus* and *Roseiflexus*-like populations present in these habitats. Genes coding for subunits of ribulose 1,5-bisphosphate carboxylase-oxygenase (RuBisCO), a key enzyme in the reductive pentose phosphate pathway (i.e., Calvin-Benson-Bassham cycle) were observed only in cyanobacterial (WC_6 and CP_7) or proteobacterial sequences (alpha-proteobacteria and *Thermochromatium* spp. in FG_16 and BLVA_20, respectively). No CO_2 fixation genes were associated with the sequences derived from the "*Ca.* T. aerophilum"-like populations from CP_7, despite the fact that other cultivated members of this phylum are capable of fixing CO_2 via the reductive tricarboxylic acid (rTCA) cycle. The average coverage of "*Ca.* T. aerophilum"-assemblies (∼3×) may not be sufficient to conclude that these Chlorobi definitively lack the capacity to fix inorganic carbon, however, metatranscriptomic studies with much deeper coverage also failed to identify key genes (i.e., ATP-citrate lyase) of the rTCA cycle in these populations at *Mushroom Spring* (Liu et al., 2012). This organism is a member of a novel, family level lineage of the Chlorobi, which are predicted to be aerobic photoheterotrophs that cannot oxidize sulfur compounds, cannot fix N_2, and do not fix CO_2 autotrophically (Liu et al., 2012).

Chlorophototrophy
Genes involved in (bacterio)chlorophyll biosynthesis and the production of photosynthetic reaction centers (here termed chlorophototrophy genes) were present in scaffold clusters corresponding to *Roseiflexus*, *Chloroflexus*, *Thermochromatium*, and *Synechococcus* spp., as well as the "*Ca.* T. aerophilum"-like population in CP_7, and other Cyanobacteria, especially in WC_6 (**Table 3**). Consequently, the dominant phototrophs within each community exhibit genomic capability for chlorophototrophic metabolism. Examination of shorter (<10 kbp) scaffolds revealed additional genes involved in chlorophototrophy, and these were assigned to specific chlorophototrophic organisms such as "*Ca.* Chloracidobacterium spp." present in WC_6, and uncultivated proteobacteria in the FG_16 subsurface mat community (**Table 3**). The high G + C% proteobacterial sequences from FG_16 averaged 74% identity (AA) to *Rhodopseudomonas palustris* and other
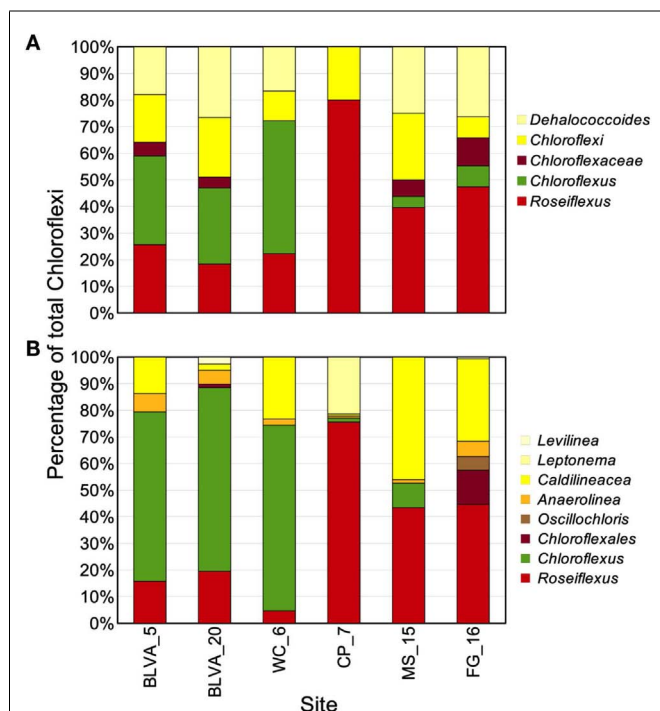
Table 3 | Phylogenetic distribution of autotrophic, phototrophic, and sulfur cycling genes in metagenomes.

Phylogenetic Group	BLVA Green (BVLA_5)	BLVA Purple (BVLA_20)	White Creek (WC_6)	Chocolate Pots (CP_7)	Mushroom Spring (MS_15)	Fairy Geyser (FG_16)
CARBON FIXATION PATHWAYS						
Roseiflexus spp.	**0.60**	**0.80**	0.20	**0.90**	**0.60**	**0.80**
Chloroflexus spp.	**1.00**	**1.00**	**0.80**			
Other Chloroflexi				0.20		**0.50**
Cyanobacteria			0.20	**0.60**		
Thermochromatium spp.		**0.67**				
Alpha-proteobacteria						**0.67**
(BACTERIO)CHLOROPHYLL BIOSYNTHESIS						
Roseiflexus spp.	**0.76**	**0.57**	0.14	**0.86**	**0.57**	**0.76**
Chloroflexus spp.	**1.00**	**0.90**	**0.76**		0.19	0.05
Other Chloroflexi			0.14	0.14	0.14	0.43
Thermochromatium spp.		0.13				
Cyanobacteria			0.45	**1.36**	0.27	0.09
Ca. Thermochlorobacter spp.			0.08	**0.83**	0.08	0.25
Ca. Chloracidobacterium spp.			0.23		0.05	
Alpha-proteobacteria						0.25
PHOTOSYSTEM REACTION CENTERS						
Roseiflexus spp.	**0.50**	0.25		**0.50**		**0.50**
Chloroflexus spp.	0.40	0.40	0.20			
Cyanobacteria			0.33	**0.97**	**0.52**	**1.00**
SULFUR CYCLING						
Thermochromatium spp.		**0.80**				

Entries represent relative completeness of indicated pathways calculated as the fraction of a unique occurrence of a gene in a taxon divided by the total number of genes known to be involved in that function (values > 0.5 are in bold). Metagenome sequences were compared to known pathways in the genome sequences of Chloroflexus aurantiacus J-10-fl, Roseiflexus sp. strain RS-1, Thermochromatium spp., Allochromatium vinosum, Synechococcus sp. strain A, "Candidatus Chloracidobacterium thermophilum," Chloroherpeton thalassium, and the alpha-proteobacterium, Rhodopseudomonas palustris TIE-1.

alpha-proteobacterial genomes, and are likely contributed from the *Hyphomicrobiaceae* population in FG_16. Genes from Chloroflexi coding for chlorophototrophic functions, but too divergent to originate from either *Chloroflexus* or *Roseiflexus* spp. (i.e., only ~70% AA ID), were present in all non-sulfidic sites, especially in FG_16 (**Table 3**). The Chloroflexi-like chlorophototrophy genes from FG_16 are phylogenetically distinct (<70% AA ID) from previously described metagenome sequences and all related sequences residing in public databases, indicating that novel uncultured phototrophic members of the Chloroflexi inhabit the mats at *Fairy Geyser*. Three deduced protein sequences from the subsurface layer in *Mushroom Spring* (MS_15) were highly similar (96–100% AA ID) to translated sequences of novel chlorophototrophy genes observed in recent "meta-omic" studies of the top-layers of this same mat type (Klatt et al., 2011; Liu et al., 2011); these observations linked these genes to a group within the Chloroflexi not previously known to contain chlorophototrophic organisms.

Iron oxidation
One goal of this study was to investigate the role of anoxygenic photosynthesis in sulfidic communities from *Bath Lake Vista Annex* and in iron mats at *Chocolate Pots*. Previous studies near the source of *Chocolate Pots* (and near CP_7) have shown that the oxidation of aqueous Fe(II) is abiotic, but mediated by the production of oxygen by cyanobacteria (Pierson et al., 1999; Trouwborst et al.,

2007). However, voltammetric microelectrode studies revealed that Fe(II) persists in deeper layers of the mat, providing a potential niche for anoxygenic phototrophs that can use Fe(II) as an electron donor for photosynthesis (photoferrotrophy) (Trouwborst et al., 2007). Query genes for both sulfur and Fe(II) oxidation (Croal et al., 2007; Jiao and Newman, 2007; Frigaard and Dahl, 2009; Grimm et al., 2011; Bryant et al., 2012) were used to search for evidence of sulfide or Fe(II) oxidation in the community from CP_7. No genes with significant similarity to the photosynthetic iron oxidation (*pio*) operon of the purple non-sulfur *Rhodopseudomonas palustris* TIE-1 (Jiao and Newman, 2007) or the *fox* operon of the purple non-sulfur *Rhodobacter ferrooxidans* SW2 (Croal et al., 2007) were observed in CP_7, or any site described in this study with the exception of one sequence in FG_16, a site that contains below detectable levels of iron (**Table 1**). This result concurs with the low numbers of alpha-proteobacterial sequences in CP_7 (**Table 3**), and the lack of Fe(II) oxidation when similar mats were illuminated with near-infrared radiation to excite bacteriochlorophylls (Trouwborst et al., 2007). To date, no thermophilic representatives of purple and green photoferrotrophs have been discovered.

Sulfur oxidation
Genes known to encode proteins involved in sulfur oxidation (*dsr* complex) in some anoxygenic phototrophs (e.g., gammaproteobacterium *Allochromatium vinosum*, Dahl et al.,

2005; Frigaard and Dahl, 2009; Gregersen et al., 2011) were identified in the *Thermochromatium*-like population from BLVA_20, and this is consistent with the high concentrations of DS (>100 μM) measured *in situ*. However, the dominant *Chloroflexus*-like populations observed in both *BLVA* samples do not contain *dsr* or *sox* genes known to be involved in the oxidation of reduced-sulfur compounds. This is consistent with the absence of these same genes in reference *Chloroflexus* and *Roseiflexus* spp. genomes (van der Meer et al., 2010; Tang et al., 2011). However, the *Chloroflexus* assemblies from BLVA_20 and *Roseiflexus* assemblies of CP_7 (as well as FAP reference genomes) contain *sqr* genes, which encode sulfide-quinone oxidoreductases and have been suggested to play a role in the oxidation of sulfide to elemental sulfur in multiple bacterial phyla (Griesbeck et al., 2002; Chan et al., 2009; Marcia et al., 2009). Consequently, it is possible that proteins encoded by *sqr* genes may enable FAPs to obtain electrons from reduced-sulfur compounds (Frigaard and Dahl, 2009; Gregersen et al., 2011; Bryant et al., 2012). In the current study, the presence of similar *Chloroflexus* as well as similar *Roseiflexus* populations across both sulfidic and non-sulfidic sites argues that utilization of sulfide as an electron source is not an obligate physiological trait across these genera.

Anaerobic metabolism

Sequence clusters corresponding to undescribed organisms from the Bacteroidetes show no evidence of chlorophototrophy, but rather contain genes suggestive of anaerobic metabolism(s). Protein-coding genes involved in the oxidation and/or fermentation of organic acids were noted in several sites. For example, acyl-CoA synthetases and lactate dehydrogenases were found in unidentified clusters from BLVA (G + C = 64%) and CP_7 (G + C = 31%) and a mixed cluster containing sequences from BLVA and CP (G + C = 36%). Subunits of a pyruvate ferredoxin: oxidoreductase (PFOR) were found in both unidentified BLVA clusters. Although important in every mat type, insufficient coverage of the less-dominant anaerobic populations present in chlorophototrophic mats precludes a thorough analysis of their metabolic potential.
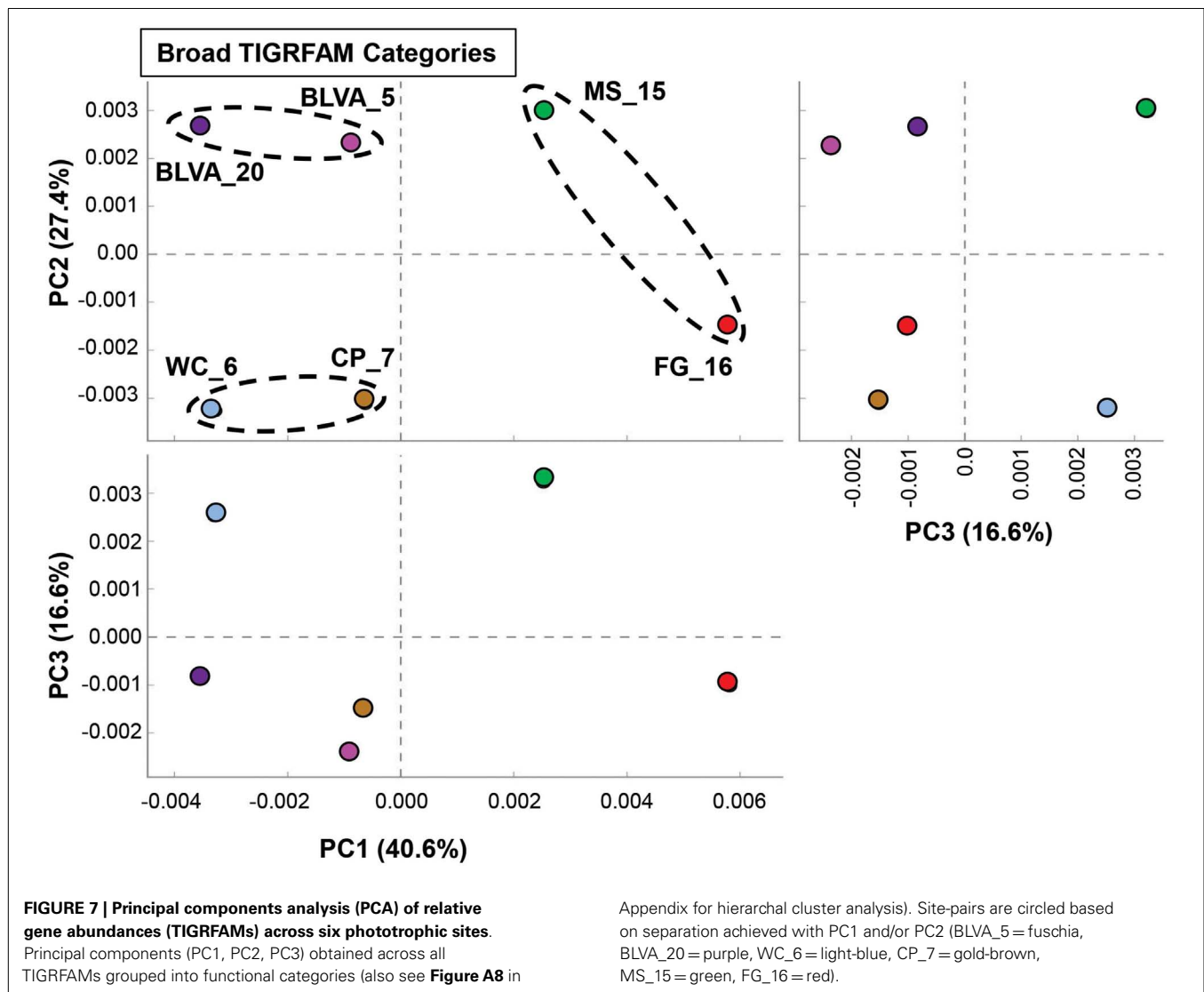
COMPARATIVE ANALYSIS OF PROTEIN FAMILIES

A complete functional analysis was performed (using multivariate statistical analysis) by assigning TIGRFAM protein families to predicted proteins within all metagenome assemblies. Differences in gene contents among the six chlorophototrophic mats should be indicative of changes in community structure and the corresponding functional attributes of dominant community members. PCA was used to examine the relative differences among sites based on all TIGRFAM categories (**Figure 7**). Factor 1 (PC1, accounting for ~41% of the relative functional variation across sites) separates subsurface from surface mat communities, while PC2 (~27% of variation) separates the sites according to different levels of oxygen (or sulfide) and the presence of oxygenic phototrophs. Factor 3 (PC3, ~17% of variation) emphasizes functional similarities between MS_15 and WC_6 that are difficult to separate based only an examination of the abundance of different phylotypes across these sites (e.g., **Figure 2**). For example, although both sites contained cyanobacteria (e.g., low sulfide),

MS_15 contained more sequences related to *Roseiflexus* spp., while WC_6 contained numerous *Chloroflexus*-like sequences. These populations may be organotrophic in this environment and not dependent on sulfide or elemental sulfur (**Table 1**; **Figure 6**).

Specific TIGRFAM categories responsible for differences across sites were also evaluated using hierarchical cluster analysis. Two approaches were evaluated using either a smaller set of TIGRFAM categories related to "energy metabolism" (**Figure 8**) or all TIGRFAM families (**Figure A8** in Appendix). In each case, communities (sites) clustered as expected based on replication of specific variables such as sulfide/oxygen, temperature, and mat sample depth (Inskeep et al., 2013). The relative abundance of TIGRFAMs associated with "energy metabolism" was evaluated and included genes related to sugar degradation, glycolysis/gluconeogenesis, pentose phosphate pathway, fermentative processes, electron transport, and chemolithoautotrophy (**Figure 8**). Site clustering using these TIGRFAMs confirmed greater metabolic potential for processes such as aerobic metabolism and oxygenic photosynthesis in CP_7 and WC_6, samples that contained the most cyanobacteria (e.g., *Synechococcus*, *Fischerella*). Conversely, the subsurface mat communities (FG_16 and MS_15) exhibited a greater abundance of genes related to the Entner-Doudoroff pathway and fermentative processes, which are expected to be more important in subsurface environments occurring just below the predominant cyanobacterial populations (See Materials and Methods). Relative abundance within the TIGRFAM category "aerobic metabolism" revealed greater numbers of these genes in sites that contained significant levels of dissolved oxygen (i.e., no DS) compared to sulfidic sites (BLVA_5, 20). Moreover, TIGRFAMs associated with "anaerobic metabolism" as well as "chemoautotrophy" were higher in the sulfidic sites (BLVA sites 5 and 20) (**Figure 8**), although some of these TIGRFAMs are also present in subsurface mat communities. As should be clear, specific inferences on the basis of a TIGRFAM assignment must be followed with further analysis of the specific gene or set of genes responsible for the abundance estimates within a category.

Hierarchical cluster analysis across all TIGRFAMs grouped into 52 functional categories showed generally similar results regarding site clustering, but the number of TIGRFAM categories used in the analysis precludes a full description of all protein families (**Figure A8** in Appendix). Based on clear differences in the phylotypes observed in sulfidic (hypoxic) vs. oxic samples, the TIGRFAM abundance profiles from BLVA (sites 5 and 20), and those from CP_7 and WC_6 formed separate clusters as expected. However, relative TIGRFAM abundance profiles of the subsurface mat communities (FG_16 and MS_15) did not form a separate cluster, as these sites simply do not exhibit greater similarity to one another compared to similarity among all sites (e.g., organisms similar to *Roseiflexus* spp. are present in all sites). Despite similarities in physical context, the two subsurface communities (MS_15, FG_16) revealed different functional signatures consistent with substantial differences in community composition described above (**Figure 2**), and that are likely due to differences in geochemistry and temperature between the two samples (FG_16 is ~15°C cooler than MS_15 and exhibits higher pH values, above pH 9). Consequently, the functional profiles across all TIGRFAM groupings are consistent with, and provide further support for, the

FIGURE 7 | Principal components analysis (PCA) of relative gene abundances (TIGRFAMs) across six phototrophic sites. Principal components (PC1, PC2, PC3) obtained across all TIGRFAMs grouped into functional categories (also see **Figure A8** in Appendix for hierarchal cluster analysis). Site-pairs are circled based on separation achieved with PC1 and/or PC2 (BLVA_5 = fuschia, BLVA_20 = purple, WC_6 = light-blue, CP_7 = gold-brown, MS_15 = green, FG_16 = red).

differences in community structure between MS_15 and FG_16 (**Figure A8** in Appendix).

DISCUSSION

The six sites investigated in this study are representative of three general types of geothermal springs in Yellowstone National Park that support bacterial chlorophototrophic communities and include (i) alkaline-siliceous chloride springs (pH 7.5–9; e.g., WC_6, MS_15, and FG_16), (ii) sulfidic-carbonate springs (pH 6–7; e.g., BLVA_5 and BLVA_20), and (iii) mildly acidic (pH 6) non-sulfidic springs containing high aqueous Fe(II) (e.g., CP_7) (Rowe et al., 1973; McClesky et al., 2005). The major physical and geochemical constraints that have been postulated to control the distribution of phototrophs (and photosynthesis) in these thermal springs are pH, temperature, sulfide concentration, and gradients in light and/or other chemicals existing as a function of mat depth (Brock, 1967, 1978; Cox et al., 2011; Boyd et al., 2012). The upper temperature limit of cyanobacterial photosynthesis is known to occur at ∼74°C (Brock, 1973), and the grazing of

these microbial mats by eukaryotic organisms typically only occurs at temperatures below 50°C. Most springs that support bacterial chlorophototrophic mats occur at pH > 5, with rare exceptions such as the acid-tolerant, purple non-sulfur phototrophs related to *Rhodopila* sp. observed in *Nymph Lake* (YNP) and in small sulfidic, acidic (pH 3.5–4.5) springs near the *Gibbon River* (Pfennig, 1974; Madigan et al., 2005). The bulk aqueous pH at CP_7 is near the lower limit observed for thermophilic cyanobacteria (Brock, 1973), and microelectrode measurements of the CP_7 mat revealed that it was constantly flushed by vent water with a pH ∼ 6 (Trouwborst et al., 2007). Even at pH 6, CP_7 supports an active community of cyanobacteria that are similar to *Synechococcus* sp. B′-like populations observed in *Mushroom* and *Octopus Spring* (pH > 8) phototrophic mats (**Figure A1** in Appendix).

DISTRIBUTION OF ANOXYGENIC PHOTOTROPHS

Anoxygenic chlorophototrophs are known to colonize sulfidic springs of YNP (van Niel and Thayer, 1930; Castenholz, 1969, 1977; Madigan, 1984; Giovannoni et al., 1987), and this was

FIGURE 8 | Hierarchical cluster analysis of relative abundances of genes in TIGRFAMs associated with "Energy Metabolism" classified by functional category. Data was standardized by functional category before clustering to avoid biasing analysis by a few categories with high gene abundance. Pearson correlation was used as the distance measure for average linkage agglomerative clustering.

confirmed in samples from BLVA in which concentrations of DS exceeded 100 μM. However, the only population with genes supporting a complete, well-studied sulfide-oxidization pathway (Dahl et al., 2005) was the *Thermochromatium*-like organisms present in BLVA_20. The other prominent anoxygenic chlorophototrophs included populations of *Chloroflexus* and *Roseiflexus*-like spp. (identified across all sites). The abundance of chlorophototrophic Chloroflexi across sites is reflective of their previously established physiological diversity, including photoheterotrophy with organic acids such as acetate and propionate, photoautotrophy, photomixotrophy, and oxic and anoxic chemoorganotrophy (Madigan et al., 1974; Pierson and Castenholz, 1974; Giovannoni et al., 1987; Hanada et al., 2002; van der Meer et al., 2003, 2010; Zarzycki and Fuchs, 2011). While these organisms are generally photoheterotrophic, their metabolic flexibility contributes in part to their ability to colonize a broad spectrum of slightly acidic to neutral pH environments at 50–70°C (Castenholz and Pierson, 1995). Highly similar (>98% average NT ID) *Roseiflexus*-like organisms were abundant in all sites, independent of bulk sulfide concentration. Moreover, *Chloroflexus*-like populations were found in both sulfidic (BLVA) and oxic systems (WC_6). The presence of *Roseiflexus* spp. sequences in BLVA_5 and _20 and the larger proportion of *Chloroflexus* spp. in WC_6 compared to *Roseiflexus* spp. was unexpected, as it has been shown that *Chloroflexus* spp. tolerate higher levels of sulfide in

culture (Madigan et al., 1974; Giovannoni et al., 1987; van der Meer et al., 2010). These results suggest that sulfide concentration is not a deterministic variable explaining niche partitioning between *Chloroflexus* spp. and *Roseiflexus* spp. This inconsistency with expected distribution patterns implies that factors other than sulfide and/or oxygen are important in controlling the relative abundance of *Chloroflexus* and *Roseiflexus* spp. in YNP phototrophic mat environments. Finally, sequences assigned to "*Ca.* C. thermophilum" (phylum Acidobacteria) (Bryant et al., 2007) were most abundant in the oxic communities of WC_6 and MS_15 (∼8 and 3% of sequences, respectively). Although small numbers of sequences (<1%) assigned to this organism (BLASTN, >50% NT ID) were observed in other sites, genes encoding enzymes of (B)Chl biosynthesis and belonging to "*Ca.* C. thermophilum" were only found in WC_6 and MS_15 (**Table 3**).

The observed differences in functional gene content between the two subsurface mat communities (MS_15 and FG_16) were of further interest, in part due to the presence of different poorly understood organisms in both sites. "Red-layer" communities (FG_16) have been shown to contain novel phototrophs (Boomer et al., 2000, 2002), whose pigments exhibit unusual *in vivo* absorption spectra (Boomer et al., 2000). Indeed, the FG_16 sample contained a high G + C (∼68–70%) alpha-proteobacterial population not observed in any other site (**Figure 2**). The 16S rRNA sequences from FG_16 indicated the presence of an alphaproteobacterium (family *Hyphomicrobiaceae*), some members of which are known to produce BChl *b* (Hiraishi, 1997). BChl *b* pigments were detected in solvent-based extractions from *Fairy Geyser* mat samples (M. Pagel and D. A. Bryant, unpublished data) and suggest that the phototrophs producing these pigments may exhibit light-harvesting properties that differ from those of other chlorophototroph populations in the mats.

Differences in community composition between the two subsurface mat communities may be driven by differences in temperature (60 vs. 36–40°C in MS_15 and FG_16, respectively). However, the MS_15 subsurface community was also distinct from surface (top 1–2 mm) communities sampled from the same mats at the same temperature (Klatt et al., 2011). For example, the abundance of *Thermotoga* spp. in the subsurface communities may be driven primarily by lower oxygen levels shown to exist 2 mm below the mat surface (Jensen et al., 2011) and is consistent with their physiology as microaerophilic heterotrophs (van Ooteghem et al., 2004). Anaerobic fermentation by *Thermotoga* spp. could constitute a major source of H_2 that could enable photomixotrophic metabolism by *Chloroflexus* and *Roseiflexus* spp. (Klatt et al., 2013). Moreover, compared to the phototrophic surface layers of these mats, MS_15 subsurface communities contained fewer *Synechococcus* spp., greater *Roseiflexus* spp., and greater numbers of likely anaerobic or fermentative organisms within the Bacteroidetes and Thermodesulfobacteria.

TROPHIC INTERACTIONS

Trophic interactions between FAPs and cyanobacteria have been studied in phototrophic geothermal mats, and it has been shown that photoheterotrophs (FAPs) utilize organic acids produced by autotrophic cyanobacteria (Anderson et al., 1987; Nold and

Ward, 1996; van der Meer et al., 2005). Moreover, it has been proposed that *Thermochromatium* spp. (purple-sulfur bacteria) are primary producers in sulfidic springs and cross-feed low-molecular weight organic acids to FAPs (Madigan et al., 1989, 2005). This is analogous to the cyanobacterial primary production and trophic interactions documented to occur in *Octopus Spring* and *Mushroom Spring* (van der Meer et al., 2005). However, this hypothesis is not supported by the relatively heavy carbon isotope composition of Chloroflexaceae-specific lipid biomarkers in sulfidic springs ($\delta^{13}C = -8.9$ to -18.5 ‰, van der Meer et al., 2003). These isotopic compositions have been interpreted to be too heavy to originate from compounds cross-fed from *Thermochromatium* spp., which use the Calvin-Benson-Bassham cycle for carbon dioxide fixation ($\delta^{13}C = -20$ to -35 ‰). The lipid signatures are more readily explained by direct carbon dioxide fixation by *Chloroflexus* and *Roseiflexus* spp. via the 3-HP pathway (Holo and Sirevåg, 1986; Strauss and Fuchs, 1993; van der Meer et al., 2000, 2010). Metagenome sequence assemblies obtained in the current study showed that these uncultivated *Chloroflexus* and *Roseiflexus* spp. contained all genes necessary for CO_2 fixation via the 3-HP pathway (**Table 3**), and is consistent with earlier evidence at *BLVA* of short-term, sulfide-stimulated $^{14}CO_2$ incorporation by FAPs (Giovannoni et al., 1987). Collectively, these observations support the hypothesis that all major chlorophototrophs contribute to primary productivity in sulfidic-carbonate springs (**Table 3**). It remains to be determined whether FAPs are more important contributors to primary productivity in these systems when purple-sulfur bacteria (i.e., *Thermochromatium*) and cyanobacteria are both absent (such as observed in BLVA_5).

This study highlights several of the major differences in community composition and structure, and potential function of chlorophototrophic microbial mats sampled from high-temperature systems (40–60°C) containing high sulfide, high Fe(II), or high dissolved oxygen. The distribution of chlorophototrophic organisms, as would be expected, is dependent on the presence or absence of high sulfide (cyanobacteria, purple-sulfur bacteria), and position within laminated mats (e.g., FAPs, Bacteroidetes, and Firmicutes). Temperature was not particularly well constrained as a consistent parameter for comparisons across the sites included in this study. However, the ubiquity of *Chloroflexus* and *Roseiflexus* spp. across all sites emphasizes their ability to tolerate large differences in not only temperature, but extremes between high and low levels of DS and/or oxygen. Assemblies of a novel Chlorobi population ("*Ca*. T. aerophilum") from the high iron site at *Chocolate Pots* (CP_7) were similar to those obtained from *Mushroom Spring* and *Octopus Spring* (Liu et al., 2012). These populations deserve further study, especially considering their phylogenetic distance and different functional attributes compared to other currently described members of the Chlorobi. The dominant cyanobacteria observed across these sites (found exclusively in non-sulfidic systems) included *Synechococcus* spp. (CP_7, MS_15) and *Fischerella* (*Mastigocladus*) spp. (WC_6). Consequently, sulfide is a critical geochemical variable that selects against the presence of cyanobacteria and provides niche opportunities for other chlorophotoautotrophs. Other poorly represented organisms in the current study include bacteria from the phyla Firmicutes and

Bacteroidetes, and although the assemblies for organisms within these phyla were not particularly large, a sufficient number of genes were found to infer that their role in these communities may involve fermentation and the degradation of complex carbon compounds. Additional sequence assembly and/or isolation of these populations, coupled with site-specific studies, are necessary to clarify the important carbon cycling functions that these populations conduct and the processes that drive interactions among primary producers and secondary consumers in chlorophototrophic mats.

MATERIALS AND METHODS

SAMPLE COLLECTION AND GEOCHEMICAL ANALYSES

Six different samples were obtained from five hot springs between August 2007 and May 2008 (**Table 1**; Table S2 in Inskeep et al., 2013) and immediately frozen in liquid N_2. Phototrophic mats were sampled at different locations relative to the source of each respective spring, and two samples were obtained from subsurface mat layers [*Mushroom Spring* (MS_15) and *Fairy Geyser* (FG_16)]. The subsurface layers were obtained by careful removal of the top 2 mm green layer with a sterile scalpel and separation of a definitive under-layer in each mat type (e.g., Boomer et al., 2000, 2002; Nübel et al., 2002). Geochemical characterization was performed on bulk spring water at each sampling location after filtration (0.2 μm). Total dissolved ions were determined using inductively coupled plasma spectrometry and major anions determined using ion chromatography as described previously (Macur et al., 2004; Inskeep et al., 2005). Temperature, pH, total DS, total soluble Fe, and dissolved oxygen were determined immediately in the field. Dissolved gases (CO_2, CH_4, and H_2) were determined using headspace gas chromatography of filtered field samples (Inskeep et al., 2005).

DNA EXTRACTION AND PREPARATION

Environmental DNA was extracted as described in Inskeep et al. (2013). Briefly, 0.5–1 g of frozen mat samples were processed using separate parallel DNA extractions with an enzymatic method (Proteinase K (1 mg/ml) with Na-dodecyl sulfate (SDS) (0.3% w/v) for 0.5 h at 37°C) and a mechanical method (bead-beating with 2% w/v SDS and 15% v/v Tris-HCl-equilibrated phenol, shaken at 5.5 m/s for 30 s) for cell lysis. The resulting cell lysates were pooled and subsequent DNA extractions were performed with phenol:chloroform:isoamyl alcohol (25:24:1), and chloroform:isoamyl alcohol (24:1). This procedure removed DNA extraction bias that has been shown to occur when only mechanical or enzymatic protocols are used for cell lysis (Klatt et al., 2007, 2011). All samples were treated with RNAse I (Promega, Madison, WI, USA), and DNA was precipitated with ethanol and Na-acetate. Small-insert (3 kb) metagenome libraries were constructed as described in Inskeep et al. (2013). About 820 bp was sequenced at each end of the inserts in the library clones, which produced pairs of linked sequences (424,982 sequences) that represented a total dataset of ∼320.6 Mbp. Ribosomal (16S rRNA) gene sequence libraries were constructed by PCR amplification using universal primers targeting domains Archaea (4aF, TCCGGTTGATC-CTGCCRG; 1391R, GACGGGCRGTGWGTRCA) and Bacteria (27F, AGAGTTTGATCCTGGCTCAG and 1391R). Amplicons

were cloned using the TOPO TA Cloning Kit (Invitrogen, Carlsbad CA USA) and sequenced using Big Dye v3.1 chemistry (Applied Biosystems, Foster City, CA, USA).

PRE-ASSEMBLY METAGENOME SEQUENCE ANALYSES

All metagenome sequences were used as queries in a "blastx" (Camacho et al., 2009) search against the NCBI nr database (accessed 22 March 2011) with default parameters. The results were parsed and visualized with the MEGAN software version 2.3.2 (Huson et al., 2007) with the default parameters (MinScore = 35.0, TopPercent = 10.0, MinSupport = 5) and taxonomic assignments of the top "blastx" matches were extracted. Comparative analysis was also completed using several relevant reference genomes available after this date (e.g., *Fischerella* sp. and "*Ca.* T. aerophilum"; Liu et al., 2012).

SEQUENCE ASSEMBLY AND ANNOTATION

Metagenomic scaffolds of overlapping end sequences were constructed separately for each of the six samples using the Celera assembler (Miller et al., 2008; Inskeep et al., 2013). This resulted in 206,469 scaffolds containing 183.2 Mbp (27–33 Mbp per site) of assembled sequence, or a 57% compression of the raw sequence data. The DOE-JGI annotation pipeline was used as an initial step for inferring functions for predicted ORFs on metagenome scaffolds, and included open reading frame (ORF) prediction, BLAST alignments, and hidden Markov model analysis (Mavromatis et al., 2009). Translated peptide sequences from predicted ORFs were analyzed with the AMPHORA package (Wu and Eisen, 2008), which identified homologs of 31 different genes (mostly predicted to encode ribosomal proteins or enzymes with housekeeping functions) that could be used as phylogenetic markers. Genes encoding particular functions were identified by BLASTP using reference sequences as queries, with the additional requirement that candidate sequences had a top BLASTP match to a sequence with the same annotated function in the NCBI nr database. All annotated metagenome sequence assemblies (Celera/PGA) discussed in the current manuscript are available through the DOE-JGI IMG/M (Markowitz et al., 2012) website (http://img.jgi.doe.gov/m) under IMG taxon OID numbers as follows: YNPSite06 (2022920004/2013515000), Site07 (2022920013/2014031006), Site15 (2022920016/2015219002), Site16 (2022920018/2016842003), Site05 (2022920003/2013954000), Site20 (2022920020/2016842008), and Site17 (2022920021/2016842005).

RIBOSOMAL RNA SEQUENCE ANALYSES

All bacterial 16S rRNA sequences from the 16S rRNA-specific PCR clone libraries were aligned and screened for chimeras with Bellerophon (Huber et al., 2004) with subsequent manual curation. OTUs were determined using the CAP3 assembler (Huang and Madan, 1999) at the 99% demarcation level. Rarefaction curves were determined, and the Chao1 and ACE richness indexes and the Fisher's alpha, Shannon-Weaver, and Simpson's diversity indexes were calculated for each library (EcoSim version 7.0, Gotelli and Entsminger, 2001; EstimateS v. 8.0, Colwell, 2009). The RDP Bayesian Classifier (Wang et al., 2007) was used to assign taxonomy to 16S rRNA sequences at the 80% confidence level (**Figures 5B** and **6B**), and all sequences belonging

to the Chloroflexi were aligned with reference sequences corresponding to *Escherichia coli* positions 29–1349 (1321 positions). Alignments were masked with bacterial complexity filters in ARB (Ludwig et al., 2004). A phylogenetic tree was produced using the BioNJ algorithm (Gascuel, 1997) (**Figure 2**) and bootstrapped with 1000 replicates. Reference sequences shorter than the initial alignment were subsequently added to the tree using the ARB parsimony tool. Consensus maximum-likelihood trees were produced from 1000 replicate trees using RaxML (Stamatakis, 2006). A maximum-likelihood tree based upon amino acid alignments of PscD sequences was constructed using PhyML (Guindon et al., 2010).

STATISTICAL ANALYSES

A distance matrix of environmental variables was constructed by calculating Gower coefficients using the R statistical environment (R Development Core Team, 2012). The Gower coefficient allows for different data types (qualitative presence/absence vs. quantitative numerical) with different dimensional scales to be combined into a general dissimilarity metric (Gower, 1971). Geochemical variables were treated as factors and were correlated to this distance matrix using the envfit function of the vegan package (Oksanen et al., 2012). Metagenomic scaffolds larger than 10 kbp were subjected to analysis using oligonucleotide composition. All possible tri-, tetra-, penta-, and hexanucleotides were counted with custom perl scripts, and normalized to the length of the scaffold. Normalized oligonucleotide composition matrices were subjected to k-means clustering with a range of k = 4–12 with 100 trials each. Clusters were reported when at least 10 scaffolds grouped together in 90% or greater Monte-Carlo simulations. The composite summary of these k-means trials was displayed as an interaction network using the program Cytoscape 2.8.1 (Shannon et al., 2003).

BROAD FUNCTIONAL ANALYSIS OF METAGENOME SEQUENCES

Assembled sequence from each of the phototrophic sites was annotated as described in Inskeep et al. (2010) and predicted proteins from the scaffolds were assigned TIGRFAM protein families (Selengut et al., 2007) using HMMER 3 (Eddy, 2011) with *e*-value cutoff of 1e−6. PCA and statistical analysis of site group differences was performed using the STAMP v2.0 software (Parks and Beiko, 2010). The White's non-parametric *T*-test and ANOVA tests were used to test for differences between two site groups and multiple site groups respectively. Two-way clustering was performed using row-standardized (across sites) average TIGRFAM category abundance data using the Euclidean distance metric and complete-linkage hierarchical clustering in MeV 4.8 (Saeed et al., 2003) software. Other details regarding TIGRFAM analysis are described in this issue (Inskeep et al., 2013).

REFERENCES

Anderson, K. L., Tayne, T. A., and Ward, D. M. (1987). Formation and fate of fermentation products in hot spring cyanobacterial mats. *Appl. Environ. Microbiol.* 53, 2343–2352.

Badger, J. H., Hoover, T. R., Brun, Y. V., Weiner, R. M., Laub, M. T., Alexandre, G., et al. (2006). Comparative genomic evidence for a close relationship between the dimorphic prosthecate bacteria *Hyphomonas neptunium* and *Caulobacter crescentus*. *J. Bacteriol.* 188, 6841–6850.

Becraft, E. D., Cohan, F. M., Kuhl, M., Jensen, S. I., and Ward, D. M. (2011). Fine-scale distribution patterns of *Synechococcus* ecological diversity in microbial mats of Mushroom Spring, Yellowstone National Park. *Appl. Environ. Microbiol.* 77, 7689–7697.

Bhaya, D., Grossman, A. R., Steunou, A.-S., Khuri, N., Cohan, F. M., Hamamura, N., et al. (2007). Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. *ISME J.* 1, 703–713.

Boomer, S. M., Lodge, D. P., Dutton, B. E., and Pierson, B. (2002). Molecular characterization of novel red green nonsulfur bacteria from five distinct hot spring communities in Yellowstone National Park. *Appl. Environ. Microbiol.* 68, 346–355.

Boomer, S. M., Pierson, B. K., Austinhirst, R., and Castenholz, R. W. (2000). Characterization of novel bacteriochlorophyll-a-containing red filaments from alkaline hot springs in Yellowstone National Park. *Arch. Microbiol.* 174, 152–161.

Boyd, E. S., Fecteau, K. M., and Peters, J. W. (2012). Modeling the habitat range of phototrophs in Yellowstone National Park: toward the development of a comprehensive fitness landscape. *Front. Microbiol.* 3:221. doi:10.3389/fmicb.2012.00221

Brock, T. D. (1967). Life at high temperatures. *Science* 158, 1012–1019.

Brock, T. D. (1973). Lower pH limit for the existence of blue-green algae: evolutionary and ecological implications. *Science* 179, 480–483.

Brock, T. D. (1978). *Thermophilic Microorganisms and Life at High Temperatures.* New York: Springer Verlag.

Bryant, D. A., Costas, A. M. G., Maresca, J. A., Chew, A. G. M., Klatt, C. G., Bateson, M. M., et al. (2007). *Candidatus Chloracidobacterium thermophilum*: an aerobic phototrophic Acidobacterium. *Science* 317, 523–526.

Bryant, D. A., Klatt, C. G., Frigaard, N.-U., Liu, Z., Li, T., Zhao, F., et al. (2012). "Comparative and functional genomics of anoxygenic green bacteria from the taxa Chlorobi, Chloroflexi, and Acidobacteria," in *Functional Genomics and Evolution of Photosynthetic Systems, Vol. 33, Advances in Photosynthesis and Respiration*, eds R. L. Burnap and W. Vermaas (Dordrecht: Springer), 47–102.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi:10.1186/1471-2105-10-421

Castenholz, R. W. (1969). Thermophilic blue-green algae and the thermal environment. *Bacteriol. Rev.* 33, 476–504.

Castenholz, R. W. (1976). The effect of sulfide on the blue-green algae of hot springs. I. New Zealand and Iceland. *J. Phycol.* 12, 54–68.

Castenholz, R. W. (1977). The effect of sulfide on the blue-green algae of hot springs II. Yellowstone National Park. *Microb. Ecol.* 3, 79–105.

Castenholz, R. W., and Pierson, B. K. (1995). "Ecology of thermophilic anoxygenic phototrophs," in *Anoxygenic Photosynthetic Bacteria*, eds R. E. Blankenship, M. T. Madigan, and C. E. Bauer (Dordrecht: Kluwer Academic Publishers), 87–103.

Chan, L.-K., Morgan-Kiss, R. M., and Hanson, T. E. (2009). Functional analysis of three sulfide:quinone oxidoreductase homologs in *Chlorobaculum tepidum*. *J. Bacteriol.* 191, 1026–1034.

Cohen, Y., and Rosenberg, E. (eds) (1989). *Microbial Mats: Physiological Ecology of Benthic Microbial Communities.* Washington, DC: American Society for Microbiology.

Colwell, R. K. (2009). *EstimateS: Statistical Estimation of Species Richness and Shared Species from Samples. Version 8.0.* Available at: http://viceroy.eeb.uconn.edu/estimates.

Cox, A., Shock, E. L., and Havig, J. R. (2011). The transition to microbial photosynthesis in hot spring ecosystems. *Chem. Geol.* 280, 344–351.

Croal, L. R., Jiao, Y., and Newman, D. K. (2007). The *fox* operon from *Rhodobacter* strain SW2 promotes phototrophic Fe(II) oxidation in *Rhodobacter capsulatus* SB1003. *J. Bacteriol.* 189, 1774–1782.

Dahl, C., Engels, S., Pott-Sperling, A. S., Schulte, A., Sander, J., Lübbe, Y., et al. (2005). Novel genes of the *dsr* gene cluster and evidence for close interaction of Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium *Allochromatium vinosum*. *J. Bacteriol.* 187, 1392–1404.

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi:10.1371/journal.pcbi.1002195

Frigaard, N., and Dahl, C. (2009). Sulfur metabolism in phototrophic sulfur bacteria. *Adv. Microb. Physiol.* 54, 103–200.

Garcia Costas, A. M., Liu, Z., Tomsho, L. P., Schuster, S. C., Ward, D. M., and Bryant, D. A. (2012). Complete genome of *Candidatus* Chloracidobacterium thermophilum, a chlorophyll-based photoheterotroph belonging to the phylum Acidobacteria. *Environ. Microbiol.* 14, 177–190.

Gascuel, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol. Biol. Evol.* 14, 685–695.

Giovannoni, S., Revsbech, N. P., Ward, D. M., and Castenholz, R. W. (1987). Obligately phototrophic *Chloroflexus*: primary production in anaerobic hot spring microbial mats. *Arch. Microbiol.* 147, 80–87.

Gotelli, N. J., and Entsminger, G. L. (2001). *EcoSim: Null Models Software for Ecology. Version 7.0. Acquired Intelligence Inc. and Kesey-Bear.* Available at: http://garyentsminger.com/ecosim/index.htm.

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics* 27, 857–871.

Gregersen, L. H., Bryant, D. A., and Frigaard, N.-U. (2011). Mechanisms and evolution of oxidative sulfur metabolism in green sulfur bacteria. *Front. Microbiol.* 2:116. doi:10.3389/fmicb.2011.00116

Griesbeck, C., Schütz, M., Schödl, T., Bathe, S., Nausch, L., Mederer, N., et al. (2002). Mechanism of sulfide-quinone reductase investigated using site-directed mutagenesis and sulfur analysis. *Biochemistry* 41, 11552–11565.

Grimm, F., Franz, B., and Dahl, C. (2011). Regulation of dissimilatory sulfur oxidation in the purple sulfur bacterium *Allochromatium vinosum*. *Front. Microbiol.* 2:51. doi:10.3389/fmicb.2011.00051

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321.

Hanada, S., Takaichi, S., Matsuura, K., and Nakamura, K. (2002). *Roseiflexus castenholzii* gen. nov., sp. nov., a thermophilic, filamentous, photosynthetic bacterium that lacks chlorosomes. *Int. J. Syst. Evol. Microbiol.* 52, 187–193.

Heidelberg, J. F., Nelson, W. C., Schoenfeld, T., and Bhaya, D. (2009). Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLoS ONE* 4:e4169. doi:10.1371/journal.pone.0004169

Henry, E. A., Devereux, R., Maki, J. S., Gilmour, C. C., Woese, C. R., Mandelco, L., et al. (1994). Characterization of a new thermophilic sulfate-reducing bacterium *Thermodesulfovibrio yellowstonii*, gen. nov. and sp. nov.: its phylogenetic relationship to *Thermodesulfobacterium commune* and their origins deep within the bacterial domain. *Arch. Microbiol.* 161, 62–69.

Hiraishi, A. (1997). Transfer of the bacteriochlorophyll *b*-containing phototrophic bacteria *Rhodopseudomonas viridis* and *Rhodopseudomonas sulfoviridis* to the genus *Blastochloris* gen. nov. *Int. J. Syst. Bacteriol.* 47, 217–219.

Holo, H., and Sirevåg, R. (1986). Autotrophic growth and CO_2 fixation of *Chloroflexus aurantiacus*. *Arch. Microbiol.* 145, 173–180.

Huang, X., and Madan, A. (1999). CAP3: a DNA sequence assembly program. *Genome Res.* 9, 868–877.

Huber, T., Faulkner, G., and Hugenholtz, P. (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* 20, 2317–2319.

Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Res.* 17, 377–386.

Inskeep, W. P., Ackerman, G. G., Taylor, W. P., Kozubal, M., Korf, S., and Macur, R. E. (2005). On the energetics of chemolithotrophy in nonequilibrium systems: case studies of geothermal springs in Yellowstone National Park. *Geobiology* 3, 297–317.

Inskeep, W. P., Jay, Z. J., Tringe, S. G., Herrgard, M. J., Rusch, D. B., et al. (2013). The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem. *Front. Microbiol.* 4:67. doi:10.3389/fmicb.2013.00067

Inskeep, W. P., Rusch, D. B., Jay, Z. J., Herrgard, M. J., Kozubal, M. A., Richardson, T. H., et al. (2010). Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. *PLoS ONE* 5:e9773. doi:10.1371/journal.pone.0009773

Jackson, T. J., Ramaley, R. F., and Meinschein, W. G. (1973). *Thermomicrobium*, a new genus of extremely thermophilic bacteria. *Int. J. Syst. Bacteriol.* 23, 28–36.

Jensen, S. I., Steunou, A.-S., Bhaya, D., Kuhl, M., and Grossman, A. R. (2011). In situ dynamics of O_2, pH and cyanobacterial transcripts associated with CCM, photosynthesis and detoxification of ROS. *ISME J.* 5, 317–328.

Jiao, Y., and Newman, D. K. (2007). The *pio* operon is essential for phototrophic Fe(II) oxidation in *Rhodopseudomonas palustris* TIE-1. *J. Bacteriol.* 189, 1765–1773.

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282.

Klatt, C. G., Bryant, D. A., and Ward, D. M. (2007). Comparative genomics provides evidence for the 3-hydroxypropionate autotrophic pathway in filamentous anoxygenic phototrophic bacteria and in hot spring microbial mats. *Environ. Microbiol.* 9, 2067–2078.

Klatt, C. G., Liu, Z., Ludwig, M., Kühl, M., Jensen, S. I., Bryant, D. A., et al. (2013). Temporal metatranscriptomic patterning in phototrophic Chloroflexi inhabiting a microbial mat in a geothermal spring. *ISME J.* doi:10.1038/ismej.2013.52

Klatt, C. G., Wood, J. M., Rusch, D. B., Bateson, M. M., Hamamura, N., Heidelberg, J. F., et al. (2011). Community ecology of hot spring cyanobacterial mats: predominant

populations and their functional potential. *ISME J.* 5, 1262–1278.

Liu, Z., Klatt, C. G., Ludwig, M., Rusch, D. B., Jensen, S. I., Kühl, M., et al. (2012). '*Candidatus Thermochlorobacter aerophilum*:' an aerobic chlorophotoheterotrophic member of the phylum Chlorobi defined by metagenomics and metatranscriptomics. *ISME J.* 6, 1869–1882.

Liu, Z., Klatt, C. G., Wood, J. M., Rusch, D. B., Ludwig, M., Wittekindt, N., et al. (2011). Metatranscriptomic analyses of chlorophototrophs of a hot-spring microbial mat. *ISME J.* 5, 1279–1290.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004). ARB: a software environment for sequence data. *Nucleic Acids Res.* 32, 1363–1371.

Macur, R. E., Langner, H. W., Kocar, B. D., and Inskeep, W. P. (2004). Linking geochemical processes with microbial community analysis: successional dynamics in an arsenic-rich, acid-sulfate-chloride geothermal spring. *Geobiology* 2, 163–177.

Madigan, M. T. (1984). A novel photosynthetic purple bacterium isolated from a Yellowstone hot spring. *Science* 225, 313–315.

Madigan, M. T., and Brock, T. D. (1975). Photosynthetic sulfide oxidation by *Chloroflexus aurantiacus*, a filamentous, photosynthetic, gliding bacterium. *J. Bacteriol.* 122, 782–784.

Madigan, M. T., Jung, D. O., Karr, E. A., Sattley, W. M., Achenbach, L. A., and van der Meer, M. T. J. (2005). "Diversity of anoxygenic phototrophs in contrasting extreme environments," in *Geothermal Biology and Geochemistry in Yellowstone National Park*, eds W. P. Inskeep and T. R. McDermott (Bozeman: Montana State University Publications), 203–219.

Madigan, M. T., Petersen, S. R., and Brock, T. D. (1974). Nutritional studies on *Chloroflexus*, a filamentous photosynthetic, gliding bacterium. *Arch. Microbiol.* 100, 97–103.

Madigan, M. T., Takigiku, R., Lee, R. G., Gest, H., and Hayes, J. M. (1989). Carbon isotope fractionation by thermophilic phototrophic sulfur bacteria: evidence for autotrophic growth in natural populations. *Appl. Environ. Microbiol.* 55, 639–644.

Marcia, M., Ermler, U., Peng, G., and Michel, H. (2009). The structure of *Aquifex aeolicus* sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9625–9630.

Markowitz, V. M., Chen, I.-M. A., Chu, K., Szeto, E., Palaniappan, K., Grechkin, Y., et al. (2012). IMG/M: the integrated metagenome data management and comparative analysis system. *Nucleic Acids Res.* 40, D123–D129.

Mavromatis, K., Ivanova, N., Chen, A., Szeto, E., Markowitz, V., and Kyrpides, N. C. (2009). Standard operating procedure for the annotations of microbial genomes by the production genomic facility of the DOE JGI. *Stand. Genomic Sci.* 1, 63–67.

McClesky, R. B., Ball, J. A., Nordstrom, D. K., Holloway, J. M., and Taylor, H. E. (2005). *Water-chemistry Data for Selected Hot Springs, Geysers, and Streams in Yellowstone National Park, Wyoming, 2001-2002.* Open-File Report 2004-1316. Reston, VA: U.S. Geological Survey.

Melendrez, M. C., Lange, R. K., Cohan, F. M., and Ward, D. M. (2011). Influence of molecular resolution on sequence-based discovery of ecological diversity among *Synechococcus* populations in an alkaline siliceous hot spring microbial mat. *Appl. Environ. Microbiol.* 77, 1359–1367.

Miller, J. R., Delcher, A. L., Koren, S., Venter, E., Walenz, B. P., Brownley, A., et al. (2008). Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics* 24, 2818–2824.

Miller, S. R., Castenholz, R. W., and Pedersen, D. (2007). Phylogeography of the thermophilic cyanobacterium *Mastigocladus laminosus*. *Appl. Environ. Microbiol.* 73, 4751–4759.

Miller, S. R., Purugganan, M. D., and Curtis, S. E. (2006). Molecular population genetics and phenotypic diversification of two populations of the thermophilic cyanobacterium *Mastigocladus laminosus*. *Appl. Environ. Microbiol.* 72, 2793–2800.

Miller, S. R., Strong, A. L., Jones, K. L., and Ungerer, M. C. (2009). Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Appl. Environ. Microbiol.* 75, 4565–4572.

Nold, S. C., and Ward, D. M. (1996). Photosynthate partitioning and fermentation in hot spring microbial mat communities. *Appl. Environ. Microbiol.* 62, 4598–4607.

Nübel, U., Bateson, M. M., Vandieken, V., Wieland, A., Kuhl, M., and Ward, D. M. (2002). Microscopic examination of distribution and phenotypic properties of phylogenetically diverse *Chloroflexaceae*-related bacteria in hot spring microbial

mats. *Appl. Environ. Microbiol.* 68, 4593–4603.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., et al. (2012). *Vegan: Community Ecology Package. R Package Version 2.0-5.* Available at: http://CRAN.R-project.org/package=vegan

Parks, D. H., and Beiko, R. G. (2010). Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 26, 715–721.

Pfennig, N. (1974). Rhodopseudomonas globiformis, sp. n., a new species of the *Rhodospirillaceae*. *Arch. Microbiol.* 100, 197–206.

Pierson, B. K., and Castenholz, R. W. (1974). A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. *Arch. Microbiol.* 100, 5–24.

Pierson, B. K., Parenteau, M. N., and Griffin, B. M. (1999). Phototrophs in high-iron-concentration microbial mats: physiological ecology of phototrophs in an iron-depositing hot spring. *Appl. Environ. Microbiol.* 65, 5474–5483.

R Development Core Team. (2012). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*, Vienna. Available at: URL http://www.R-project.org.

Rappé, M. S., and Giovannoni, S. J. (2003). The uncultured microbial majority. *Annu. Rev. Microbiol.* 57, 369–394.

Rowe, J. J., Fournier, R. O., and Morey, G. W. (1973). *Chemical Analysis of Thermal Waters in Yellowstone National Park, Wyoming, 1960–65.* USGS Open File Report. Available at: http://pubs.er.usgs.gov/

Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., et al. (2007). The sorcerer II global ocean sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* 5:e77. doi:10.1371/journal.pbio.0050077

Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., et al. (2003). TM4: a free, open-source system for microarray data management and analysis. *BioTechniques* 34, 374–378.

Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W. C., et al. (2007). TIGR-FAMs and genome properties: tools for the assignment of molecular function and biological process in

prokaryotic genomes. *Nucleic Acids Res.* 35, D260–D264.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.

Strauss, G., and Fuchs, G. (1993). Enzymes of a novel autotrophic CO_2 fixation pathway in the phototrophic bacterium *Chloroflexus aurantiacus*, the 3-hydroxypropionate cycle. *Eur. J. Biochem.* 215, 633–643.

Taffs, R., Aston, J. E., Brileya, K., Jay, Z., Klatt, C. G., McGlynn, S., et al. (2009). In silico approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC Syst. Biol.* 3:114. doi:10.1186/1752-0509-3-114

Tang, K.-H., Barry, K., Chertkov, O., Dalin, E., Han, C., Hauser, L., et al. (2011). Complete genome sequence of the filamentous anoxygenic phototrophic bacterium *Chloroflexus aurantiacus*. *BMC Genomics* 12:334. doi:10.1186/1471-2164-12-334

Teeling, H., Meyerdierks, A., Bauer, M., Amann, R., and Glöckner, F. O. (2004). Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environ. Microbiol.* 6, 938–947.

Trouwborst, R. E., Johnston, A., Koch, G., Luther, G. W. I. I. I., and Pierson, B. K. (2007). Biogeochemistry of Fe(II) oxidation in a photosynthetic microbial mat: implications for precambrian Fe(II) oxidation. *Geochim. Cosmochim. Acta* 71, 4629–4643.

van der Meer, M. T., Schouten, S., de Leeuw, J. W., and Ward, D. M. (2000). Autotrophy of green non-sulphur bacteria in hot spring microbial mats: biological explanations for isotopically heavy organic carbon in the geological record. *Environ. Microbiol.* 2, 428–435.

van der Meer, M. T. J., Klatt, C. G., Wood, J., Bryant, D. A., Bateson, M. M., Lammerts, L., et al. (2010). Cultivation and genomic, nutritional, and lipid biomarker characterization of *Roseiflexus* strains closely related to predominant in situ populations inhabiting Yellowstone hot spring microbial mats. *J. Bacteriol.* 192, 3033–3042.

van der Meer, M. T. J., Schouten, S., Bateson, M. M., Nübel, U., Wieland, A., Kuhl, M., et al. (2005). Diel variations in carbon metabolism by green nonsulfur-like bacteria in alkaline siliceous hot spring microbial mats from Yellowstone National Park. *Appl. Environ. Microbiol.* 71, 3978–3986.

van der Meer, M. T. J., Schouten, S., Sinninghe Damste, J. S., de Leeuw, J. W., and Ward, D. M. (2003). Compound-specific isotopic fractionation patterns suggest different carbon metabolisms among *Chloroflexus*-like bacteria in hotspring microbial mats. *Appl. Environ. Microbiol.* 69, 6000–6006.

van Niel, C. B., and Thayer, L. A. (1930). *Report on Preliminary Observations on the Microflora in and Near the Hot Springs in Yellowstone National Park and Their Importance for the Geological Formations.* YNP Lib File 731.2, Mammoth, WY.

van Ooteghem, S. A., Jones, A., Van Der Lelie, D., Dong, B., and Mahajan, D. (2004). H_2 production and carbon utilization by *Thermotoga neapolitana* under anaerobic and microaerobic growth conditions. *Biotechnol. Lett.* 26, 1223–1232.

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267.

Ward, D. M., and Castenholz, R. W. (2000). "Cyanobacteria in geothermal habitats," in *Ecology of Cyanobacteria*, eds B. A. Whitton, and M. Potts (Dordrecht: Kluwer Academic Publishers), 37–59.

Ward, D. M., Castenholz, R. W., and Miller, S. R. (2012). "Cyanobacteria in geothermal habitats," in *Ecology of cyanobacteria II. Their diversity in space and time*, ed. B. A. Whitton (Heidelberg, GE: Springer), 39–63.

Ward, D. M., Cohan, F. M., Bhaya, D., Heidelberg, J. F., Kühl, M., and Grossman, A. (2008). Genomics, environmental genomics and the issue of microbial species. *Heredity* 100, 207–219.

Ward, D. M., Ferris, M. J., Nold, S. C., and Bateson, M. M. (1998). A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiol. Mol. Biol. Rev.* 62, 1353–1370.

Ward, D. M., Weller, R., Shiea, J., Castenholz, R. W., and Cohen, Y. (1989). "Hot spring microbial mats: anoxygenic and oxygenic mats of possible evolutionary significance," in *Microbial Mats: Physiological Ecology of Benthic Microbial Communities*, eds Y. Cohen, and E. Rosenberg (Washington, DC: American Society of Microbiology), 3–15.

Wu, M., and Eisen, J. (2008). A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* 9, R151.

Zarzycki, J., and Fuchs, G. (2011). Coassimilation of organic substrates via the autotrophic 3-hydroxypropionate bi-cycle in *Chloroflexus aurantiacus*. *Appl. Environ. Microbiol.* 77, 6181–6188.

Zeikus, J. G., and Wolfe, R. S. (1972). *Methanobacterium thermoautotrophicus* sp. n., an anaerobic, autotrophic, extreme thermophile. *J. Bacteriol.* 109, 707–713.

APPENDIX

Table A1 | Community diversity estimated from 16S sequence libraries.

Site_N umber	N	99% OTUs	Singletons	ACE	Chao1	SD	95U	95D	alpha	SD	Shannon	Simpson
BLVA_5	305	69	32	120.9	96.6	12.9	134.8	80.5	27.8	2.5	3.2	10.4
BLVA_20	364	130	102	773.9	873.1	318.3	1791.1	462.5	72.3	6.0	3.8	14.1
White Creek_6	367	69	34	131.3	141.3	36.6	253.3	97.3	25.1	2.1	3.4	19.2
Chocolate Pots_7	380	127	100	402	421.1	94.0	669.1	286.6	66.9	5.4	3.4	9.2
Mushroom Spring _15 (Bacteria)	314	112	66	214.1	220.9	37.7	322.5	168.3	62.3	5.6	4.1	37.5
Mushroom Spring _15 (Archaea)	265	35	16	65	77.7	33.2	199.5	46.1	10.8	1.1	2.7	8.9
Fairy Geyser_16	360	137	96	390.3	408.1	87.5	639.4	283.2	80.7	6.8	4.1	32.8

Richness indexes ACE, Chao1 (w/95% confidence intervals) and diversity indexes (Fisher's alpha, Shannon-Weaver, and Simpson's Index).

FIGURE A1 | Fragment recruitment of metagenome sequence reads from six phototrophic mat samples to the most relevant reference genomes (*y*-axis = 50–100% (nt); *x*-axis = genome position) currently available in reference databases. Site Colors: *White Creek* (WC_6) = blue; *Chocolate Pots* (CP_7) = brown; *Mushroom Spring* (MS_15) = green; *Fairy Geyser* (FG_16) = red; *Bath Lake Vista Annex* (BLVA_5 and 20) = light and dark purple, respectively. Reference genomes: *Synechococcus* sp. B′ (JA-2-3B′a) and A (JA-3-3Ab); *Chloroflexus* sp Y-400-fl and *C. aurantiacus* J-10-fl; *Roseiflexus* sp RS-1 and *R. castenholzii* DSM 13941 (Plot constructed using JCVI bioinformatic utilities; Rusch et al., 2007).

FIGURE A2 | BLASTN recruitment of White Creek (WC_6) metagenome sequence reads to representative cyanobacterial genomes.
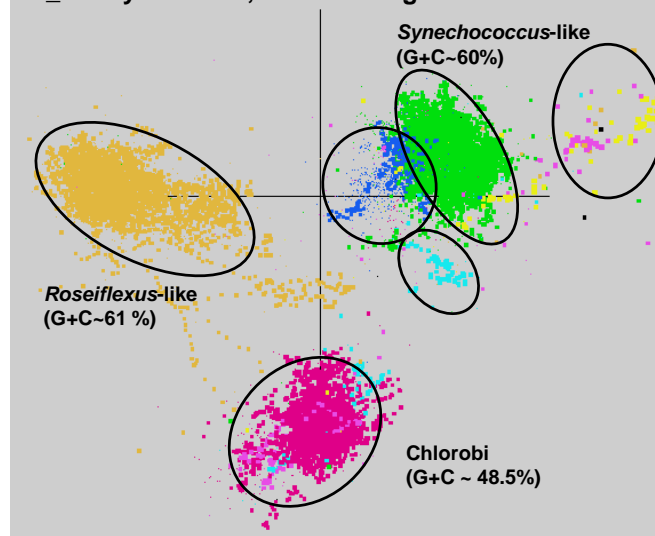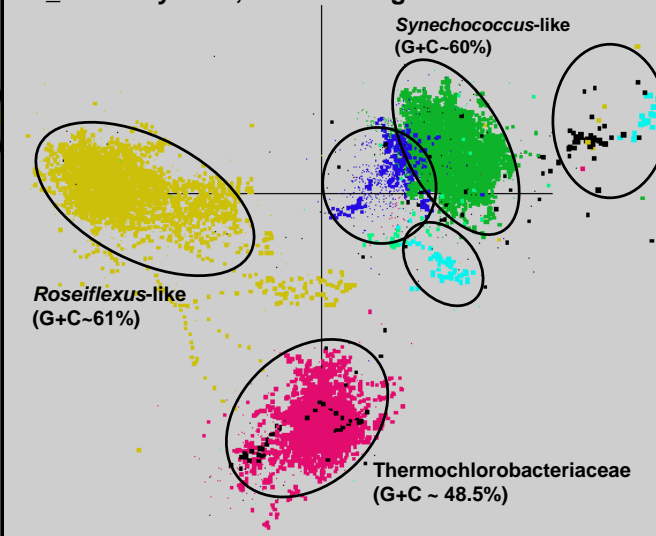
FIGURE A3 | Nucleotide word-frequency PCA of assembled sequence from Chocolate Pots (CP_7). This community contains predominant phylotypes of *Roseiflexus*-, *Synechococcus*-, "*Ca.* T. aerophilum"- (Chlorobi), and *Leptospirillum*-like populations as well as minor contributions from the Firmicutes, Proteobacteria, and Bacteroidetes [green = Cyanobacteria

(*Synechococcus* spp.); gold = Chloroflexi (*Roseiflexus* spp.); maroon = Chlorobi/Thermochlorobacteriaceae/"*Ca.* T. aerophilum", dark-blue = Spirochetes/*Leptospiraceae*; light-blue = Bacteroidetes/*Flexibacteraceae*; yellow = Firmicutes; light-purple = Proteobacteria].

FIGURE A4 | Maximum-likelihood phylogenetic tree of Chlorobi PscD sequences from metagenomes and sequenced genomes. The tree was produced using the program PhyML from PscD alignments containing 172 positions, with amino acid comparisons corrected using the Jones-Taylor-Thornton model (Jones et al., 1992). Bootstrap support for ≥50% of 1000 replicate trees are shown at nodes.

FIGURE A5 | G + C Composition of Scaffold Clusters. Scaffold clusters greater than 10 kbp were demarcated using oligonucleotide frequencies as depicted in **Figure 4**.

FIGURE A6 | Predominant metagenome sequence assemblies [(A) >10 kb contigs and (B) >20 kb contigs only] from Bath Lake Vista Annex (BLVA_20) sampled during a bloom of purple-sulfur bacteria (Gamma-proteobacteria). Major clusters are identified at the class-level (purple = Gamma-proteobacteria; gold = class Chloroflexi; yellow = Clostridia, light-blue = Nitrospira; green = Acidobacteria; maroon = Betaproteobacteria; dark-blue = Deltaproteobacteria).

FIGURE A7 | Neighbor-joining phylogenetic tree of Chloroflexi 16S rRNA sequences from all clone libraries. (A) Sub-branch of tree corresponding to *Roseiflexus* spp. **(B)** Sub-branch of tree corresponding to FAPs related to *Chloroflexus* spp. and other organisms capable of producing bacteriochlorophyll *c*. Sequences are color coded according to spring origin, and numbers adjacent to polygons indicate the number of clones in each clade. Bootstrap support for ≥50% of 1000 replicate trees are shown at nodes.
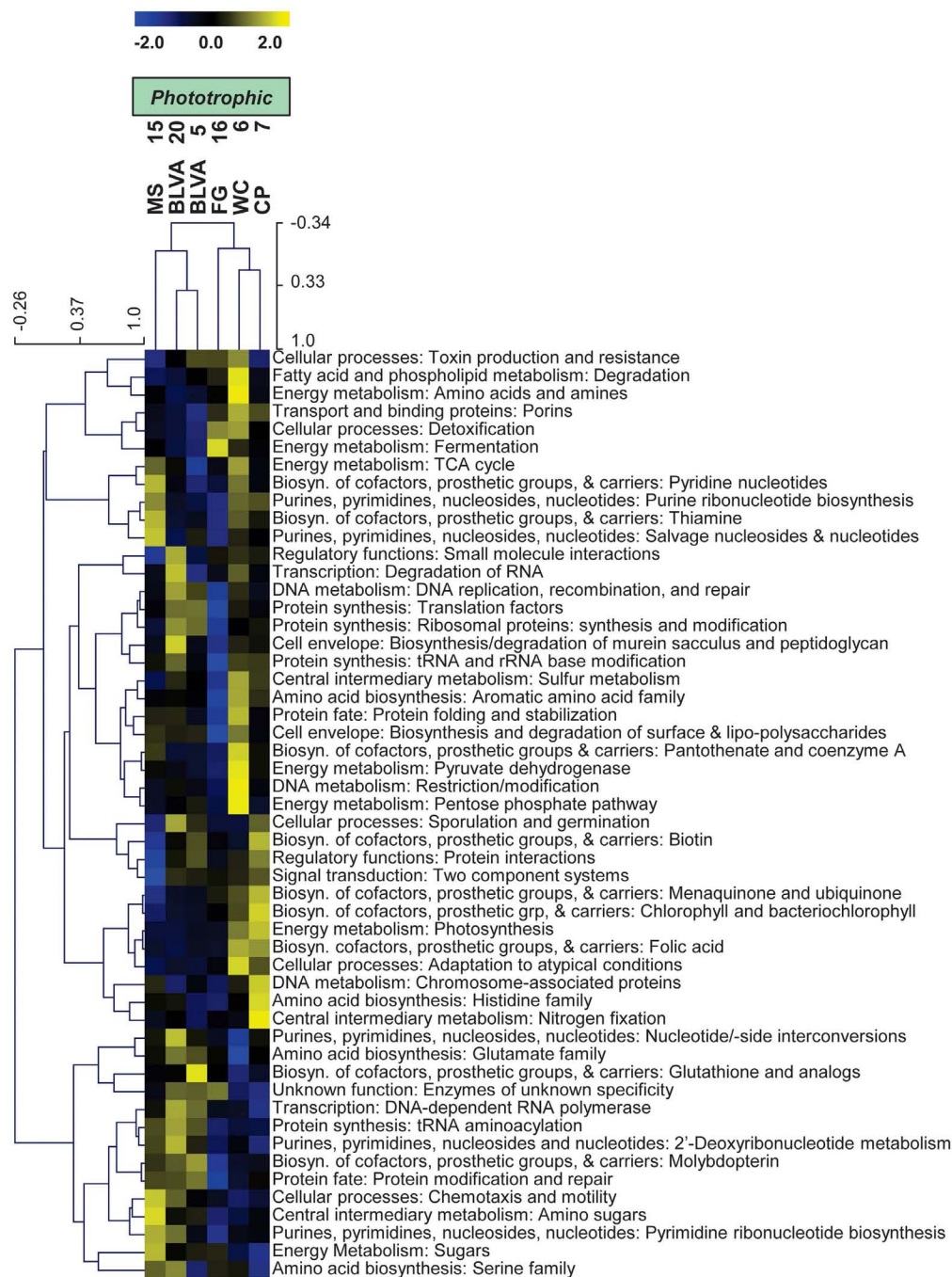
FIGURE A8 | Hierarchical cluster analysis of relative abundances of genes in all TIGRFAMs grouped into functional categories. Data were standardized by functional category before clustering to avoid biasing analysis by a few categories with high gene abundance. Pearson correlation was used as the distance measure for average linkage agglomerative clustering.

The epsomitic phototrophic microbial mat of Hot Lake, Washington: community structural responses to seasonal cycling

Stephen R. Lindemann[1], James J. Moran[2], James C. Stegen[1], Ryan S. Renslow[3], Janine R. Hutchison[2], Jessica K. Cole[1], Alice C. Dohnalkova[3], Julien Tremblay[4], Kanwar Singh[4], Stephanie A. Malfatti[4], Feng Chen[4], Susannah G. Tringe[4], Haluk Beyenal[5] and James K. Fredrickson[1]*

[1] Biological Sciences Division, Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA
[2] Chemical, Biological, and Physical Sciences Division, National Security Directorate, Pacific Northwest National Laboratory, Richland, WA, USA
[3] Scientific Resources Division, William R. Wiley Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA
[4] Lawrence Berkelely National Laboratory, Joint Genome Institute, Walnut Creek, CA, USA
[5] The Gene and Linda Voiland School of Chemical Engineering and Bioengineering, Washington State University, Pullman, WA, USA

Phototrophic microbial mats are compact ecosystems composed of highly interactive organisms in which energy and element cycling take place over millimeter-to-centimeter-scale distances. Although microbial mats are common in hypersaline environments, they have not been extensively characterized in systems dominated by divalent ions. Hot Lake is a meromictic, epsomitic lake that occupies a small, endorheic basin in north-central Washington. The lake harbors a benthic, phototrophic mat that assembles each spring, disassembles each fall, and is subject to greater than tenfold variation in salinity (primarily Mg^{2+} and SO_4^{2-}) and irradiation over the annual cycle. We examined spatiotemporal variation in the mat community at five time points throughout the annual cycle with respect to prevailing physicochemical parameters by amplicon sequencing of the V4 region of the 16S rRNA gene coupled to near-full-length 16S RNA clone sequences. The composition of these microbial communities was relatively stable over the seasonal cycle and included dominant populations of *Cyanobacteria*, primarily a group IV cyanobacterium (*Leptolyngbya*), and *Alphaproteobacteria* (specifically, members of *Rhodobacteraceae* and *Geminicoccus*). Members of *Gammaproteobacteria* (e.g., *Thioalkalivibrio* and *Halochromatium*) and *Deltaproteobacteria* (e.g., *Desulfofustis*) that are likely to be involved in sulfur cycling peaked in summer and declined significantly by mid-fall, mirroring larger trends in mat community richness and evenness. Phylogenetic turnover analysis of abundant phylotypes employing environmental metadata suggests that seasonal shifts in light variability exert a dominant influence on the composition of Hot Lake microbial mat communities. The seasonal development and organization of these structured microbial mats provide opportunities for analysis of the temporal and physical dynamics that feed back to community function.

Keywords: Hot Lake, phototrophic microbial mats, 16S tag sequencing, phylogenetic turnover, microbial diversity, seasonal cycling, community assembly, magnesium sulfate

INTRODUCTION

Microbial mats are macroscale communities of metabolically linked organisms (Taffs et al., 2009; Klatt et al., 2013) occupying a shared biogenic ultrastructure typically composed of an organic exopolymeric matrix (Decho et al., 2005; Braissant et al., 2009). As such, microbial mats exist as entire ecosystems where complete energy and element cycles, otherwise taking place over large distances, occur on millimeter scales (reviewed in Franks and Stolz, 2009; Paerl and Yannarell, 2010). Consequently, the diverse metabolic activities of the community members impose steep physical and chemical gradients and create niches with fine spatiotemporal resolution (Dupraz et al., 2009). Sunlight drives strong vertical community structuring of phototrophic microbial mats as photons of specific wavelengths are selectively harvested with depth (e.g., Pierson et al., 1987; Jorgensen and Des Marais, 1988). These mats experience significant variation in their physicochemical environments and, therefore, the interspecies interactions operating within them, as light availability changes over diel and seasonal cycles (Van der Meer et al., 2005; Villanueva et al., 2007; Steunou et al., 2008; Dillon et al., 2009). Cyanobacteria often populate the upper, photic areas in these mats where they capture solar energy, fixing carbon and producing O_2 as a byproduct; both products of photosynthesis are subsequently cycled by heterotrophs (Paerl et al., 2000). Cyanobacterial mats are common in hypersaline systems worldwide, where elevated salinities restrict grazers and allow accretion of biomass (Oren, 2010).

The biology of mat communities occupying athalassohaline environments, especially those dominated by Mg-Na-SO₄

brines, remains understudied considering their widespread global occurrence. Epsomitic hypersaline lakes and playas are common features of the inter-range semi-arid plateau between the Rocky Mountains and the Pacific Coast and Cascade Ranges that stretches from eastern Washington and Oregon through British Columbia (Bauld, 1981; Renaut, 1990) and within the endorheic Ebro Basin in Spain (Guerrero and De Wit, 1992; Jonkers et al., 2003). Insofar as athalassohaline mat systems in western North America have been studied in detail, the focus has predominantly been on carbonate precipitation (Renaut, 1993; Power et al., 2007, 2009) or the detection of signatures of life from an astrobiological aspect (Foster et al., 2010).
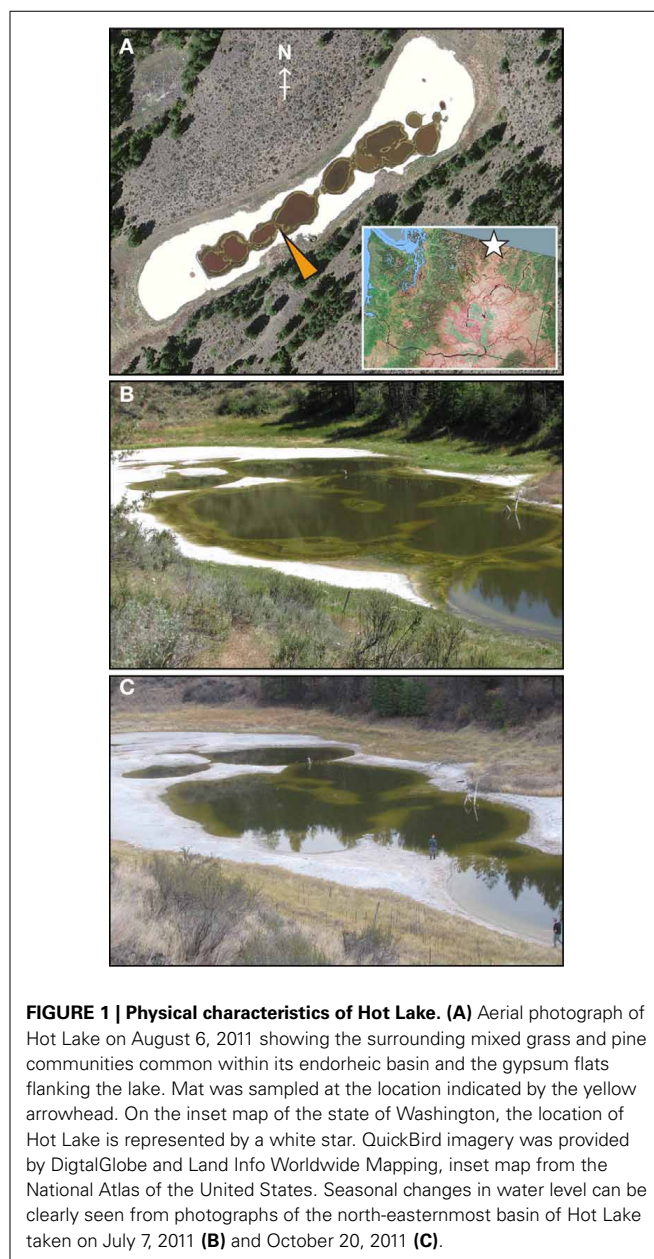
Hot Lake is a heliothermal hypersaline lake in extreme north-central Washington near Oroville that seasonally harbors a benthic phototrophic microbial mat. It is constrained within a glacially-carved, endorheic basin in a semi-arid climatic zone. The basin drains less than a one-half square mile ($<1.3\,km^2$) watershed and is underlain by metamorphic rock, dolomites, and shales (Jenkins, 1918). Sulfuric acid produced by the oxidation of neighboring pyrite and pyrrhotite deposits releases magnesium, calcium, and sulfate ions that are transported into the lake (Jenkins, 1918). Due to the endorheic nature of the basin, these salts accumulate in Hot Lake, causing it to become magnesium-dominated as calcium sulfate precipitates. Positive net evaporation through the summer and early fall months causes significant decreases in water volume and concurrent increases in salinity (Anderson, 1958). Hot Lake is meromictic with a relatively fresh mixolimnion situated atop a more saline monimolimnion (Anderson, 1958; Walker, 1974). The lake exhibits an inverse thermal gradient, sometimes exceeding 50°C at maximum, due to peak light absorption in the upper monimolimnion and insulation by the overlying water column (Anderson, 1958). Retention of heat in the monimolimnion likewise causes some of the warmest solar-heated waters on Earth at Solar Lake (Cohen et al., 1977), which is also home to a well-studied benthic cyanobacterial mat that exhibits seasonal cycling (Krumbein et al., 1977; Jorgensen et al., 1979, 1986; Frund and Cohen, 1992).

Hot Lake has previously been studied in some detail for its unique limnology and geology (Handy, 1916; Jenkins, 1918; Anderson, 1958; Bennett, 1962; Walker, 1974), as well as for the flora (St. John and Courtney, 1924; McKay, 1935; Anderson, 1958) and fauna (Anderson, 1958; Broch, 1969) that inhabit the lake. Recent work also includes the microbes of its marginal soils (Kilmer et al., in press). To date, however, only Anderson mentioned Hot Lake's benthic microbial mat; his study identified the mat's cyanobacteria but did not characterize the non-cyanobacterial microbial populations inhabiting the mat. In 1955, the mat was present at depths exceeding 1.0 m and extended into the upper reaches of the monimolimnion (Anderson, 1958, 2012). In this work, we interrogate the mat community's compositional variation as it responds to the highly dynamic environmental conditions of Hot Lake throughout the seasonal cycle of 2011. Additionally, we examine the community's phylogenetic turnover with respect to the environmental metadata and infer processes driving community variation.

MATERIALS AND METHODS

SAMPLING AND ENVIRONMENTAL CHARACTERIZATION

Benthic mat samples were collected on April 21, July 7, September 1, and October 20, 2011 at the same sampling station, located at 48.973062°N, 119.476876°W at an elevation of ~576 m. Mixolimnion water and ice were also collected from the same location on December 1, 2011. The collected mat was visually representative of mat observed ringing the entire lake (See **Figure 1A**). The mat was operationally defined as the portion that remained intact when lifted off the underlying sediments and was typically 3–5 mm in thickness. Two samples of mat (~50 cm^2 each) were collected per time point, cryoprotected by immersion in 2.3 M sucrose, and immediately frozen on dry ice. Mat collected for microscopic analysis was fixed in the field with 4%

FIGURE 1 | Physical characteristics of Hot Lake. (A) Aerial photograph of Hot Lake on August 6, 2011 showing the surrounding mixed grass and pine communities common within its endorheic basin and the gypsum flats flanking the lake. Mat was sampled at the location indicated by the yellow arrowhead. On the inset map of the state of Washington, the location of Hot Lake is represented by a white star. QuickBird imagery was provided by DigtalGlobe and Land Info Worldwide Mapping, inset map from the National Atlas of the United States. Seasonal changes in water level can be clearly seen from photographs of the north-easternmost basin of Hot Lake taken on July 7, 2011 **(B)** and October 20, 2011 **(C)**.

paraformaldehyde in lake water and held at 4°C for at least 24 h to ensure complete fixation. Paired 50-mL water samples were taken from the same depth as sampled mat. The water temperature was immediately recorded using a WTW 3400i Multi-Parameter Field Meter (WTW, Inc., College Station, TX) prior to storing the samples at 4°C. Samples were then filtered in the laboratory using a 0.22-μm polyethersulfone syringe filter unit (EMD Millipore, Billerica, MA) and held at 4°C. Filtrate was assayed for pH, total dissolved solids (TDS), alkalinity, major cations (magnesium, sodium, potassium, calcium), major anions (sulfate, chloride), dissolved organic carbon (DOC), nitrate, ammonium, and o-phosphate by Huffman Laboratories (Golden, CO). Irradiance data were obtained from OVLW1, a remote automated weather station ~1.5 km from Hot Lake at an elevation of ~440 m. These data were provided by the U. S. Bureau of Land Management & Boise Interagency Fire Center and hosted by MesoWest, a project of the Department of Atmospheric Sciences at the University of Utah (http://mesowest.utah.edu).

FIBER OPTIC MICROPROFILING

A custom-built fiber optic microprobe was used to quantify light penetration by depth in the mat. The fiber optic microprobes, which had tapered tips, were formed using a variation on previously described techniques (Gao et al., 1995; Beyenal et al., 2000), as detailed in Lewandowski and Beyenal (2007). Briefly, the insulation near the tip of a 9 μm core fiber with numerical aperture (NA) of 0.11 (Corning® SMF-28® ULL optical fiber, Corning, NY, USA) was mechanically stripped, the tip cleaned with isopropyl alcohol and cleaved. The cleaved fiber was held vertically in a precision linear positioner and lowered into unstirred, 37.5% hydrofluoric acid. After etching for 2–15 min at room temperature, the fiber was removed and rinsed in deionized water. Ambient light intensity was then measured; stable, reproducible readings and inspection using a scanning electron microscope indicated the successful formation of the fiber tip. The optical fiber cable was connected using an FC connector to an Ocean Optics Torus Miniature Spectrometer (Dunedin, FL, USA). The fiber optic microprobe was placed on a micromanipulator controlled by a stepper motor controller (PI M-230.10S Part No. M23010SX, Physik Instrumente, Auburn, MA, USA) and custom Microprofiler® software. Spectra were taken at 0.25-mm increments throughout the mat. Light intensity directly above the mat surface was recorded as a reference. Intensity values at different depths are reported as percent transmission relative to the surface illumination for each wavelength.

CRYOSECTIONING

To prepare cross sections for microscopic analysis, paraformaldehyde-fixed mat was cryoprotected overnight with 2.3 M sucrose at 4°C. Blocks of cryoprotected mat were excised from the center of the mat sample and embedded in Tissue-Tek O.C.T. Compound in Tissue-Tek 10 × 10 × 5-mm Cryomolds (Electron Microscopy Sciences, Hatfield, PA). 50-μm-thick sections were cut using a Leica CM1520 cryostat, transferred to slides, mounted in VECTASHIELD mounting medium (Vector Laboratories, Burlingame, CA) and imaged using a Nikon Optiphot-2 epifluorescence microscope. Depth-resolved sections

for DNA extraction were prepared by excising blocks from the center of frozen, cryoprotected mat samples and embedding them in O.C.T. Compound in 25 × 20 × 5-mm Tissue-Tek Cryomolds (Electron Microscopy Sciences, Hatfield, PA) oriented with the pinnacled top facing up. The mat was then sectioned into 50-μm-thick sections, 10 of which were pooled to span 500 μm total depth, and nucleic acids were extracted as described below.

SUBSAMPLING AND DNA EXTRACTION

Seasonal cycling of the mat community was examined by extracting genomic DNA from cryoprotected samples of whole mat. Frozen, cryoprotected mat was subsampled into 3 × 3 grids, each of the nine subsamples being 0.5 mm on a side. For each time point, three sections of each grid were randomly chosen from each of two plates using a random number generator (www.random.org, see **Figure 5**A). DNA was extracted according to the enzymatic protocol (EP) previously described (Ferrera et al., 2010) with the following modifications: prior to purification, ~100-mg mat samples were washed with molecular biology grade 0.5 M EDTA at pH 8.0 (Life Technologies, Carlsbad, CA) to remove excess magnesium and resuspended in lysis buffer (50 mM Tris at pH 8.0, 25 mM EDTA pH 8.0). Samples were then incubated at 85°C for 5 min to inactivate native nucleases and slowly cooled to 37°C. Chemical and enzymatic lysis then proceeded as described by Ferrera et al. Briefly, samples were treated with 1 mg/ml lysozyme at 37°C for 45 min, at which point 1:10 vol 10% SDS and 0.2 mg/ml proteinase K were added prior to incubation at 56°C for 1 h. Post-lysis, DNA was extracted with phenol-chloroform-isoamyl alcohol (25:24:1, vol:vol:vol) and then chloroform-isoamyl alcohol (24:1). Sodium acetate at pH 5.5 was added to a final concentration of 0.3 M. The DNA was then precipitated in 50% isopropanol, washed in 70% ethanol, dried, and resuspended in TE buffer (10 mM Tris-HCl at pH 8.0, 1 mM EDTA). DNA was extracted from cryosectioned laminar sections using the same protocol with the exception that, prior to EDTA washing, samples were washed three times with 50 mM Tris at pH 8.0 in 25% sucrose to remove the O.C.T. Compound.

CLONE LIBRARY CONSTRUCTION, SEQUENCING, AND PROCESSING

Near-full-length *rrnA* genes were PCR amplified from genomic DNA harvested from a ~25-mm^2 (238 mg) whole-mat sample collected on July 7, 2011 using universal bacterial primers 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1492R (5′-GGYTACCTTGTTACGACTT-3′) (Lane, 1991). PCR was performed using Phusion polymerase (New England Biolabs, Ipswitch, MA) in HF Buffer and 3% dimethyl sulfoxide according to the manufacturer's instructions at an annealing temperature of 55°C for 27 cycles. PCR product was cloned using the Zero Blunt TOPO PCR cloning kit (Life Technologies, Carlsbad, CA) according to the manufacturer's directions. Plasmids were isolated from clones and their 16S rRNA genes were sequenced using Sanger dideoxy chain-termination sequencing by Functional Biosciences (Madison, WI) from pCR-II-TOPO's SP6 and T7 promoter regions. Using the ContigExpress algorithm of Vector NTi Advance v. 11.0 (Life Technologies, Carlsbad, CA), sequence ends were trimmed until the initial and final 25 bases contained no ambiguities or bases with a Phred quality score of

less than 20. Sequences were then checked for vector contamination and assembled into contigs. Assemblies were curated and mismatches resolved manually.

Post assembly, sequences were aligned using the mothur-formatted SILVA-based bacterial reference alignment (http://www.mothur.org/w/images/9/98/Silva.bacteria.zip, updated April 22, 2012) in mothur v. 1.29 (Schloss et al., 2009). These aligned sequences were filtered to remove non-informative columns and clustered to account for the expected error for a Phred score of 20 (1%, allowing 12 differences across the alignment). Sequences were then checked for chimeras using UCHIME (Edgar et al., 2011) as implemented in mothur 1.29 both in self-referential mode and using the SILVA gold alignment as a reference. Chimeras detected using the reference sequences were manually examined to prevent the inadvertent removal of sequences without good reference sequences. Near-full-length clones that were observed at least twice in the clone library (at >99% identity) or that mapped >0.1% of the Itag sequences were chosen for more thorough analysis, and these were manually examined for chimeras prior to submission to GenBank (see Supplemental Table 1 for accession numbers). The full-length, 50,000-column alignment of these representative sequences was incorporated into the reference alignment used in the Itag analysis in order to promote the alignment of Itag sequences similar to these near-full-length sequences. In addition, it was degapped and used as a reference to map Itag sequences (see following section).

Itag SEQUENCING

Short 16S rRNA tag (Itags) sequencing was done on an Illumina MiSeq instrument at the Joint Genome Institute, Walnut Creek, CA. Primer design for universal amplification of the V4 region of 16S rDNA was based on a protocol published by Caporaso and co-workers (Caporaso et al., 2011). The forward primer (515F, 5′- AATGATACGGCGACCACCGAGATCTACAC TATGGTAATT GT GTGCCAGCMGCCGCGGTAA) remained unchanged and the barcoded reverse primers are largely similar to the Caporaso V4 reverse primer (806R), but with 0–3 random bases and the Illumina sequencing primer binding site added between the amplification primer and the Illumina adapter sequence. For each sample, three separate 16S rRNA amplification reactions targeting the V4 hypervariable region were performed, pooled together, cleaned up using AMPureXP (Beckman Coulter) magnetic beads, and quantified with the Qubit HS assay (Invitrogen). Some samples were also analyzed with a BioAnalyzer 2100 (Agilent) instrument to confirm appropriate amplicon size. Pooled amplicons were then diluted to 10 nM and quantified by qPCR prior to sequencing according to JGI's standard procedures. A total of 3,184,278 (1,592,139 forward and 1,592,139 reverse reads) barcoded paired-end reads where obtained after computational removal of PhiX and contaminant reads (reads containing Illumina adapters). Reads were then paired-end assembled using FLASH (Magoc and Salzberg, 2011). All sequences were then trimmed from both 5′ and 3′ ends using a sliding window of 10 bp and quality score threshold of 33. Reads having more than 5 ambiguous bases, an average quality score lower than 30, or more than 10 nucleotides having a quality score lower than 15 were

rejected. We ended with a total of 1,634,356 quality-filtered tag sequences that were used for downstream analyses.

Itag SEQUENCE PROCESSING AND ANALYSIS

Sequences were processed using mothur v. 1.29 as previously described (Schloss et al., 2011), though some modifications were made to accommodate 2 × 250 cycle paired-end MiSeq sequences, and 454-specific portions of the protocol were omitted. Paired, phiX-decontaminated reads were sorted into samples by barcode using a custom Perl script and mothur-formatted FASTA and group files were generated. The FASTA file was filtered to remove those sequences with ambiguities or those with lengths shorter than 251 nts. Thereafter, processing closely followed the protocol of Schloss et al. (2011), with the exception that sequences were subsampled prior to distance matrix generation. Briefly, sequences were aligned to the SILVA-based bacterial reference alignment, which was augmented with the Hot Lake mat near-full-length sequences (see Clone library construction, sequencing, and processing above). Sequences were then screened to remove those that did not align to positions 13871–23444 of the reference alignment, filtered to remove non-informative columns, pre-clustered to >99% identity (allowing two differences), and dereplicated. Sequences were then checked for chimeras using UCHIME as implemented in mothur 1.29 in self-referential mode and identified chimeras were removed.

The resulting set of filtered sequences was then classified using a Wang (Bayesian) approach with the Ribosome Database Project training set v. 9 (updated March 20, 2012 and formatted for mothur) as a reference. Sequences of unknown classification at the kingdom level were removed. Each group was then subsampled to the size of the smallest group (14,562 sequences). Sequences were clustered into operational taxonomic units (OTUs) using an average neighbor algorithm with a 3% cutoff classified at a cutoff bootstrap value of 80%. Alpha (species observed, inverse Simpson, and Simpson evenness) and beta diversity metrics (Bray-Curtis) were computed in mothur using subsampled sequences ($n = 14.562$). Twenty clones from the Hot Lake mat clone library were selected based upon abundance in the library and representation of phyla and evenly pooled to generate a mock community. The mock community was amplified by PCR as described above and sequenced alongside the other Itag samples to compute the sequencing error rate.

SHORT READ MAPPING AND PHYLOGENY RECONSTRUCTION

Unique Itag reads were mapped to the near-full-length *rrnA* sequences from the Hot Lake mat clone library using the nucmer algorithm in MUMmer v. 3.23 (Kurtz et al., 2004). A match was defined as a minimum identity of 99% across at least 243 nts. These sequences and their corresponding counts were compared with their OTU assignments, and the percent of the reads composing each OTU that mapped to each sequence from the clone library was calculated. Near-full-length sequences that were mapped by >1% of the reads of the most abundant OTUs and their near neighbors were aligned. A phylogeny was then reconstructed using a neighbor-joining algorithm assuming a maximum composite likelihood substitution model, including

transitions and transversions at uniform rates among sites and pairwise deletion of gaps, within MEGA5.1 (Tamura et al., 2011). MEGA5.1 was also used to reconstruct a maximum likelihood phylogeny using the nearest-neighbor interchange heuristic and general time reversible (GTR) substitution model assuming uniform substitution rates at all sites. The robustness of both phylogenies was tested using the bootstrap method with 1000 replications.

PHYLOGENETIC NULL MODEL ANALYSIS

The OTU table was rarefied and representative sequences for each of the 993 most abundant OTUs were retrieved from the Itags, placed within a maximum-likelihood phylogeny using FastTree 2.1 (Price et al., 2010), and used to quantify Bray and Curtis (1957) dissimilarity for all between-community pairwise comparisons. Mantel tests were used to relate Bray-Curtis values to between-community environmental differences in order to evaluate the degree to which community composition varied with environmental conditions. Bray-Curtis values were related to each environmental variable independently, and significance was evaluated with a permutation-based test to control for data non-independence.

A non-significant relationship between Bray-Curtis and a given environmental variable suggests that the environmental variable being evaluated does not strongly influence community composition. All environmental variables were, however, significantly related to Bray-Curtis (see Results). Bray-Curtis analyses therefore, provided relatively little information regarding the identity of environmental variables that most strongly influence community composition. To gain more insight, we coupled turnover in the phylogenetic structure of communities with a randomization approach that provides an expected magnitude of phylogenetic turnover when community composition is governed primarily by stochastic factors (for conceptual and technical details see Stegen et al., 2012, 2013; Swenson et al., 2012).

Phylogenetic turnover was quantified as the abundance-weighted-mean phylogenetic distance among closest relatives occurring in two communities, the β-Mean Nearest Taxon Distance (βMNTD) (for details see Fine and Kembel, 2011; Webb et al., 2011; Stegen et al., 2012). To derive ecological information from phylogenetic turnover, we compared observed βMNTD to expected βMNTD under a model of stochastic community assembly. A distribution of expected values was found using 999 iterations of a randomization that moved OTU names across tips of the phylogeny.

The β-Nearest Taxon Index (βNTI) quantifies the difference between observed and expected βMNTD in units of standard deviations; negative and positive βNTI values indicate less than and greater than expected phylogenetic turnover, respectively. Stochastic aspects of community assembly are controlled for in the randomization, such that a significant increase in βNTI over increasing environmental differences provides good evidence that variation in environmental conditions causes alterations in community composition by selecting for particular OTUs (Stegen et al., 2012). To complement the Bray-Curtis analyses,
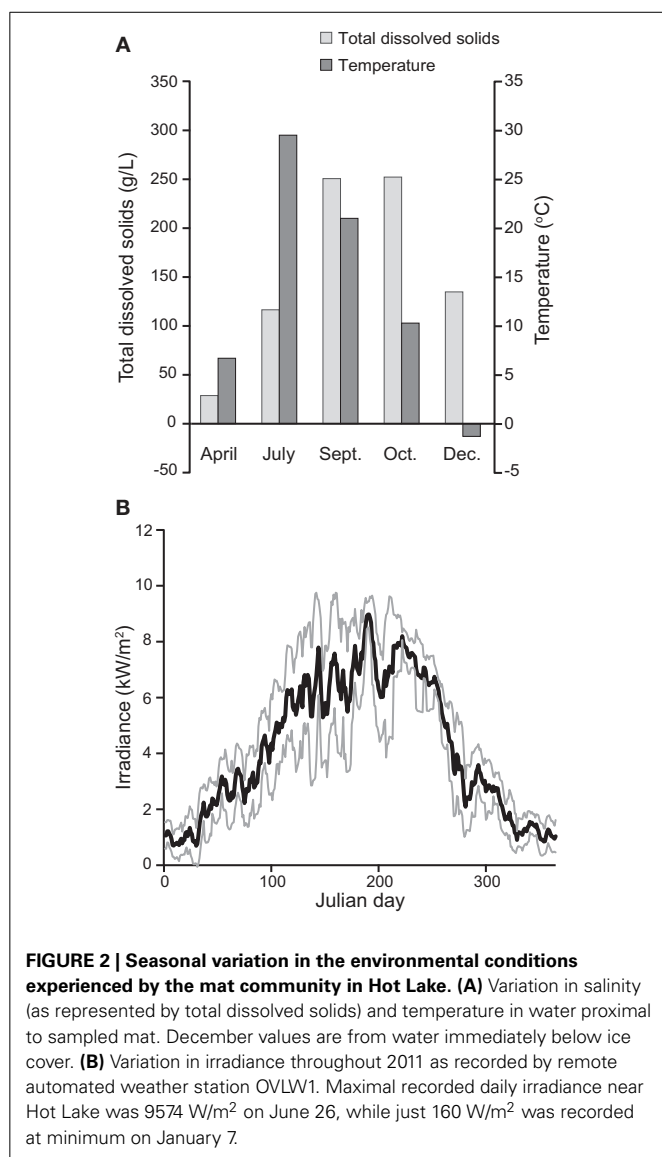
we therefore, used Mantel tests to relate βNTI to environmental variables one at a time, and permutations were used to evaluate significance.

It is important to note that the use of βNTI to arrive at ecological inferences makes the assumption that phylogenetic relationships carry ecological information. This assumption was tested using a phylogenetic Mantel correlogram (as in Stegen et al., 2013; Wang et al., 2013), which relates among-OTU ecological differences to among-OTU phylogenetic distances. OTU ecological distances were quantified as in Stegen et al. (2013) and Wang et al. (2013) using all measured environmental variables. When OTU ecological differences are significantly related to between-OTU phylogenetic distances, there is said to be "phylogenetic signal" (Losos, 2008).

RESULTS

DESCRIPTION OF HOT LAKE AND ITS PHOTOTROPHIC MAT

We observed very different water levels at Hot Lake than those described by Anderson. Most notably, the maximum water level in 2011 was ∼1 m lower than in 1955 (Anderson, 1958); surfaces that Anderson reported as submerged <1 m deep we found to be exposed and covered by fine white crystals (**Figure 1A**). Periodic descriptions of Hot Lake by others over the course of a half-century (St. John and Courtney, 1924; McKay, 1935; Walker, 1974), coupled with aerial photography, suggest that Hot Lake's water level in 2011 was more typical of modern trends. As the first half of the 1950s exhibited colder-than-average temperatures and elevated levels of precipitation in the region (Anderson, 2012), it is likely that Anderson observed Hot Lake near the upper bounds of its water volume. The white efflorescent salts on the surface of the lake's dehydrated banks, which others have previously described (Jenkins, 1918; St. John and Courtney, 1924; McKay, 1935), we determined to be primarily composed of gypsum, epsomite, hexahydrite, aragonite, and magnesite by X-ray diffraction analysis (data not shown). The salinity of Hot Lake (reported as TDS) of water collected at equal depths with the sampled mat was at its seasonal minimum in spring after significant inflow from precipitation and snowmelt (**Figure 2A**, **Table 1**). Salinity increased throughout 2011, driven by escalating evaporation and decreasing water levels (**Figures 1B,C**) over the summer and into fall. Day-to-day variability in irradiance was most strongly affected by cloud cover, which was less influential during late summer and fall than earlier in the year (**Figure 2B**). Mat-level water temperature was closely associated with irradiance (cf. **Figure 2** and Anderson, 1958). The concentrations of major cations (Mg^{2+}, Na^+, K^+), anions (SO_4^{2-}, Cl^-), and alkalinity all correlated to TDS and displayed strong evidence of evaporitic concentration throughout the seasonal cycle (**Table 1**). The DOC in Hot Lake also showed evidence of evaporative concentration, reaching 23.5 mM in mixolimnion water in September of 2011, which indicates the system was unlikely to be carbon-limited. In contrast, dissolved nitrogen sources (i.e., NO_3^- and NH_4^+) and o-phosphate concentrations were very near or below the detection limits, suggesting either may be limiting for mat growth. In the case of phosphate, this effect is likely imposed by the sparing solubility of magnesium and calcium phosphates.

FIGURE 2 | Seasonal variation in the environmental conditions experienced by the mat community in Hot Lake. (A) Variation in salinity (as represented by total dissolved solids) and temperature in water proximal to sampled mat. December values are from water immediately below ice cover. **(B)** Variation in irradiance throughout 2011 as recorded by remote automated weather station OVLW1. Maximal recorded daily irradiance near Hot Lake was 9574 W/m² on June 26, while just 160 W/m² was recorded at minimum on January 7.

community assembly-disassembly cycle again from April, 2012 to April, 2013.

The initial assembly of the mat began with stabilization of benthic sediments by a thin and gelatinous ~1 mm-thick, light-green biofilm lacking apparent lamination. As the season progressed, this biofilm matured into a coherent microbial mat characterized by a firm, rubbery texture and three to four visibly-apparent lamina (**Figure 3**). The dorsal surface layer of the mat was orange (**Figure 3A**), which microscopic examination revealed to be dominated by filamentous cyanobacteria (**Figure 3B**) occasionally interspersed with diatoms (data not shown). Over the seasonal cycle, the orange color of the surface layer intensified. Immediately below the orange layer was a ~1–2-mm thick, green layer dominated by filamentous cyanobacteria. The green layer was typically underlain by a pink layer composed of highly pigmented microclusters of microorganisms. Magnification (100X) also revealed a ~200–400 μm-thick brown layer sandwiched between the green and pink layers (**Figure 3B**). A patchy gray-black layer was occasionally observed underneath the pink layer. Inclusions of calcium and magnesium carbonates and other mineral phases were interspersed throughout the mat as observed by x-ray diffraction and electron microscopy (data not shown). Light penetration profiles measured using fiber optic microprobes revealed rapid attenuation (>99% within the first 1.0–1.5 mm, **Figure 4A**) of wavelengths strongly absorbed by chlorophyll a (with absorbance maxima of ~440 and ~675 nm) and phycocyanin (with maximum absorbance of ~625 nm). In contrast, near-infrared light ($\lambda = 805$ nm) reached the bottom of the mat, though an inflection point in the curve between 3–4 mm in depth (**Figure 4B**) suggested utilization by mat phototrophs. These transmission curves were generally consistent with vertical variations in pigmentation (cf. **Figures 3**, **4**).

COMMUNITY STRUCTURE OF THE HOT LAKE MAT AROUND THE SEASONAL CYCLE

To interrogate the spatiotemporal variability of the mat community's structure around the seasonal cycle, we collected two independent mat samples from the same location on April 21, July 7, September 1, and October 20, 2011. In each case, the mat sampled was morphologically consistent with the mature mat described above. Amplicons from the V4 hypervariable region within the 16S rRNA gene were sequenced to assay mat community structure, yielding a total of 1,470,056 assembled contigs generated from paired-end reads (exclusive of mock communities). We retained 1,207,584 quality-filtered sequences after processing. The calculated per-base error rate after completion of processing and subsampling to the size of the smallest group (14,562 reads) was ~0.029%.

To assess the spatial heterogeneity of the mat, three subsamples were randomly chosen from each of the two larger mat samples (a "plate") from every time point (**Figure 5A**), and the distance between each of these communities was compared using the Bray-Curtis β-diversity metric (Bray and Curtis, 1957). The mean distances were then compared across time points by sample interrelationship (i.e., having a shared edge or corner, subsampled from the same plate or time point, or from different time

The distribution of the microbial mat relative to depth varied over the seasonal cycle. As the year progressed, the mat gradually colonized increasingly shallower benthic surfaces, beginning near the thermocline and proceeding upward toward the water line. In April, these sediments were free of mat and consisted of a thin layer of gypsum and carbonate. At that time, the mat was present at a minimum depth of 60 cm. By early July, it had colonized approximately the lower half of the benthic surfaces in contact with the mixolimnion and proceeded to occupy all submerged sediments above the thermocline by September. The declining water level left shallow mat exposed by late summer; desiccated mat was widespread in October. Though ice covered the lake in December, frozen mat was present in the ice and immediately below it. Once ice cover receded the following April (2012), we again found no evidence of a benthic mat in the oxic mixolimnion. As the water level varied significantly throughout 2011, we used fixed points of reference to correlate mat coverage to absolute position on the lake bottom. We observed this mat

points). No significant difference ($p > 0.05$) was observed for samples collected at the same time point, no matter their spatial relationship; however, communities were significantly more closely related to others from the same time point than those from other time points ($p \sim 1 \times 10^{-27}$, **Figure 5B**). A dendrogram of the Bray-Curtis distance matrix is represented in **Figure 5C**. In general, subsamples clustered strongly with others from the same time point, though the July 2-2 subsample clustered with those from October. This clustering was likely driven by a substantially smaller number of reads in July 2-2 from OTUs otherwise observed near the bottom of the mat in July (e.g., OTU 223, 229, and 231, see **Figure 7**) as compared with other July samples. For samples collected at the same time point, inter-plate relationships were not significantly different from intra-plate relationships with the exception of the September subsamples. In this case, the difference was again driven by a significantly reduced ($p < 0.05$) abundance of reads in subsamples derived from plate 1 associated with OTUs commonly observed near the bottom of the mat (e.g., OTU 261). These data suggest either that the mat was relatively homogeneous laterally or that the heterogeneity of community structure generally occurred at a spatial resolution much smaller than the sample size ($25\,mm^2$).

The mat community remained relatively stable in composition from April to October of 2011, despite an approximately tenfold increase in salinity. At the phylum level, members of *Cyanobacteria-Chloroplast* and *Proteobacteria* dominated the community throughout the seasonal cycle (**Figure 6A**). A statistically significant ($p < 0.05$) increase in *Proteobacteria* and concomitant decrease in *Cyanobacteria-Chloroplast* was observed in September, driven largely by a spike in sequences associated with OTU 219 and a decrease in those within OTU 218, a group IV cyanobacterium (**Figures 6B, 7**). OTU 219 was classified as *Geminicoccus* (**Figures 6C, 7**), which has not been placed into a family within *Alphaproteobacteria* (*incertae familiae*, Foesel et al., 2007). Within *Cyanobacteria*, while OTU 218 was dominant throughout 2011, the group XIII cyanobacterium OTU 221 was more abundant during periods of low irradiance. In contrast, OTU 220, classified as *Cyanobacteria incertae ordinis*, exhibited a pattern inverse to that of OTU 221. Sequences were classified using the RDP Classifier (Wang et al., 2007), which groups cyanobacteria based upon molecular similarity. Cyanobacterial sequences that did not map to near-full-length 16S sequences obtained from the clone library were compared using NCBI BLAST to all non-redundant sequences in GenBank to identify the most similar cultured representative. From that analysis, the nearest cultured neighbors of OTU 221 and 220 were found to be *Phormidium* sp. UTCC 487 (99.6% identical, Casamatta et al., 2005) and *Leptolyngbya rubra* PCB9602-6 (98% identical, A. Lopez-Cortes, unpublished data), respectively. Sequences derived from diatom chloroplasts of genus *Halamphora* and *Nitzschia*

Table 1 | Aqueous geochemical analysis of mat-level waters.

			Sample collection date					
			4 April	**7 July**	**1 September**	**20 October**	**1 December**	**1 December**[a]
General water chemistry	pH	units	8.37	8.52	8.15	8.10	8.37	8.65
	TDS[b]	(g/L)	28.8	116.6	250.6	252.1	134.7	53.1
	Alkalinity	(mM)	3.00	9.19	18.08	18.78	9.16	3.80
Major cations	Mg^{2+}	(mM)	158	671	1453	1444	728	295
	Na^+	(mM)	81	318	713	718	364	146
	K^+	(mM)	12	43	103	102	49	20
	Ca^{2+}	(mM)	9	16	13	13	10	4
Major anions	SO_4^{2-}	(mM)	198	756	1758	1796	884	356
	Cl^-	(mM)	19	57	115	116	66	28
Dissolved carbon	DOC[c]	(mM C)	3.2	12.7	23.5	23.4	11.5	4.8
	Bicarbonate	(mM)	5.6	10.0	11.5	14.6	9.1	3.9
	Carbonate	(mM)	<0.8	4.2	12.3	11.5	4.6	1.8
Nutrients	Nitrate	(mM)	<0.36	<0.36	<0.36	<0.36	<0.36	<0.36
	Ammonia	(mM)	0.04	0.03	0.03	0.02	0.04	0.04
	o-phosphate	(mM)	<0.52	<0.52	<0.52	<0.52	<0.52	<0.52

Samples collected over a seasonal cycle highlight the large differences in concentration of dominant cations, anions, and dissolved carbon species. While organic and inorganic carbon concentrations remain high throughout the sampling period, nutrient levels (including nitrate, ammonia, and o-phosphate) are persistently low and may limit microbial growth. The strong correlation between TDS and most major cations and anions are consistent with evaporitic concentration. Charge balance was checked and was within 5% at every sampling point.

[a] *Melted ice from lake surface.*

[b] *Total dissolved solids.*

[c] *Dissolved organic carbon.*

(Pillet et al., 2011), were commonly observed in spring and fall but rare during the summer (**Figure 6B**).

We observed the most significant ($p < 0.05$) variations in mat community composition at the phylum level between April and July and again between September and October. Between April and July, reads attributed to *Chloroflexi* (∼41-fold), *Verrucomicrobia* (∼3.1-fold), and *Acidobacteria* (of which no reads were observed in any of the six April samples) significantly increased, while those attributed to *Firmicutes* diminished four-fold. Between September and October, *Spirochaetes*, *Chloroflexi*, and *Verrucomicrobia* reads increased ∼3.5-, 2.4-, and 2.2-fold, respectively, and reads attributed to *Actinobacteria* decreased slightly more than twofold. The rise in *Chloroflexi*

was driven almost entirely by one OTU classified within family *Anaerolinaceae* (**Figure 7A**, OTU 234); very few reads attributed to family *Chloroflexaceae* were observed. We detected less than two-fold variation in relative abundance between time points for reads attributed to all other phyla accounting for ≥0.5% of reads. While the primer set employed in this study (515F-806R) is known to broadly cover archaea (Walters et al., 2011), attributed reads did not exceed 0.5% at any point. Stability in the mat community's structure throughout 2011 was also generally observed at higher taxonomic resolution (**Figure 7**).

Within the phylum *Proteobacteria*, reads attributed to clades affiliated with sulfur cycling strongly increased from April to July and held steady throughout the summer but had decreased precipitously by mid-autumn. *Deltaproteobacteria* and *Gammaproteobacteria* diminished approximately five- and two-fold, respectively, from September to October. The majority of the loss borne by *Gammaproteobacteria* occurred in families

FIGURE 3 | Ultrastructure of the Hot Lake mat sampled on September 1, 2011. (A) Cross-section of the Hot Lake mat at the millimeter scale. Orange (O), green (G), and pink (P) lamina are readily apparent along with interspersed carbonate minerals (C). **(B)** Ultrastructure of a 50 μm-thick section of the Hot Lake mat (100X magnification).
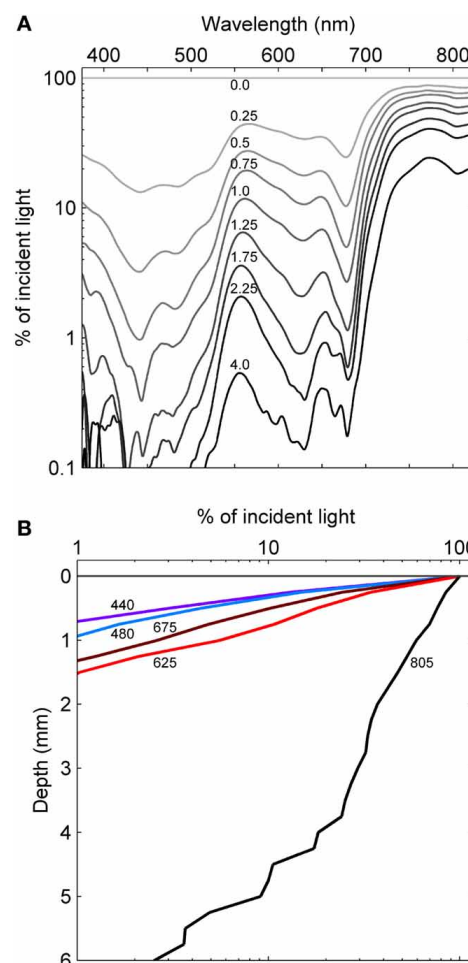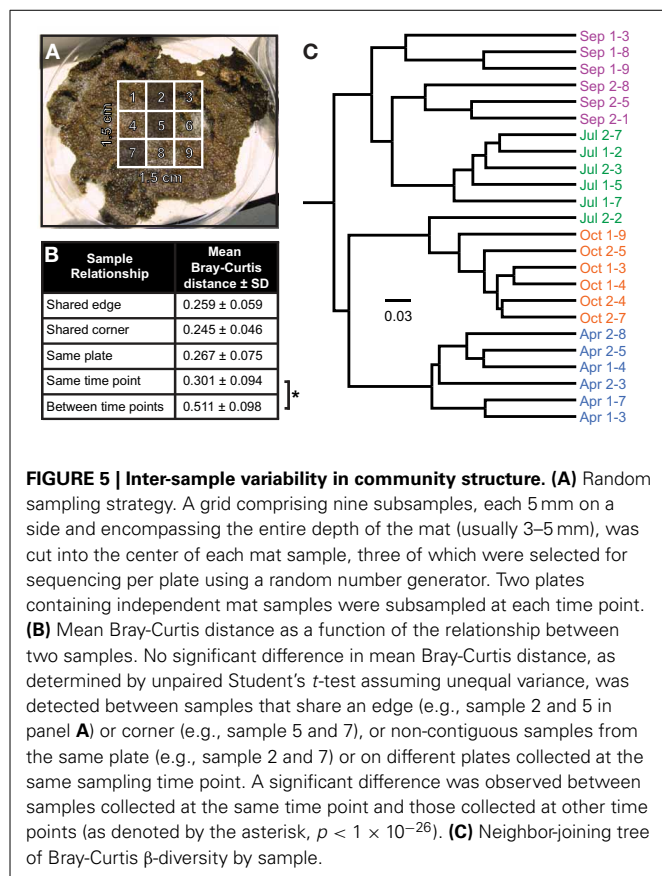

FIGURE 4 | Light penetration into the Hot Lake mat. (A) Spectrally-resolved transmission of light through the mat as measured by fiber-optic microprobe. Numbers above the curve represent the depth the probe was inserted into the mat in millimeters. **(B)** Attenuation of wavelengths representing local maxima in absorbance. Values denote wavelength in nanometers.

FIGURE 5 | Inter-sample variability in community structure. (A) Random sampling strategy. A grid comprising nine subsamples, each 5 mm on a side and encompassing the entire depth of the mat (usually 3–5 mm), was cut into the center of each mat sample, three of which were selected for sequencing per plate using a random number generator. Two plates containing independent mat samples were subsampled at each time point. (B) Mean Bray-Curtis distance as a function of the relationship between two samples. No significant difference in mean Bray-Curtis distance, as determined by unpaired Student's t-test assuming unequal variance, was detected between samples that share an edge (e.g., sample 2 and 5 in panel A) or corner (e.g., sample 5 and 7), or non-contiguous samples from the same plate (e.g., sample 2 and 7) or on different plates collected at the same sampling time point. A significant difference was observed between samples collected at the same time point and those collected at other time points (as denoted by the asterisk, $p < 1 \times 10^{-26}$). (C) Neighbor-joining tree of Bray-Curtis β-diversity by sample.

Chromatiaceae and *Ectothiorhodospiraceae*, whose members are frequently involved in sulfur cycling, diminishing approximately six- and ten-fold, respectively (**Figure 6D**). Much of this diminution was focused within OTUs 229 and 231, classified within the aerobic sulfide-oxidizing genus *Thioalkalivibrio* and the purple sulfur bacterium *Halochromatium* (**Figure 7**). Coupled with the significant concurrent reductions in phylotypes associated with sulfate reduction, as exemplified by OTU 261 (*Desulfofustis*), the data suggest that the rate of sulfur cycling may have substantially diminished between September and October.
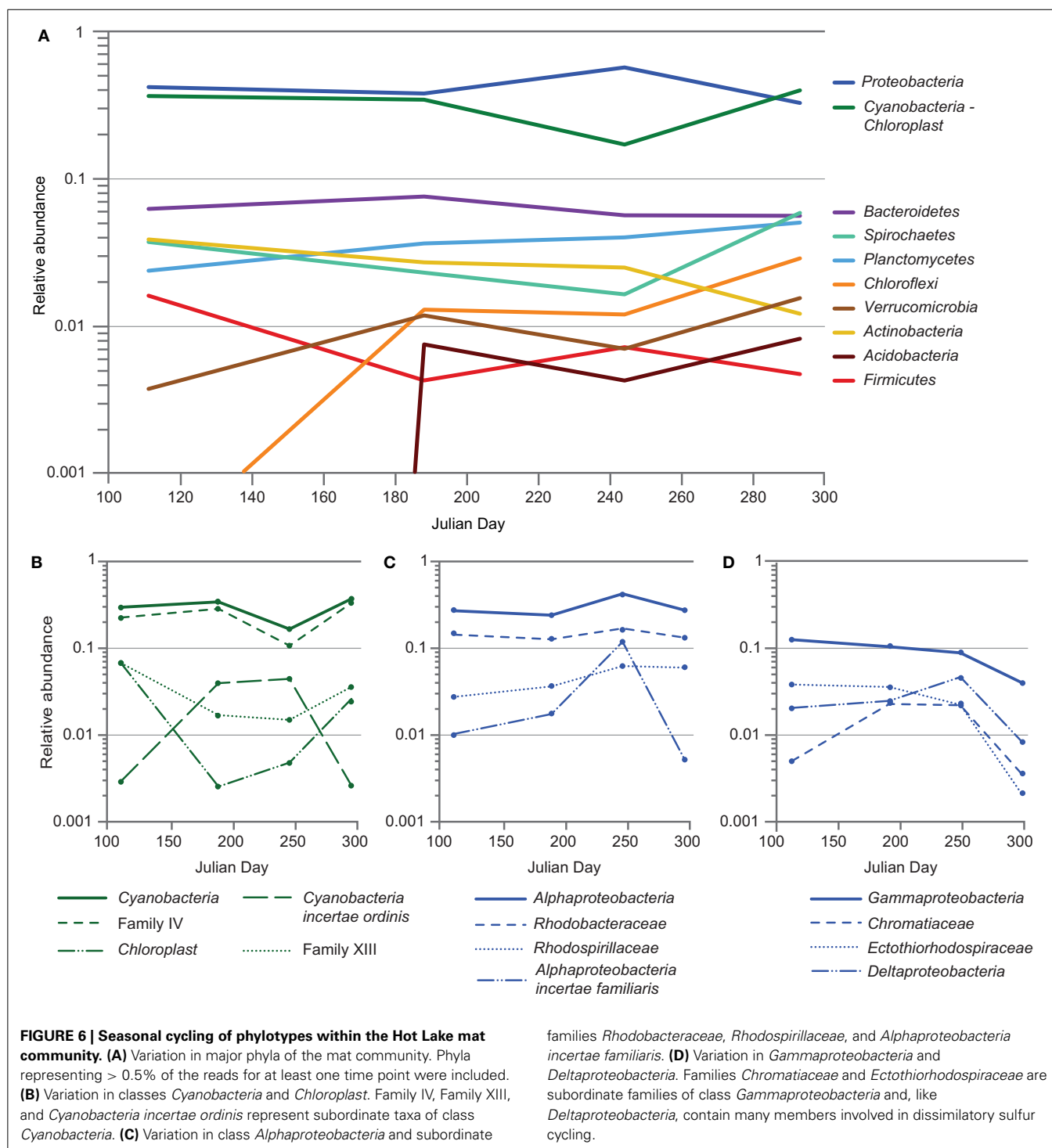
The reduction in sulfur-cycling phylotypes was part of a larger trend in reduced α-diversity in the mat community in October. After a summer season of gradually increasing trends in species observed and the inverse Simpson index, both of these metrics, as well as the Simpson evenness index, significantly declined in October (**Figure 8A**). In July, depth-resolved phylotype abundances revealed that members of *Cyanobacteria* (OTUs 218, 221, and 220) rapidly diminish between depths of 2.5 and 3.0 mm, where phylotypes associated with sulfide-oxidizing or anaerobic metabolisms increased sharply (for example, OTUs 229, 231, 261, **Figure 7B**). These data suggest the presence of a sharp chemocline at this position in the mat, which is consistent with the light profile in suggesting a termination of oxygenic photosynthesis below this depth (**Figure 4**). The bottom of the mat exhibited significantly increased α-diversity by all metrics (species observed, inverse Simpson, and Simpson evenness indices, **Figure 8B**), and OTUs observed at

or near the bottom of the mat in July were largely absent in October (**Figure 7B**). These data suggest that mat community disassembly, defined as the combined processes of biomass turnover and dispersion of cells as mat exopolymer is degraded, is associated with the loss of the phylotypes that inhabit the bottom of the mat.

In order to improve the phylogenetic resolution afforded by the short reads, we mapped them to the analogous region of the near-full-length clone sequences. The result of mapping reads from the most abundant OTUs to the clones is detailed in **Figure 7** (≥99% identity). In most cases, the OTUs were dominated by a single sequence, and thus, mapped to a single clone (e.g., OTUs 219, 225, 222). In other cases, the OTUs contained sequences mapping to several clones (e.g., OTUs 218, 226, 227) or a single clone recruited only a small fraction of the reads in an OTU (e.g., OTUs 224, 233, 232). In many cases, mapping to the clones allowed the OTUs to be classified phylogenetically at much greater resolution. A neighbor-joining phylogeny of clones mapped by the reads from the most abundant OTUs and their nearest neighbors is represented in **Figure 9**. In some cases, this mapping made otherwise inscrutable relationships apparent; for example, clones HL7711_P1F1 and HL7711_P1A2 are 98.2% identical, having only two regions of difference, one of which is covered by the Itags. Mapping to the longer clones revealed that the ratio of reads mapping to HL7711_P1F1 vs. HL7711_P1A2 is ∼2:1 in all samples examined (95% confidence interval 1.76–2.98), suggesting that they are from divergent 16S rRNA genes on the same cyanobacterial chromosome. In contrast, OTU 229 was dominated by two sequences sharing 98.8% identity and of roughly equal abundance, one for which a matching clone (HL7711_P3F6) was available, and one that did not match a clone at ≥99% identity. The ratio between these sequences was inconsistent (averaging 3.00 ± 6.53, standard deviation) and the abundance of each was decoupled in space (depth-resolved abundance in July) and time. Reads from OTU 229 mapping to HL7711_P3F6 were more abundant in April and July, while the other sequence was more frequently observed in September (data not shown), implying that these sequences represent different organisms. Collectively, these data suggest that the depth of coverage afforded by Itag sequencing allows increased dissection of the internal structure of an OTU's composite sequences. In some cases, this provides more detailed insight into the variation of individual species or ecotypes that may be combined into single OTUs by clustering algorithms.

INFERENCE OF ECOLOGICAL DRIVERS OF COMMUNITY STRUCTURE BY PHYLOGENETIC TURNOVER

We examined variation at the OTU level in light of sample meta-data using phylogenetic turnover analysis to infer the ecological parameter(s) most responsible for driving variations in community structure. One set of analyses used Bray-Curtis, which quantifies turnover in the relative abundance of OTUs. A second set of analyses used βNTI, which measures the deviation between observed and expected phylogenetic turnover, reported as βMNTD. Phylogenetic turnover (i.e., βMNTD) quantifies the difference in phylogenetic composition between a given pair of

FIGURE 6 | Seasonal cycling of phylotypes within the Hot Lake mat community. (A) Variation in major phyla of the mat community. Phyla representing > 0.5% of the reads for at least one time point were included. **(B)** Variation in classes *Cyanobacteria* and *Chloroplast*. Family IV, Family XIII, and *Cyanobacteria incertae ordinis* represent subordinate taxa of class *Cyanobacteria*. **(C)** Variation in class *Alphaproteobacteria* and subordinate families *Rhodobacteraceae*, *Rhodospirillaceae*, and *Alphaproteobacteria incertae familiaris*. **(D)** Variation in *Gammaproteobacteria* and *Deltaproteobacteria*. Families *Chromatiaceae* and *Ectothiorhodospiraceae* are subordinate families of class *Gammaproteobacteria* and, like *Deltaproteobacteria*, contain many members involved in dissimilatory sulfur cycling.

communities. For example, βMNTD will be small if OTUs within one community are closely related to the OTUs in a second community. Likewise, βMNTD will be large when OTUs within one community are distantly related to OTUs in a second community. Randomizations provide an expected level of βMNTD under the assumption that the observed magnitude of Bray-Curtis is due to stochastic changes in OTU abundances. The value of βNTI is the difference between observed and expected βMNTD. In turn, increasingly large βNTI values indicate an increasing influence of deterministic processes that select upon environmentally-determined fitness to cause differences in OTU relative abundances.

All Mantel tests relating Bray-Curtis to environmental variables were significant ($p \ll 0.05$ for all), while βNTI was
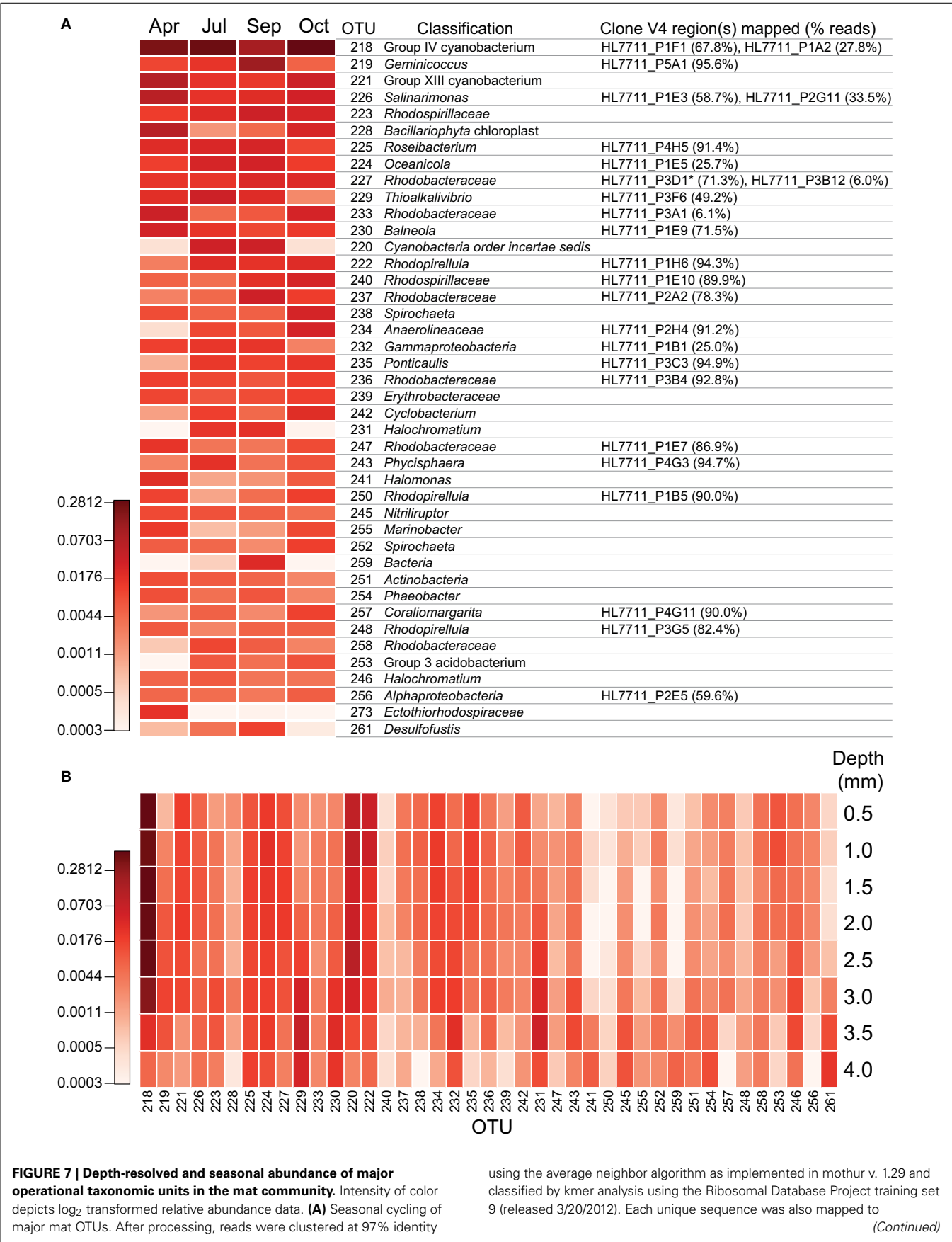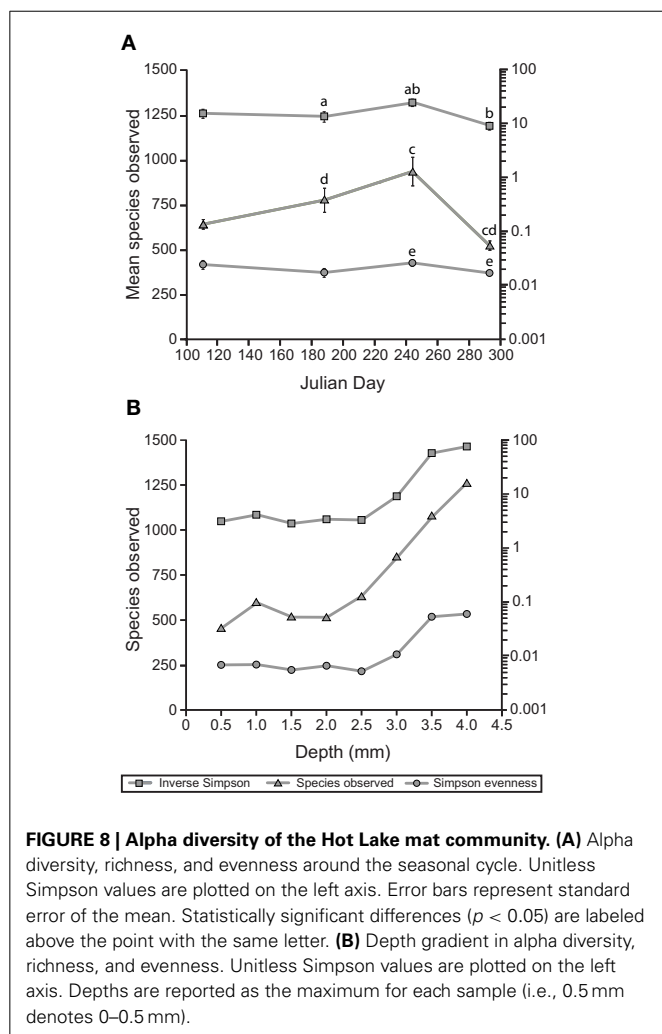
FIGURE 7 | Depth-resolved and seasonal abundance of major operational taxonomic units in the mat community. Intensity of color depicts log₂ transformed relative abundance data. **(A)** Seasonal cycling of major mat OTUs. After processing, reads were clustered at 97% identity using the average neighbor algorithm as implemented in mothur v. 1.29 and classified by kmer analysis using the Ribosomal Database Project training set 9 (released 3/20/2012). Each unique sequence was also mapped to

(Continued)

the corresponding regions of the near full-length 16S sequences using the nucmer algorithm (as implemented in MUMmer 3.23). Short reads were considered to match full-length sequences if they were >99% identical across the entire amplified region. As near full-length sequences were also classified using the same protocol as the short reads, the classification with the best taxonomic resolution or bootstrap value was reported for an OTU as

long as > 50% of its reads mapped to the corresponding near full-length sequence. *HL7711_P3D1 shares its V4 region with HL7711_P3F7 and HL7711_P3G11. **(B)** Depth-resolved abundance of major OTUs within mat sampled on July 7, 2011 and cryosectioned. OTUs are identical to those in panel A with the exception of OTU 273, which is omitted due to a lack of reads in the depth-resolved samples. Depths reported are the maxima for each sample and represent a 0.5 mm-thick laminar section.

FIGURE 8 | Alpha diversity of the Hot Lake mat community. (A) Alpha diversity, richness, and evenness around the seasonal cycle. Unitless Simpson values are plotted on the left axis. Error bars represent standard error of the mean. Statistically significant differences ($p < 0.05$) are labeled above the point with the same letter. **(B)** Depth gradient in alpha diversity, richness, and evenness. Unitless Simpson values are plotted on the left axis. Depths are reported as the maximum for each sample (i.e., 0.5 mm denotes 0–0.5 mm).

is most likely supported across short phylogenetic distances (see also Stegen et al., 2012, 2013).

Our interpretations of the Mantel test results are necessarily conservative because communities were only sampled across four points in time such that there are only four independent estimates of environmental conditions. There are, nonetheless, patterns that point toward specific environmental factors that drive variation in community composition. In particular, Mantel tests using Bray-Curtis or βNTI both suggest that temporal variation in light availability most strongly influenced the community composition of the Hot Lake mat among measured environmental variables. Two variables in the Bray-Curtis analysis had noticeably higher correlation coefficients relative to all other variables, and both were related to variation in light. Only one environmental variable was significantly (albeit very weakly) related to βNTI, and this variable was again related to variation in light. Taken together, these data suggest that the structure of the Hot Lake mat community was more strongly influenced by the dynamics of photic energy than by changes in either water temperature or salinity.

DISCUSSION

Within Hot Lake, a single mat community is annually exposed to nearly 10-fold changes in the concentrations of Mg^{2+}, SO_4^{2-}, and other dissolved ions. Although the role of increasing salinity in restricting microbial diversity and activity within mat communities has been well-established (e.g., Pinckney et al., 1995; Benlloch et al., 2002; Sorensen et al., 2005; Severin et al., 2012), relatively few studies have investigated the impact of salinity on the structure of mat communities exposed to naturally occurring salt concentration dynamics. Previous studies examining the impacts of variable salinity on community structure have frequently focused upon the sequential pools of solar salterns (reviewed in Oren, 2009) where salinity is relatively well-controlled, and high-evaporation intertidal mats such as those near Abu Dhabi (Abed et al., 2007). In the case of solar saltern systems, the mats of sequential concentrating pools are largely end-members (vis-à-vis salinity) and must be treated as discrete neighboring communities. In the case of tide pool salinity cycling, the mat community is repeatedly exposed to maximal salinity for relatively short durations. In contrast, like other microbial mats exposed to significant natural variation in salinity (e.g., Yannarell et al., 2006; Desnues et al., 2007; Yannarell and Paerl, 2007), the Hot Lake microbial mat community must annually adapt to salinity conditions ranging from brackish to extremely hypersaline.

Given that previous work (Jungblut et al., 2005; Rothrock and Garcia-Pichel, 2005; Abed et al., 2007) has demonstrated a salinity limitation on species diversity in cyanobacterial mats, we sought to determine whether the seasonally-increasing salinity of Hot Lake would promote a succession of cyanobacteria with
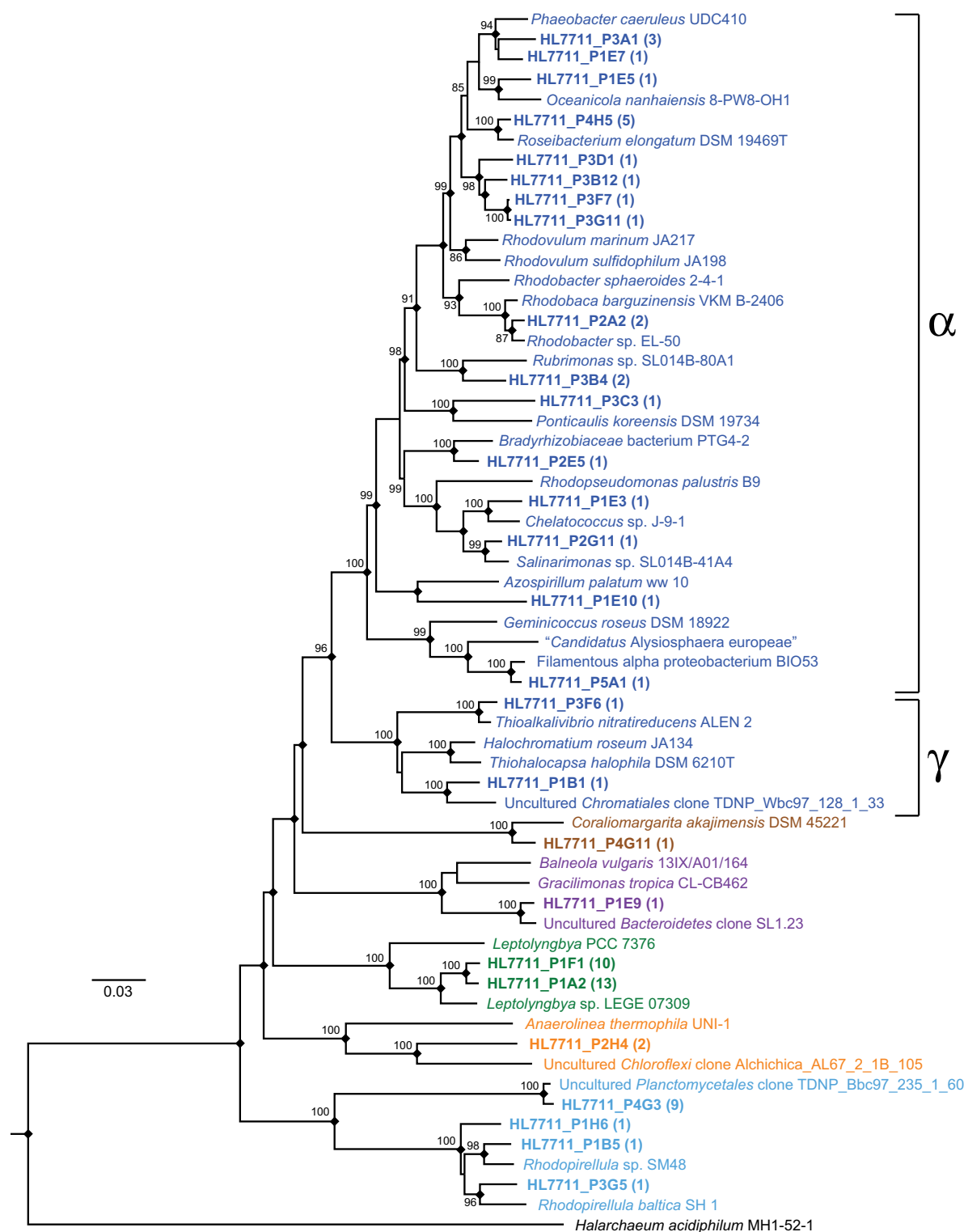
significantly related only to the standard deviation across one week of irradiance. In addition, while Bray-Curtis increased with environmental distance for all variables, βNTI decreased with increasing environmental distance for over half of the environmental variables. The use of phylogenetic turnover to make ecological inferences was supported by significant phylogenetic signal, but only within relatively short phylogenetic distance classes, as has been previously observed (Andersson et al., 2009; Stegen et al., 2012, 2013; Wang et al., 2013). Significant phylogenetic signal across short phylogenetic distances specifically supports the use of βMNTD and βNTI, as these metrics quantify phylogenetic turnover among nearest phylogenetic neighbors; our analyses indicate that the assumption of phylogenetic signal

FIGURE 9 | Phylogenetic reconstruction of near full-length 16S sequences from the Hot Lake mat representing major OTUs. Clones were generated from mat sampled on July 7, 2011 and are in bold. Clusters of sequences with >99% identity are represented by a single sequence; the number of sequences represented by each is noted parenthetically. While a neighbor-joining tree is depicted above, nodes duplicated using a maximum-likelihood algorithm employing the general time-reversible model are notated with a diamond. Values near nodes represent neighbor-joining bootstrap values greater than 80. Terminal node colors denote phyla according to the same scheme used in **Figure 6A**. Classes *Alphaproteobacteria* and *Gammaproteobacteria* are enclosed in brackets.

increasing epsotolerance (Nübel et al., 2000, cf. **Table 1**). Our data suggest, rather, that a single cyanobacterium (*Leptolyngbya*) is dominant throughout the seasonal cycle. While other, less abundant cyanobacteria and diatom chloroplasts exhibit significant seasonal variation (**Figure 7A**, OTUs 221, 228, and 220), their patterns of variation correlate more strongly with irradiance and/or temperature than with salinity. In general, the cyanobacterial species occupying the Hot Lake mat appear to be similar to those in communities observed in high-latitude and polar mats (Jungblut et al., 2005, 2009; Fernandez-Carazo et al., 2011; Kleinteich et al., 2012; Martineau et al., 2013) with dominant populations of *Phormidium* (e.g, OTU 221) and *Leptolyngbya* (OTUs 218 and 220) species. Of note is the absence of the nearly-ubiquitous mat-building cyanobacterium *Coleofasciculus chthonoplastes* (Guerrero and De Wit, 1992; Jonkers et al., 2003). Although Hot Lake cycles through salinities well known to be permissive for *Coleofasciculus*, there was no microscopic or molecular evidence for the presence of this cyanobacterium. The cyanobacteria detected in our study are consistent with the microscopic observations of Anderson and collaborators, suggesting that the same cyanobacteria may have anchored the mat community for the past 55 years despite significant changes in lake level over that time (Anderson, 1958).

In general, the non-cyanobacterial fraction of the mat community also exhibits relative stability over the course of the seasonal cycle at fine taxonomic resolution. One notable exception is the loss of OTUs likely to be involved in sulfur cycling (i.e., *Deltaproteobacteria*, and, within *Gammaproteobacteria*, families *Ectothiorhodospiraceae* and *Chromatiaceae*, **Figure 6D**) and other OTUs populating the lower regions of the mat late in the seasonal cycle. This loss occurred during a period of little change in the salinity of overlying water and contributed to reductions in species observed, Simpson evenness, and inverse Simpson metrics between late summer and late fall (**Figure 7A**). As the bottom ~1 mm of the mat is considerably more diverse than the overlying cyanobacterially-dominated laminae, elimination of niches near the mat-sediment interface has an amplified impact upon the overall α-diversity of the mat community.

Mantel tests relating βNTI to environmental variables did not support a strong linkage between community composition and salt concentration. One possible explanation is that salinity increases rapidly enough in Hot Lake throughout the seasonal cycle that organisms with broad epsotolerance are positively selected. The cyclical nature of this selective pressure may have generated a regional pool of broadly epsotolerant potential mat members that numerically dominate and, through priority effects, exclude organisms with higher fitness across a narrower range of salt concentrations. If true, this predicts that the competitiveness of mat taxa should not vary systematically across the range of salt concentrations endemic to Hot Lake. Our analyses also suggest that the observed reductions in diversity between September 1 and October 20, 2011 are part of a larger correlation between increased variation in irradiance and decreased OTU richness. Classic coexistence theory (e.g., Chesson, 2000) predicts that temporally variable environments should promote coexistence and, in turn, increase OTU richness. We related OTU

richness to the standard deviation in irradiance, which exhibited a strong negative relationship. This pattern leads us to hypothesize that community composition shifts through time, at least in part due to the exclusion of some taxa by ecological selection imposed by temporal variation in light. This selection might be mediated by the availability of metabolizable organic carbon at low or variable irradiances, as some evidence suggests cyanobacteria exude increased amounts of low molecular weight organic carbon molecules under elevated irradiances (Zlotnik and Dubinsky, 1989; Lee and Rhee, 1999).

Our ability to infer the ecological drivers of the mat community structure is limited by the temporal resolution in sampling with respect to changes in environmental parameters. For example, we measured salt concentrations of ~28, 118, 251, and 252 mg/L, such that, effectively, there were only three distinct levels evaluated in our analyses. Furthermore, because βNTI is sensitive to relative abundance, it is also possible that the high abundance and relative stability of the cyanobacterial and alphaproteobacterial compartment of the Hot Lake mat obscure the effect of important salinity-mediated changes in less abundant but functionally significant OTUs (e.g, in *Gammaproteobacteria* and *Deltaproteobacteria*). These changes in taxa associated with sulfur cycling may be harbingers or drivers of overwinter mat community disassembly processes. In order to more robustly evaluate the hypothesis that variation in light, rather than salinity, is structuring the Hot Lake microbial mat community, additional measurements of the mat community structure and local physicochemical properties taken at greater temporal resolution will be required.

Taken together, our observations suggest two hypotheses for the loss of the Hot Lake mat's relatively rich and even understructure. The first is that the seasonal reduction and increased variability of irradiance (**Figure 2B**), which result in diminished photosynthesis, reduce the amount of energy and reduced carbon available for the maintenance of heterotroph biomass. This effect might be felt both in the growth rate of primary producer biomass that can be recycled and the amount of low molecular weight photosynthate reaching the bottom regions of the mat, where dissimilatory sulfate reduction and fermentation are likely to be main metabolic strategies. Primary productivity is known to diminish with increasing salinity (Pinckney et al., 1995); this may further limit the availability of reduced carbon and nitrogen species to heterotrophs and favor the net consumption of extracellular polymers (Braissant et al., 2009).

A second hypothesis is that increasing salinities, which require energetically expensive osmotic regulation, eventually exclude species with low energy-yielding metabolisms. Although the extreme sulfate concentrations of Hot Lake make dissimilatory sulfate reduction more energetically favorable than in an equisaline NaCl environment, sulfate reduction appears to exhibit a global salinity maximum for metabolic viability (Oren, 2011). If sulfate reduction (e.g., by *Deltaproteobacteria*) is negatively impacted by elevated salinities, reductions in sulfide oxidizers within *Ectothiorhodospiraceae* and *Chromatiaceae* are likely to closely follow. As the turnover rate of organisms within the mat is unknown, the phylogenetic signals may lag significantly behind decreases in metabolic activities. This may account for the

observed change in relative abundance of these phylotypes over a period of stable salinity (September 1 to October, 20 2011). Quantifying the reaction rates of photosynthesis, sulfide oxidation, and sulfate reduction with respect to the relative abundances of associated phylotypes throughout the seasonal cycle will help to discern which of these hypotheses best explain our observations.

We expect that elucidation of the major ecological variables governing the Hot Lake microbial mat community will shed light on the environmental parameters driving its seasonal assembly and disassembly. Seasonal disassembly of a microbial mat is by no means unique to Hot Lake. Mats inhabiting diverse habitats, such as the salt marshes of Sippewissett and the North Sea barrier island beaches of Mellum, are frequently destroyed over the winter (Stal et al., 1985; Franks and Stolz, 2009), and tropical mats are known to be destroyed by hurricanes (Yannarell et al., 2007). The action of wind, waves, and tides are believed to be the primary means for the physical destruction of such microbial mats. Renaut considered potential mechanisms of microbial mat destruction in saline lakes and playas of the Cariboo Plateau, British Columbia, with an eye toward their potential for preservation within the geological record (Renaut, 1993). Of the seven mechanisms he proposed for destruction of Cariboo Plateau mats, diagenetic decomposition, which we deem equivalent to mat disassembly, seems to be the most likely explanation for the Hot Lake mat's overwinter disappearance from supralittoral and benthic surfaces.

Our observations suggest the hypothesis that the mat community assembles during periods in which solar energy is abundant and the rate of photosynthesis is correspondingly high. Rates of photosynthesis that exceed consumption may drive the accumulation of carbon-rich extracellular polymers that compose the mat's matrix and provide opportunities for the recruitment of new mat members with diverse metabolic capacities and narrow physicochemical tolerances. Conversely, when the rate of heterotrophic degradation of these polymers exceeds their rate of synthesis, the mat community may begin to disassemble as the matrix is consumed and niches are lost. Hot Lake, therefore, presents a unique opportunity to study the recruitment of metabolic function to an assembling community and the corresponding loss of function as the community disassembles (Johnson et al., 2012). Metagenome-enabled study of the Hot Lake mat community may uncover the interspecies metabolic interactions responsible for mat formation and stability and aid in the elucidation of design principles for microbial community assembly.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fmicb.2013.00323/abstract

REFERENCES

Abed, R. M., Kohls, K., and de Beer, D. (2007). Effect of salinity changes on the bacterial diversity, photosynthesis and oxygen consumption of cyanobacterial mats from an intertidal flat of the Arabian Gulf. *Env. Microbiol.* 9, 1384–1392. doi: 10.1111/j.1462-2920.2007.01254.x

Anderson, G. C. (1958). Some limnological features of a shallow saline meromictic lake. *Limnol. Oceanogr.* 3, 259–270. doi: 10.4319/lo.1958.3.3.0259

Anderson, G. C. (2012). Personal communication.

Andersson, A. F., Riemann, L., and Bertilsson, S. (2009). Pyrosequencing reveals contrasting seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *ISME J.* 4, 171–181. doi: 10.1038/ismej.2009.108

Bauld, J. (1981). 8. Occurrence of benthic microbial mats in saline lakes. *Hydrobiologia* 81–82, 87–111. doi: 10.1007/BF00048708

Benlloch, S., López-López, A., Casamayor, E. O., Øvreås, L., Goddard, V., Daae, F. L., et al. (2002). Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern. *Environ. Microbiol.* 4, 349–360. doi: 10.1046/j.1462-2920.2002.00306.x

Bennett, W. A. G. (1962). *Saline Lake Deposits in Washington.* Olympia, WA: State Printing Plant.

Beyenal, H., Lewandowski, Z., Yakymyshryn, C., Lemley, B., and Wehri, J. (2000). Fiber optic microsensors to measure back scattered light intensity in biofilms. *Appl. Opt.* 39, 3408–3412. doi: 10.1364/AO.39.003408

Braissant, O., Decho, A. W., Przekop, K. M., Gallagher, K. L., Glunk, C., Dupraz, C., et al. (2009). Characteristics and turnover of exopolymeric substances in a hypersaline microbial mat. *FEMS Microbiol. Ecol.* 67, 293–307. doi: 10.1111/j.1574-6941.2008.00614.x

Bray, J. R., and Curtis, J. T. (1957). An ordination of the upland forest communities of southern wisconsin. *Ecol. Monogr.* 27, 325. doi: 10.2307/1942268

Broch, E. S. (1969). The osmotic adaptation of the fairy shrimp Branchinecta campestris Lynch to saline astatic waters. *Limnol Oceanogr.* 14, 485–492. doi: 10.4319/lo.1969.14.4.0485

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4516–4522. doi: 10.1073/pnas.1000080107

Casamatta, D. A., Johansen, J. R., Vis, M. L., and Broadwater, S. T. (2005). Molecular and morphological characterization of ten polar and near-polar strains within the oscillatoriales (cyanobacteria)1. *J. Phycol.* 41, 421–438. doi: 10.1111/j.1529-8817.2005.04062.x

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* 31, 343–366. doi: 10.1146/annurev.ecolsys.31.1.343

Cohen, Y., Krumbein, W. E., Goldberg, M., and Shilo, M. (1977). Solar Lake (Sinai). 1. Physical and chemical limnology. *Limnol. Oceanogr.* 22, 597–608. doi: 10.4319/lo.1977.22.4.0597

Decho, A. W., Visscher, P. T., and Reid, R. P. (2005). Production and cycling of natural microbial exopolymers (EPS) within a marine stromatolite. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 219, 71–86. doi: 10.1016/j.palaeo.2004.10.015

Desnues, C., Michotey, V. D., Wieland, A., Zhizang, C., Fourçans, A., Duran, R., et al. (2007). Seasonal and diel distributions of denitrifying and bacterial communities in a hypersaline microbial mat (Camargue, France). *Water Res.* 41, 3407–3419. doi: 10.1016/j.watres.2007.04.018

Dillon, J. G., Miller, S., Bebout, B., Hullar, M., Pinel, N., and Stahl, D. A. (2009). Spatial and temporal variability in a stratified hypersaline microbial mat community. *FEMS Microbiol. Ecol.* 68, 46–58. doi: 10.1111/j.1574-6941.2009.00647.x

Dupraz, C., Reid, R. P., Braissant, O., Decho, A. W., Norman, R. S., and Visscher, P. T. (2009). Processes of carbonate precipitation in modern microbial mats. *Earth Sci. Rev.* 96, 141–162. doi: 10.1016/j.earscirev.2008.10.005

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Fernandez-Carazo, R., Hodgson, D. A., Convey, P., and Wilmotte, A. (2011). Low cyanobacterial diversity in biotopes of the Transantarctic Mountains and Shackleton Range (80–82°S), Antarctica. *FEMS Microbiol. Ecol.* 77, 503–517. doi: 10.1111/j.1574-6941.2011.01132.x

Ferrera, I., Massana, R., Balague, V., Pedros-Alio, C., Sanchez, O., and Mas, J. (2010). Evaluation of DNA extraction methods from complex phototrophic biofilms. *Biofouling* 26, 349–357. doi: 10.1080/08927011003605870

Fine, P. V. A., and Kembel, S. W. (2011). Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography* 34, 552–565. doi: 10.1111/j.1600-0587.2010.06548.x

Foesel, B. U., Gößner, A. S., Drake, H. L., and Schramm, A. (2007). Geminicoccus roseus gen. nov., sp. nov., an aerobic phototrophic Alphaproteobacterium isolated from a marine aquaculture biofilter. *Syst. Appl. Microbiol.* 30, 581–586. doi: 10.1016/j.syapm.2007.05.005

Foster, I. S., King, P. L., Hyde, B. C., and Southam, G. (2010). Characterization of halophiles in natural MgSO4 salts and laboratory enrichment samples: astrobiological implications for mars. *Planet. Space Sci.* 58, 599–615. doi: 10.1016/j.pss.2009.08.009

Franks, J., and Stolz, J. F. (2009). Flat laminated microbial mat communities. *Earth Sci. Rev.* 96, 163–172. doi: 10.1016/j.earscirev.2008.10.004

Frund, C., and Cohen, Y. (1992). Diurnal cycles of sulfate reduction under oxic conditions in cyanobacterial mats. *Appl. Environ. Microbiol.* 58, 70–77.

Gao, H., Chen, Z., Kumar, J., Tripathy, S., and Kaplan, D. (1995). Tapered fiber tips for fiber optic biosensors. *Opt. Eng.* 34, 3465–3470. doi: 10.1117/12.215379

Guerrero, M. C., and De Wit, R. (1992). Microbial mats in the inland saline lakes of Spain. *Limnetica* 8, 197–204.

Handy, F. M. (1916). *An Investigation of the Mineral Deposits of Northern Okanogan County*. The State College of Washington Bulletin Service. Olympia, WA.

Jenkins, O. P. (1918). Spotted lakes of epsomite in Washington and British Columbia. *Am. J. Sci.* 46, 638–644. doi: 10.2475/ajs.s4-46.275.638

Johnson, D. R., Goldschmidt, F., Lilja, E. E., and Ackermann, M. (2012). Metabolic specialization and the assembly of microbial communities. *ISME J.* 6, 1985–1991. doi: 10.1038/ismej.2012.46

Jonkers, H. M., Ludwig, R., Wit, R., Pringault, O., Muyzer, G., Niemann, H., et al. (2003). Structural and functional analysis of a microbial mat ecosystem from a unique permanent hypersaline inland lake: â▼ ↓ La Salada de Chipranaâ▼ ← (NE Spain). *FEMS Microbiol. Ecol.* 44, 175–189. doi: 10.1016/S0168-6496(02)00464-6

Jorgensen, B. B., Cohen, Y., and Revsbech, N. P. (1986). Transition from anoxygenic to oxygenic photosynthesis in a microcoleus chthonoplastes cyanobacterial mat. *Appl. Environ. Microbiol.* 51, 408–417.

Jorgensen, B. B., and Des Marais, D. J. (1988). Optical properties of benthic photosynthetic communities: fiber-optic studies of cyanobacterial mats. *Limnol. Oceanogr.* 33, 99–113. doi: 10.4319/lo.1988.33.1.0099

Jorgensen, B. B., Revsbech, N. P., Blackburn, T. H., and Cohen, Y. (1979). Diurnal cycle of oxygen and sulfide microgradients and microbial photosynthesis in a cyanobacterial mat sediment. *Appl. Environ. Microbiol.* 38, 46–58.

Jungblut, A.-D., Hawes, I., Mountfort, D., Hitzfeld, B., Dietrich, D. R., Burns, B. P., et al. (2005). Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environ. Microbiol.* 7, 519–529. doi: 10.1111/j.1462-2920.2005.00717.x

Jungblut, A. D., Lovejoy, C., and Vincent, W. F. (2009). Global distribution of cyanobacterial ecotypes in the cold biosphere. *ISME J.* 4, 191–202. doi: 10.1038/ismej.2009.113

Kilmer, B., Eberl, T., Cunderla, B., Chen, F., Clark, B., and Schneegurt, M. (in press). Molecular and phenetic characterization of the bacterial assemblage of Hot Lake, WA, an environment with high concentrations of magnesium sulfate, and its relevance to Mars. *Int. J. Astrobiol.* doi: 10.1017/S1473550413000268

Klatt, C. G., Inskeep, W. P., Herrgard, M. J., Jay, Z. J., Rusch, D. B., Tringe, S. G., et al. (2013). Community structure and function of high-temperature chlorophototrophic microbial mats inhabiting diverse geothermal environments. *Front. Microbiol.* 4:106. doi: 10.3389/fmicb.2013.00106

Kleinteich, J., Wood, S. A., Küpper, F. C., Camacho, A., Quesada, A., Frickey, T., et al. (2012). Temperature-related changes in polar cyanobacterial mat diversity and toxin production. *Nat. Clim. Change* 2, 356–360. doi: 10.1038/nclimate1418

Krumbein, W. E., Cohen, Y., and Shilo, M. (1977). Solar Lake (Sinai). 4. Stromatolitic cyanobacterial mats. *Limnol. Oceanogr.* 22, 635–656. doi: 10.4319/lo.1977.22.4.0635

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12

Lane, D. J. (1991). "16S/23S rRNA sequencing," in *Nucleic acid techniques in bacterial systematics*, eds. E. Stackebrandt and M. Goodfellow (Chichester: Wiley), 329.

Lee, D.-Y., and Rhee, G.-Y. (1999). Kinetics of growth and death in anabaena flos-aquae (cyanobacteria) under light limitation and supersaturation. *J. Phycol.* 35, 700–709. doi: 10.1046/j.1529-8817.1999.3540700.x

Lewandowski, L., and Beyenal, H. (2007). *Fundamentals of Biofilm Research*. CRC Press. ISBN-13: 978-0849335419

Losos, J. B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol. Lett.* 11, 995–1003. doi: 10.1111/j.1461-0248.2008.01229.x

Magoc, T., and Salzberg, S. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507

Martineau, E., Wood, S. A., Miller, M. R., Jungblut, A. D., Hawes, I., Webster-Brown, J., et al. (2013). Characterisation of Antarctic cyanobacteria and comparison with New Zealand strains. *Hydrobiologia* 711, 139–154. doi: 10.1007/s10750-013-1473-1

McKay, E. (1935). Salt tolerance of Ruppia maritima in lakes of high magnesium sulphate content. *Plant Physiol.* 10, 425–446. doi: 10.1104/pp.10.3.425

Nübel, U., Garcia-Pichel, F., Clavero, E., and Muyzer, G. (2000). Matching molecular diversity and ecophysiology of benthic cyanobacteria and diatoms in communities along a salinity gradient. *Environ. Microbiol.* 2, 217–226. doi: 10.1046/j.1462-2920.2000.00094.x

Oren, A. (2009). Saltern evaporation ponds as model systems for the study of primary production processes under hypersaline conditions. *Aquat. Microb. Ecol.* 56, 193–204. doi: 10.3354/ame01297

Oren, A. (2010). "Mats of filamentous and unicellular cyanobacteria in hypersaline environments," in *Microbial Mats: Modern and Ancient Microorganisms in Stratified Systems*, Cellular Origin, Life in Extreme Habitats and Astrobiology, Vol. 14, (Netherlands: Springer), 387–400. doi: 10.1007/978-90-481-3799-2_20

Oren, A. (2011). Thermodynamic limits to microbial life at high salt concentrations. *Env. Microbiol.* 13, 1908–1923. doi: 10.1111/j.1462-2920.2010.02365.x

Paerl, H. W., Pinckney, J. L., and Steppe, T. F. (2000). Cyanobacterial-bacterial mat consortia: examining the functional unit of microbial survival and growth in extreme environments. *Environ. Microbiol.* 2, 11–26. doi: 10.1046/j.1462-2920.2000.00071.x

Paerl, H. W., and Yannarell, A. C. (2010). "Environmental dynamics, community structure and function in a hypersaline microbial mat," in *Microbial Mats: Modern and Ancient Microorganisms in Stratified Systems*, Cellular Origin, Life in Extreme Habitats and Astrobiology, Vol. 14, (Netherlands: Springer), 421–442. doi: 10.1007/978-90-481-3799-2_22

Pierson, B., Oesterle, A., and Murphy, G. L. (1987). Pigments, light penetration, and photosynthetic activity in the multi-layered microbial mats of Great Sippewissett Salt Marsh, Massachusetts. *FEMS Microbiol. Lett.* 45, 365–376. doi: 10.1111/j.1574-6968.1987.tb02412.x

Pillet, L., de Vargas, C., and Pawlowski, J. (2011). Molecular identification of sequestered diatom chloroplasts and kleptoplastidy in foraminifera. *Protist* 162, 394–404. doi: 10.1016/j.protis.2010.10.001

Pinckney, J., Paerl, H. W., and Bebout, B. M. (1995). Salinity control of benthic microbial mat community production in a Bahamian hypersaline lagoon. *J. Exp. Mar. Biol. Ecol.* 187, 223–237. doi: 10.1016/0022-0981(94)00185-G

Power, I. M., Wilson, S. A., Thom, J. M., Dipple, G. M., Gabites, J. E., and Southam, G. (2009). The hydromagnesite playas of Atlin, British Columbia, Canada: a biogeochemical model for CO2 sequestration. *Chem. Geol.* 260, 286–300. doi: 10.1016/j.chemgeo.2009.01.012

Power, I. M., Wilson, S. A., Thom, J. M., Dipple, G. M., and Southam, G. (2007). Biologically induced mineralization of dypingite by cyanobacteria from an alkaline wetland near Atlin, British Columbia, Canada. *Geochem. Trans.* 8, 13. doi: 10.1186/1467-4866-8-13

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490. doi: 10.1371/journal.pone.0009490

Renaut, R. W. (1990). Recent carbonate sedimentation and brine evolution in the saline lake basins of the Cariboo Plateau, British Columbia, Canada. *Hydrobiologia* 197, 67–81. doi: 10.1007/BF00026939

Renaut, R. W. (1993). Morphology, distribution, and preservation potential of microbial mats in the hydromagnesite-magnesite playas of the Cariboo Plateau, British Columbia, Canada. *Hydrobiologia* 267, 75–98. doi: 10.1007/BF000 18792

Rothrock, M. J., and Garcia-Pichel, F. (2005). Microbial diversity of benthic mats along a tidal desiccation gradient. *Environ. Microbiol.* 7, 593–601. doi: 10.1111/j.1462-2920.2005.00728.x

Schloss, P. D., Gevers, D., and Westcott, S. L. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PloS ONE* 6:e27310. doi: 10.1371/journal.pone.0027310

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Env. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Severin, I., Confurius-Guns, V., and Stal, L. J. (2012). Effect of salinity on nitrogenase activity and composition of the active diazotrophic community in intertidal microbial mats. *Arch. Microbiol.* 194, 483–491. doi: 10.1007/s00203-011-0787-5

Sorensen, K. B., Canfield, D. E., Teske, A. P., and Oren, A. (2005). Community composition of a hypersaline endoevaporitic microbial mat. *Appl. Environ. Microbiol.* 71, 7352–7365. doi: 10.1128/AEM.71.11.7352-7365.2005

St. John, H., and Courtney, W. D. (1924). The flora of Epsom Lake. *Am. J. Bot.* 11, 100–107. doi: 10.2307/2435492

Stal, L. J., van Gemerden, H., and Krumbein, W. E. (1985). Structure and development of a benthic marine microbial mat. *FEMS Microbiol. Lett.* 31, 111–125. doi: 10.1111/j.1574-6968.1985.tb01138.x

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7, 2069–2079. doi: 10.1038/ismej. 2013.93

Stegen, J. C., Lin, X., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* 6, 1653–1664. doi: 10.1038/ismej.2012.22

Steunou, A.-S., Jensen, S. I., Brecht, E., Becraft, E. D., Bateson, M. M., Kilian, O., et al. (2008). Regulation of nif gene expression and the energetics of N2 fixation over the diel cycle in a hot spring microbial mat. *ISME J.* 2, 364–378. doi: 10.1038/ismej.2007.117

Swenson, N. G., Stegen, J. C., Davies, S. J., Erickson, D. L., Forero-Montana, J., Hurlbert, A. H., et al. (2012). Temporal turnover in the composition of tropical tree communities: functional determinism and phylogenetic stochasticity. *Ecology* 93, 490–499. doi: 10.1890/11-1180.1

Taffs, R., Aston, J. E., Brileya, K., Jay, Z., Klatt, C. G., McGlynn, S., et al. (2009). In silico approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC Syst. Biol.* 3:114. doi: 10.1186/1752-0509-3-114

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Van der Meer, M. T. J., Schouten, S., Bateson, M. M., Nübel, U., Wieland, A., Kühl, M., et al. (2005). Diel variations in carbon metabolism by green

nonsulfur-like bacteria in alkaline siliceous hot spring microbial mats from yellowstone national park. *Appl. Environ. Microbiol.* 71, 3978–3986. doi: 10.1128/AEM.71.7.3978-3986.2005

Villanueva, L., Navarrete, A., Urmeneta, J., White, D. C., and Guerrero, R. (2007). Analysis of diurnal and vertical microbial diversity of a hypersaline microbial mat. *Arch. Microbiol.* 188, 137–146. doi: 10.1007/s00203-007-0229-6

Walker, K. F. (1974). The stability of meromictic lakes in central Washington. *Limnol. Oceanogr.* 19, 209–222. doi: 10.4319/lo.1974.19.2.0209

Walters, W. A., Caporaso, J. G., Lauber, C. L., Berg-Lyons, D., Fierer, N., and Knight, R. (2011). PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics* 27, 1159–1161. doi: 10.1093/bioinformatics/btr087

Wang, J., Shen, J., Wu, Y., Tu, C., Soininen, J., Stegen, J. C., et al. (2013). Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J.* doi: 10.1038/ismej.2013.30

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM. 00062-07

Webb, C. O., Ackerly, D. D., and Kembel, S. (2011). *Phylocom: Software for the Analysis of Phylogenetic Community Structure and Character Evolution (with Phylomatic and Ecoevolve)*. User's manual, version 4.2. Available online at: http://www.phylodiversity.net/phylocom.

Yannarell, A. C., and Paerl, H. W. (2007). Effects of salinity and light on organic carbon and nitrogen uptake in a hypersaline microbial mat. *FEMS Microbiol. Ecol.* 62, 345–353. doi: 10.1111/j.1574-6941.2007.00384.x

Yannarell, A. C., Steppe, T. F., and Paerl, H. W. (2006). Genetic variance in the composition of two functional groups (diazotrophs and cyanobacteria) from a hypersaline microbial mat. *Appl. Environ. Microbiol.* 72, 1207–1217. doi: 10.1128/AEM.72.2.1207-1217.2006

Yannarell, A. C., Steppe, T. F., and Paerl, H. W. (2007). Disturbance and recovery of microbial community structure and function following Hurricane Frances. *Environ. Microbiol.* 9, 576–583. doi: 10.1111/j.1462-2920.2006.01173.x

Zlotnik, I., and Dubinsky, Z. (1989). The effect of light and temperature on DOC excretion by phytoplankton. *Limnol. Oceanogr.* 34, 831–839. doi: 10.4319/lo.1989.34.5.0831

Localized electron transfer rates and microelectrode-based enrichment of microbial communities within a phototrophic microbial mat

Jerome T. Babauta[1], **Erhan Atci**[1], **Phuc T. Ha**[1], **Stephen R. Lindemann**[2], **Timothy Ewing**[1], **Douglas R. Call**[3], **James K. Fredrickson**[2] and **Haluk Beyenal**[1]*

[1] The Gene and Linda Voiland School of Chemical Engineering and Bioengineering, Washington State University, Pullman, WA, USA
[2] Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA
[3] Paul G. Allen School for Global Animal Health, Washington State University College of Veterinary Medicine, Pullman, WA, USA

Phototrophic microbial mats frequently exhibit sharp, light-dependent redox gradients that regulate microbial respiration on specific electron acceptors as a function of depth. In this work, a benthic phototrophic microbial mat from Hot Lake, a hypersaline, epsomitic lake located near Oroville in north-central Washington, was used to develop a microscale electrochemical method to study local electron transfer processes within the mat. To characterize the physicochemical variables influencing electron transfer, we initially quantified redox potential, pH, and dissolved oxygen gradients by depth in the mat under photic and aphotic conditions. We further demonstrated that power output of a mat fuel cell was light-dependent. To study local electron transfer processes, we deployed a microscale electrode (microelectrode) with tip size ~20 μm. To enrich a subset of microorganisms capable of interacting with the microelectrode, we anodically polarized the microelectrode at depth in the mat. Subsequently, to characterize the microelectrode-associated community and compare it to the neighboring mat community, we performed amplicon sequencing of the V1–V3 region of the 16S gene. Differences in Bray-Curtis beta diversity, illustrated by large changes in relative abundance at the phylum level, suggested successful enrichment of specific mat community members on the microelectrode surface. The microelectrode-associated community exhibited substantially reduced alpha diversity and elevated relative abundances of *Prosthecochloris*, *Loktanella*, *Catellibacterium*, other unclassified members of *Rhodobacteraceae*, *Thiomicrospira*, and *Limnobacter*, compared with the community at an equivalent depth in the mat. Our results suggest that local electron transfer to an anodically polarized microelectrode selected for a specific microbial population, with substantially more abundance and diversity of sulfur-oxidizing phylotypes compared with the neighboring mat community.

Keywords: electron transfer, hot lake, microbial mats, microelectrodes, sulfur cycle, sequence analysis

INTRODUCTION

Microbial mats are highly stratified microbial communities where metabolically interacting species inhabiting strata generated by sharp physicochemical gradients transfer energy via metabolic byproducts (Franks and Stolz, 2009). The proximity and density of organisms inside microbial mats makes energy transfer highly efficient, where energy transfer refers to oxidation–reduction reactions that predominantly cycle carbon and sulfur throughout the mat. These properties make microbial mats exceptional candidates for specific applications, including carbon sequestration and wastewater treatment (Roeselers et al., 2008). Using microbial mats in this manner requires precise control of the physicochemical gradients that drive energy transfer where one possibility to control such gradients is to electrochemically change the mat environment by providing (reduction) or removing (oxidation) electrons. Theoretically, this could be accomplished inside the microbial mat by using electrodes. Because organisms in the

mat proliferate within spatially defined niches along physicochemical gradients in the mat, drawing a current is likely to cause changes in niche localization and hence alter the composition of the local microbial community. Therefore, quantifying the effect of electron transfer within microbial mats could also be used to understand the energy balance that leads to stratification of the species and associated functions within the mat.

Microbial mats are similar to microbial biofilms, except that microbial mats exhibit lamination (Guerrero et al., 2002). Biofilms tend to be composed of a single species and form readily on electrode surfaces, making them popular model systems in which to study electron transfer mechanisms. In contrast, a high diversity of species and, therefore, microbial metabolisms (e.g., photosynthesis, sulfate reduction) simultaneously operate at various locations within a mat. Consequently, efforts to detect electron transfer in the mat need to be spatially precise and take into account dynamic

physicochemical gradients. Understanding local electron transfer within a mat is made increasingly difficult because the diversity and dynamicity of metabolism generates a complex web of metabolites that shifts over time. Therefore, the common method of employing macro-scale, passive microbial fuel cells to detect power generation by either anodic or cathodic microorganisms that respond to the microbial fuel cell is not capable of directly measuring local electron transfer rates *inside* a mat. However, several studies have demonstrated microbial fuel cells that use a photosynthetic community colonized on the electrodes to produce power (He et al., 2009; Nishio et al., 2010; Bradley et al., 2012; Chandra et al., 2012; Badalamenti et al., 2013; Lan et al., 2013; Lin et al., 2013; Strycharz-Glaven et al., 2013). However, power generation is not our focus here. These techniques are limited because all use large electrodes and therefore do not possess the resolution to measure local electron transfer processes in stratified systems at the microscale. Our goal was to determine if local electron transfer inside a mat could be measured and, if so, manipulated to induce changes in the local community.

Local electron transfer can be resolved by scaling down to needle-type microscale electrodes, or simply microelectrodes (Lewandowski and Beyenal, 2014). Initially, when a microelectrode with a carbon tip is inserted into a desired location in a microbial mat, the microelectrode tip responds to redox couples and reaches an open circuit potential. Once the microelectrode tip is polarized against a reference electrode, an electrochemical gradient is generated where electroactive compounds can be oxidized or reduced. Subsequently, some bacteria will respond to an increase/decrease in electroactive compounds while other species should be able to take advantage of the electrochemical gradient directly by transferring electrons to the electrode, as happens with known dissimilatory metal-reducing bacteria (Shi et al., 2009; Borole et al., 2011; Babauta et al., 2012a). However, we note that known dissimilatory metal-reducing phylotypes (e.g., *Geobacter*, *Shewanella*) were not detected within the Hot Lake mat community over one annual cycle (Lindemann et al., 2013). Regardless of the specific mechanism, the electrochemical gradient should only cause a change if current is generated; making local current generation a useful indicator to determine if a community change has occurred. In the absence of current generation, polarization of the microelectrode tip could also cause physical association of microorganisms considering van der Waals and electrostatic interactions (Hori and Matsumoto, 2010). However, the changes in electrostatic interactions caused by charging/discharging of the electric double layer at the electrode surface is generally neutralized by ions in solution at the overpotentials used here and given the high ionic strength of Hot Lake water.

Although there is a long history of microscale measurements made in biofilms and microbial mats, these studies have not been performed with polarized microelectrodes (Jørgensen and Revsbech, 1983; Revsbech and Ward, 1984; de Beer et al., 1994; Kühl et al., 1994; Beyenal et al., 1998; Bishop and Yu, 1999; Nguyen et al., 2012). Earlier studies focused on the penetration of oxygen in biofilms and microbial mats, which allowed microaerophilic and anaerobic modes of life to exist even in oxic waters. In addition to biological consumption of oxygen,

we have characterized the electrochemical consumption of oxygen and other relevant species within biofilm on electrodes. In previous reports, we demonstrated that surface conditions significantly differ from those only hundreds of microns distal from the electrode surface. These included several 100 mV changes in redox potential, several units of pH changes, complete consumption of oxygen, and formation of hydrogen peroxide (Babauta et al., 2012b, 2013). Termed "microscale gradients" due to the spatial dimensions of such physicochemical gradients, microscale gradients related to biofilms on electrodes have only been studied recently. The relationship between these microscale gradients and electrode processes play an important role in understanding how energy transfer is possible at such long distances (Erable et al., 2012; Malvankar and Lovley, 2012; Snider et al., 2012; Renslow et al., 2013; Vargas et al., 2013). Here, we extend previous considerations of microscale gradients above electrodes to the use of polarized microelectrodes where similar current densities are expected to promote similar redox gradients.

We chose mats derived from Hot Lake as a model microbial mat system. Hot Lake is an epsomitic lake located near Oroville in north-central Washington that contains a benthic, phototrophic mat that exhibits seasonal variations in salinity, temperature, and light (Anderson, 1958; Lindemann et al., 2013). As the Hot Lake mat is exposed to the greatest sulfate concentrations of any described phototrophic mat and its potential for elevated electron flow mediated by reduced sulfur compounds was correspondingly high, this mat was an ideal system in which to study local electron transfer rates. Furthermore, previous reports have shown that Hot Lake mats harbor substantial populations of phylotypes consistent with sulfur cycling, indicative of high energy transfer. In this work, we quantify redox potential, pH and dissolved oxygen profiles in the field during day and night to demonstrate microscale gradients and demonstrate electricity generation using a mat fuel cell. We further quantify local electron transfer rates with an anodically polarized carbon microelectrode tip and analyze the microelectrode-associated community generated in comparison to the neighboring mat community. Finally, we propose a local electron transfer mechanism operating at the microelectrode tip to explain the differences between microelectrode-associated and neighboring mat communities.

MATERIALS AND METHODS

COLLECTING MICROBIAL MAT SAMPLES

Microbial mat samples for both field and laboratory experiments were collected from Hot Lake, Oroville, WA, USA (48.973062°N, 119.476876°W at an elevation of ~576 m) in July 2012 and May 2013, respectively. Mat was gently removed with 1–2″ of sediment at depth and was transferred into containers holding Hot Lake water.

FIELD MICROBIAL MATS AND DEPTH PROFILES OF REDOX POTENTIAL, pH, AND OXYGEN

For field experiments, microbial mat samples were placed in a 2-L open channel reactor and Hot Lake water pumped from the lake was circulated (~5 h^{-1}) continuously during field measurements.
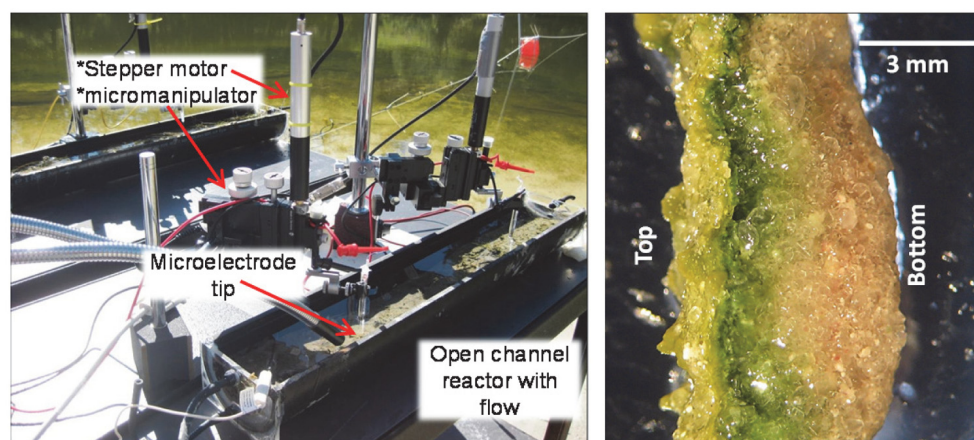
FIGURE 1 | Photo of open channel reactors containing microbial mat samples used in microprofiling (left). Stereomicroscope image of excised mat slice on its side showing stratification (right).

Redox potential, pH, and dissolved oxygen microelectrode measurements were carried out as shown in **Figure 1**. Microelectrode movements were controlled by a Mercury Step motor controller PI M-230.10S Part No. M23010SX (Physik Instrumente, Auburn, MA, USA). Each microelectrode was positioned ∼2000 μm above the mat surface and stepped down in 5 μm increments using custom microprofiler software. A Zeiss Stemi 2000 stereomicroscope was used to determine the locations of the microelectrode tip and surface of the mat. Redox potential, pH, and dissolved oxygen microelectrodes were constructed according to Lewandowski and Beyenal (2014).

LABORATORY MICROBIAL MATS

For lab-scale experiments, microbial mat samples were transferred to our laboratory at Washington State University, Pullman, WA, USA. All samples were placed in reactors mimicking lake conditions without flow. First, sediment having approximately 5 cm thickness was added to the bottom of the reactor. Then, mat samples of 3–5 mm thickness and morphologically representative of neighboring mat were added to the top of the sediment. These mats were morphologically consistent with the mat described by Lindemann et al. (2013). During lab-scale experiments, deionized water was added periodically to maintain water level and salinity in the reactor. These mat samples were incubated in an enclosed growth chamber at ambient temperature that was artificially lighted using a Reeflux 250W 12000 K double ended metal halide lamp. The light was cycled in a 14 h on/10 h off cycle.

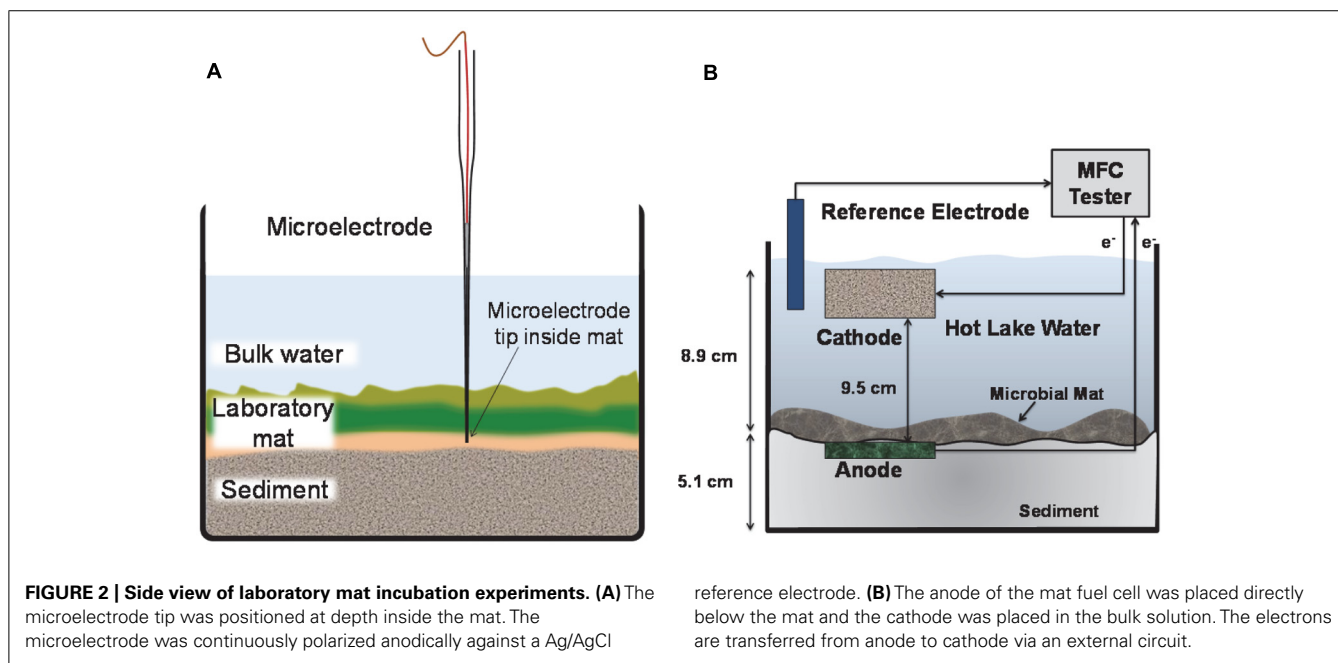### Electrochemical measurements inside laboratory microbial mats

Electrochemical measurements refer to the generation of current from either anodic or cathodic reactions using anodic microelectrode polarization or a mat fuel cell. **Figure 2** shows the set up used for a microelectrode deployed inside the microbial mat (**Figure 2A**) and a mat fuel cell (**Figure 2B**). The microelectrode tip was positioned inside the laboratory mat at depths where physicochemical gradients (i.e., redox potential, pH,

oxygen, S^{2-}) were steep. Typically, this occurred at depths of 2–3 mm from the mat surface. At depths >3–4 mm, gradients were typically less steep indicating proximity to the sediment layers underneath the mat, which is similar in trend to what is observed at the bulk/mat surface interface. The steepest gradients were always observed to span the mat thickness and we targeted all microelectrode measurements at these depths. For the mat fuel cell, however, the anode could only be placed directly under the mat and was expected to respond to the gradients that extended beyond the mat thickness or that were external to the mat.

Mat fuel cell as current collector from the microbial mat

We used graphite felt for both the anode and the cathode (HP Materials Solutions, Inc., Woodland Hills, CA, USA) and each had projected surface areas of 2.14×10^{-2} and 6.48×10^{-3} m^2, respectively. The electrodes were connected to a MFC Tester (Dewan et al., 2010) using a Grade 2, 0.635 mm diameter Ultra-Corrosion Resistant Titanium Wire (McMaster-Carr, Los Angeles, CA, USA) and 20 gauge copper insulated hookup wire. Ti wires were woven into the graphite felt and secured with nylon bolts. A mechanical and solder connection to the copper wire was sealed with silicone rubber to prevent water intrusion. The resistance of the copper wire, Ti wire and graphite felt connections were <1 Ohm at every point measured on each electrode.

The external circuit used to calculate power was a MFC Tester and was identical to the one described previously (Dewan et al., 2010). Briefly, the main components included a capacitor, two n-channel MOSFET switches to control energy storage and use, a USB-1608FS data acquisition module (Measurement Computing Corporation, Norton, MA, USA) and a PC with custom LabVIEW VI (National Instruments Corporation, Austin, TX, USA) to monitor and control the charge/discharge cycle. The generated energy was dissipated by shorting the capacitor through the ground of the data acquisition module thereby preventing electrons from being returned to the mat fuel cell.

FIGURE 2 | Side view of laboratory mat incubation experiments. (A) The microelectrode tip was positioned at depth inside the mat. The microelectrode was continuously polarized anodically against a Ag/AgCl reference electrode. (B) The anode of the mat fuel cell was placed directly below the mat and the cathode was placed in the bulk solution. The electrons are transferred from anode to cathode via an external circuit.

Power generation was monitored continuously by measuring the charge/discharge rates of the MFC Tester capacitor. We used Vc 50 mV and Vd 350 mV as charge and discharge potentials of the 1 F capacitor.

Microelectrodes as current collector from the microbial mat

Needle-type microelectrodes with 15–20 μm tip diameters were used to collect current inside the mat at depth. As described earlier for the field experiments, microelectrode tips were moved precisely inside laboratory mats using a micromanipulator with automated stepper motor and stereomicroscope. Once in position, a Gamry Interface 1000TM potentiostat (Gamry® Instruments, Warminster, PA, USA) was used to control the microelectrode tip potential. A saturated Ag/AgCl reference electrode and platinum auxillary electrode completed the three-electrode setup (not shown in **Figure 2A**). Polarizing at +400 mV$_{Ag/AgCl}$, electrons were passed from the microelectrode tip to the auxiliary electrode and the current was measured. Because the microelectrode tip was polarized above the open circuit potential to +400 mV$_{Ag/AgCl}$, it acted as an anode inside the mat and was therefore unable to reduce oxygen. Thus the generally negative effects associated with oxygen reduction (i.e., generation of hydrogen peroxide and/or oxygen radicals) were avoided.

Procedures for the microelectrode used here are described in detail elsewhere (Lewandowski and Beyenal, 2014). Carbon fiber wires having 30 μm diameter (World Precision Instrument©) were used to construct microelectrodes. Corning 8161 glass was used to make a shaft for the carbon fiber. Carbon fibers were sealed in the glass by heat pulling. The carbon tip of the microelectrode was exposed by grinding away the glass seal using a diamond grinding wheel (Narishige, Model #EG-4). The diameter of the tip decreased after pulling to 15–20 μm due to the applied heat.

COMMUNITY ANALYSIS
DNA extraction

Genomic DNA from accumulated biomass on the microelectrode tip and a 1.5 mm-thick region of neighboring mat biomass from near the tip were extracted for community analysis. Before extraction, samples were washed twice with filter-sterilized (0.2 μm) Hot Lake water collected from the aquarium followed by two washes with TE buffer (10 mM Tris pH 8.0, 1 mM EDTA pH 8.0). This assured that only firmly adherent cells remained at the tip of the microelectrode. DNA was extracted using QIAamp Kit (QIAGEN) by following manufacture's instruction. Extracted DNA products were quantified prior to sequencing using a NanoDrop.

PacBio sequencing

Fragments of 16S rRNA genes containing variable V1–V3 regions were amplified from the extracted DNA with primers 27F (GAGTTTGATCMTGGCTCAG) and 515R (TTACCGCGGCT-GCTGGCAC) (Kroes et al., 1999). Three barcodes, FB2 (TCAT-GAGTCGACACTA), FB8 (CTGCTAGAGTCTACAG), and RB2 (GCGATCTATGCACACG) were added to primers in paired asymmetric mode (FB2-RB2 for microelectrode sample and FB8-RB2 for microbial mat sample) for further sorting of each sample from pooled PacBio sequencing outcomes. Each PCR reaction was performed in duplicate 25 μl reactions containing 30 ng of DNA, 1X GoTaq@Flexi Buffer, 1.25 mM of MgCl$_2$, 0.2 mM each dNTP, 0.1 μM of each barcoded primer (IDT), and 1.25 U of *Taq* polymerase (Promega). A C1000 Touch^{TM} Thermal Cycler (BIO-RAD, CA, USA) was used for the PCR as following program: (i) an initial denaturation step at 95°C for 4 min, (ii) 25 amplification cycles (95°C for 30 min, 57°C for 10 s, and 72°C for 20 s), and (iii) final extension at 72°C for 5 min. After this PCR amplification, the amplicons were purified (QIAGEN PCR purification kit) and quantified (NanoDrop). Amplicons were pooled and sequenced

using a PacBio-RSII sequencer (Washington State University, Pullman, WA, USA). PacBio FASTAQ formatted circular consensus sequences (CCS) were processed and analyzed using Mothur v.1.32 (Schloss et al., 2009).

Sequence processing and analysis

Sequences were quality trimmed using a sliding window of 10 bp and an average quality score of 40 and sequences with one or more ambiguous bases were removed. Filtered sequences were dereplicated and aligned to SILVA-based bacterial reference alignment to which the aligned Hot Lake mat near full length sequences had been added as previously described (Lindemann et al., 2013). Sequences were then screened to remove those that did not align to positions 1044–10,241 of the reference alignment, filtered to remove non-informative columns, pre-clustered to >99% identity (allowing four differences), and dereplicated. Chimeras were identified and removed using UCHIME as implemented in mothur 1.32 in self-referential mode. Filtered sequences were then classified using a Wang approach against the RDP training set v.9 reference with 80% bootstrap cutoff and sequences of unknown classification were removed. Sequences were subsampled (to the size of the smallest group $n = 10,423$ sequences) and clustered into operational taxonomic units (OTUs) at 0.03 average distances using the average neighbor algorithm in mothur. OTUs were classified based upon the sequence classifications described above. Alpha and beta diversity metrics were calculated using subsampled sequences described above.

RESULTS AND DISCUSSION

Traditionally, microbial mat communities have been studied using chemical sensors to determine fluxes of carbon, nitrogen, sulfur, and oxygen. From a practical engineering perspective, the flux of these elements creates an elemental balance that could be related to an equivalent flux of electrons if the governing redox reactions are known. Energy transformations can be studied directly and *in situ*. In order to establish the initial conditions for detection of electron transfer occurring at depth in the Hot Lake mat samples, we measured depth profiles of redox potential, pH, and oxygen around the diel cycle. These profiles were taken in explanted mat in open-flow channels continuously drawing lake water from the same depth in the lake as the collected mat. In addition, we measured depth profiles with similar trends in the laboratory (results not shown).

REDOX POTENTIAL, pH, AND DISSOLVED OXYGEN DEPTH PROFILES

Figures 3A,B show redox potential and pH depth profiles entering from the top of the mats at two time points during a 24-h day/night cycle and represent the cyclical response of the Hot Lake mat to light availability. During the day, redox potential remained positive at all depths in the mat from a bulk value of ∼475 mV$_{Ag/AgCl}$ to an observed minimum of ∼320 mV$_{Ag/AgCl}$. Redox potential reached a minimum at a depth of 1000 μm in the mat, and began to increase with further depth. During the night, the bulk value was slightly decreased to ∼+400 mV$_{Ag/AgCl}$ but sharply decreased to an observed minimum of approximately −150 mV$_{Ag/AgCl}$. Unlike the profile observed during the day, the night redox gradient showed a continual decrease in redox potential, reaching a minimum at 1700 μm under the surface of the mat. **Figure 3B** shows a similar change in the pH trend in the top layers of the mat between gradients taken during the day and night. With a stable bulk pH of ∼9.08 for both day and at night, we observed a maximum intra-mat pH of 9.8 during the day and a decrease to a pH of 8.98 during the night.

Figure 4 shows the change in oxygen concentration in the top layers of the mat taken during the day and night and confirms the trends observed in redox potential and pH changes in the mat. Consistent with previous reports on cyanobacterial mat communities, oxygen concentration changed dramatically over the diel cycle. During the day, oxygen reached supersaturation at ∼100 μM (headspace ∼21% O$_2$) and during the night was essentially undetectable.

During the day, oxygen saturated the top layers of the mat and likely controlled redox potential (cf. **Figures 3A** and **4**). However, the dip in redox potential in the mat, although still in the oxic range (Timothy et al., 2011), is not easily explained. Since steep physicochemical gradients span the mat, it is possible that unknown redox-active compounds slightly affect the primarily oxygen-dominated redox potential. At night, in the absence of light energy, reduced carbon is likely turned over until oxidant is depleted. Fermentation products then control redox potential at depth where oxidant (especially, oxygen) is
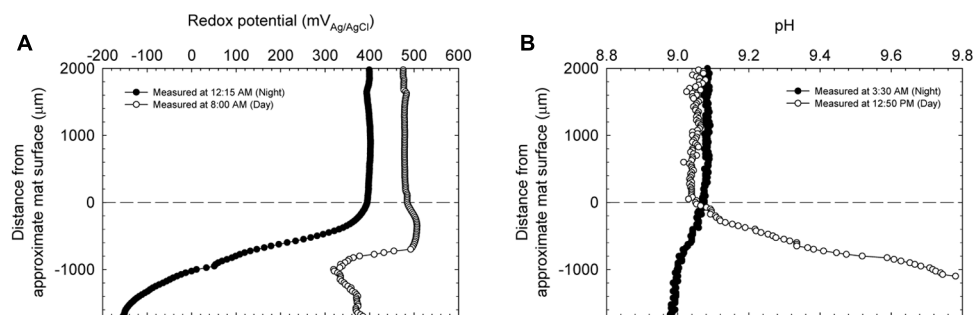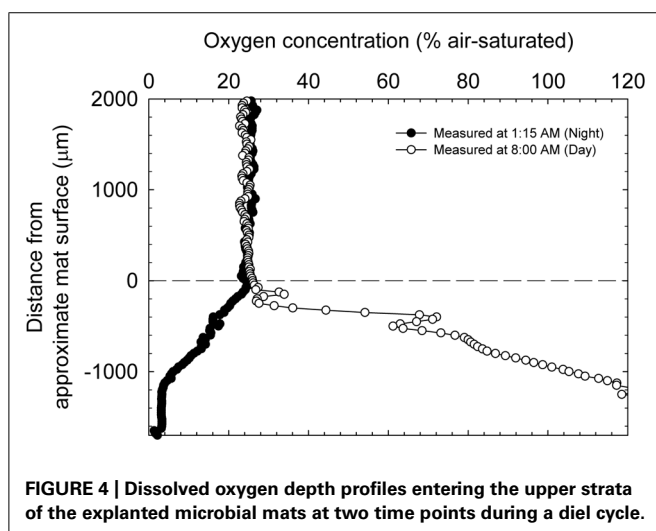

FIGURE 3 | Redox potential (A) and pH (B) depth profiles entering the upper strata of the explanted microbial mats at two time points during a diel cycle. Zero indicates the approximate location of the mat's surface.

FIGURE 4 | Dissolved oxygen depth profiles entering the upper strata of the explanted microbial mats at two time points during a diel cycle.

depleted. Because fermentation occurs in highly reducing environments, redox potential at night is expected to decrease with depth in the mat. Consistent with this interpretation, oxygen concentration inside the mat at night decreased with depth, eventually to below the detection limit (**Figure 4**). The changes in pH likely follow changes in alkalinity with carbon dioxide equilibrium, including carbonate precipitation, during the diel cycle (Lindemann et al., 2013). pH closely mimics the dissolved oxygen curve as a spatial indication of photosynthesis in the upper strata of the mat. During the night, CO_2 is regenerated in the mat by respiration with oxygen (at the top), sulfate, and also by fermentation. Following the gradient of fermentation and sulfate reduction, pH decreases from top to bottom in the mat.

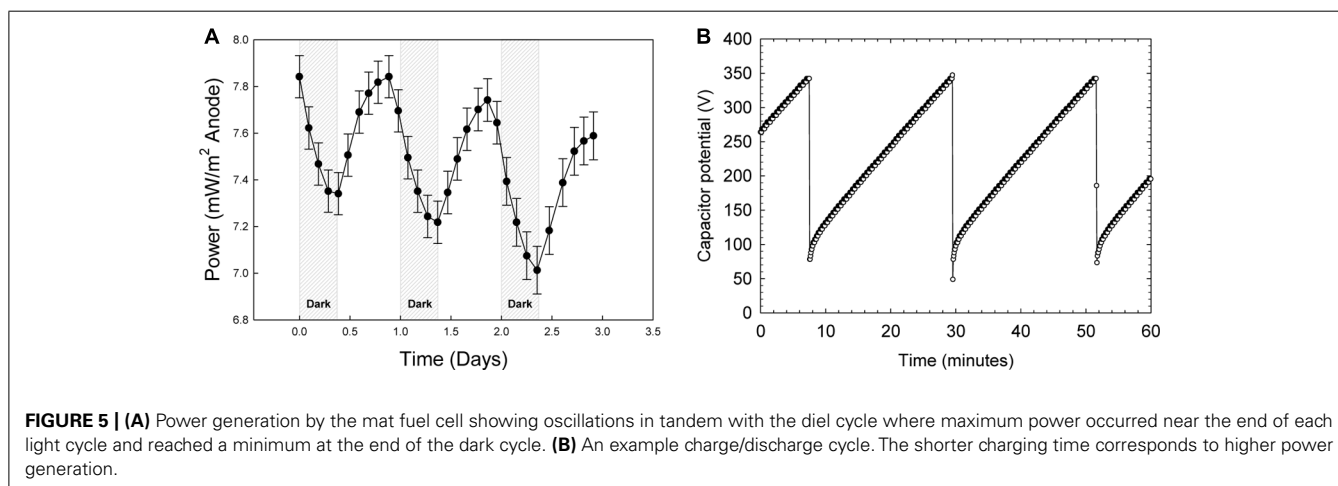## PHYSICOCHEMICAL GRADIENTS AND ELECTRON TRANSFER

Translating redox potential gradients into electron transfer rates could be achieved through the use of a mat fuel cell. The mat fuel cell harnesses electrons from the mat system externally using macro-scale electrodes and the power generation with the diel

cycle is monitored. This is similar to sediment microbial fuel cells as have been extensively described (Reimers et al., 2006); however, instead of burying the anode in the sediment, we placed it directly below the mat. We should note that inserting electrodes into the mat ecosystem is difficult and induces modest disruption of existing mats. In the following sections, we demonstrate utilizable electron transfer inside the mat and that selectively drawing off electrons intra-mat may cause changes in the local mat community.

Mat fuel cell power generation

Figure 5 shows power generation by a mat fuel cell. Power oscillated in tandem with the diel cycle where maximum power occurred near the end of each light cycle and reached a minimum at the end of the dark cycle. Each data point in **Figure 5A** corresponds to a charge/discharge cycle. The potential variation for an example charge discharge cycle is shown in **Figure 5B**. Because power was calculated by measuring the rate of charge/discharge of the MFC Tester, the cyclical nature of the measured power from the mat fuel cell directly translates to increases and decreases in the rate of charge/discharge. As the cathode was placed in the bulk and bulk oxygen concentration was nearly identical over the diel cycle (**Figure 4**), the majority of the oscillation was primarily localized to the anode directly below the mat. This suggests that the shift in gradients internal to the mat as discussed earlier and the gradients that extended outside the mat caused the power to oscillate.

Interestingly, observing increasing power in the presence of light is counter-intuitive given the trend in oxygen availability within the mat. As shown in **Figure 4**, oxygen concentration in the mat during the day reaches supersaturation and therefore should cause the anode directly under the mat to "discharge" as electrons are siphoned off to oxygen reduction, especially because the anode open circuit potential is only sustainable in the *absence* of oxygen. Admittedly, the contradiction is perplexing; however, we have shown previously that for mixed species environmental biofilms, oxygen gradients inside the biofilm can be quite sharp (Babauta et al., 2013). In that work, oxygen concentration depth profiles (using the same technique described here) showed that

FIGURE 5 | (A) Power generation by the mat fuel cell showing oscillations in tandem with the diel cycle where maximum power occurred near the end of each light cycle and reached a minimum at the end of the dark cycle. **(B)** An example charge/discharge cycle. The shorter charging time corresponds to higher power generation.

oxygen concentration could increase to nearly double the bulk concentration approximately 300 μm below the biofilm surface and subsequently decrease to undetectable levels within 2 mm. Considering that concentrations of oxygen and other metabolically relevant compounds reach maximal concentrations inside the mat and not necessarily at the base of the mat provides a valid reason as to why intra-mat processes cannot be fully explained by power cycling of the mat fuel cell. It is possible that the mat is not uniformly oxic and instead retains anoxic regions even in the presence of light; this is suggested by both the attenuation of wavelengths used for oxygenic photosynthesis and a dramatic increase in phylotypes known to employ microaerobic and anaerobic metabolisms below a depth of 2.5-3 mm in the Hot Lake mat (see **Figures 4** and **7**, Lindemann et al., 2013). Similar behavior was observed in several reports on phototrophic microbial fuel cells where current output in the first month primarily increased during light exposure but after 5 months decreased during light exposure (He et al., 2009). The authors attributed the change in trends to the dynamic interactions between photosynthetic microorganisms and heterotrophic bacteria. With the same reasoning, the

FIGURE 6 | Increase in current at a microelectrode tip placed 3 mm inside mat and polarized at +400 mV$_{Ag/AgCl}$.

complex interactions and energy transfer between dominant communities at different depths in the mat may cause the transient response of the mat fuel cell. Exactly which dominant community the anode is responding to is not determinable from **Figure 5** and only broad interpretations can be given. It is also important to note that Lindemann et al. (2013) documents seasonal changes in the mat community that were driven predominantly by changes in light availability. Such changes were hypothesized to rely upon the balance between photosynthetic production and heterotrophic consumption, especially in the diverse bottom layers of the mat.

New method to scale down to micro-scale electron transfer analysis

To alleviate issues with scale and study local electron transfer, we scaled down from large electrodes to very precise microelectrodes. As shown in **Figure 2A**, microelectrodes were positioned such that the microelectrode tip (active surface) was placed 3 mm deep in the mat each time and were polarized to +400 mV$_{Ag/AgCl}$. **Figure 6** shows that current generally increased during the day and decreased overnight. However, unlike in **Figure 5**, the maximum current was observed halfway through the light exposure period whereas for the mat fuel cell it was towards the end of the light exposure period. One possible explanation for the shift in current maximum was the proximity of the microelectrode tip to the oxygenic phototrophs at the top of the mat. Initially, inside the mat, reduced carbon compounds accumulated in the early stages of light exposure are oxidized vigorously by heterotrophs, which cause the increase in current (van der Meer et al., 2005; Behrens et al., 2008). Midway through the light cycle, oxygen concentration driven by oxygenic phototrophs, which should increase in magnitude and breadth during light cycles, began to influence the region surrounding the microelectrode tip and subsequently decrease the current. The results from this technique suggest that the observation of the maximal current within the light exposure period may depend greatly upon the depth at which the microelectrode tip is placed, which aligns well with the conclusions by He et al. (2009).

Surprisingly, these current oscillations were minor in comparison to the increase in baseline current. From a baseline of 0.2 nA, current increased over ~5 days to a maximum current

FIGURE 7 | Distribution of phyla for the microbial mat sample (A) and microelectrode tip sample (B).
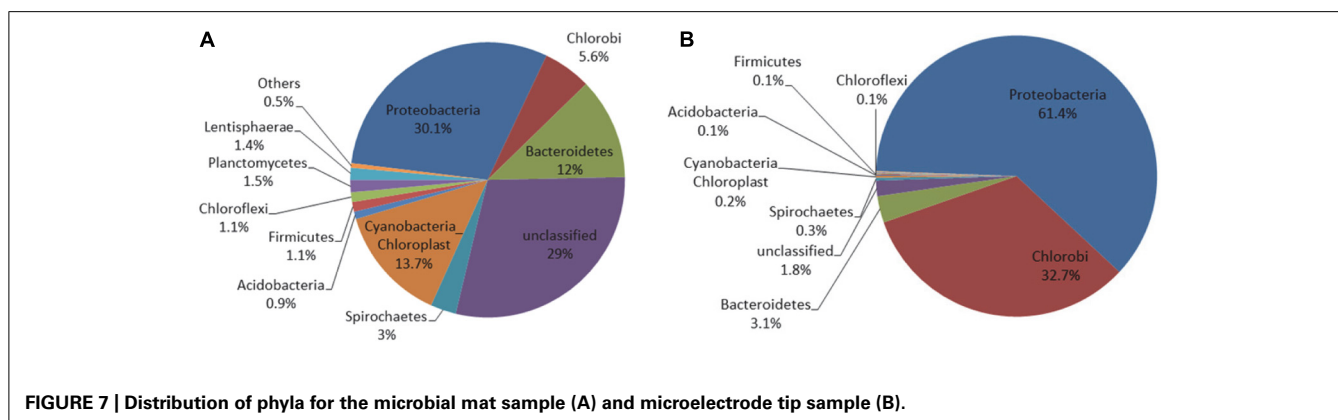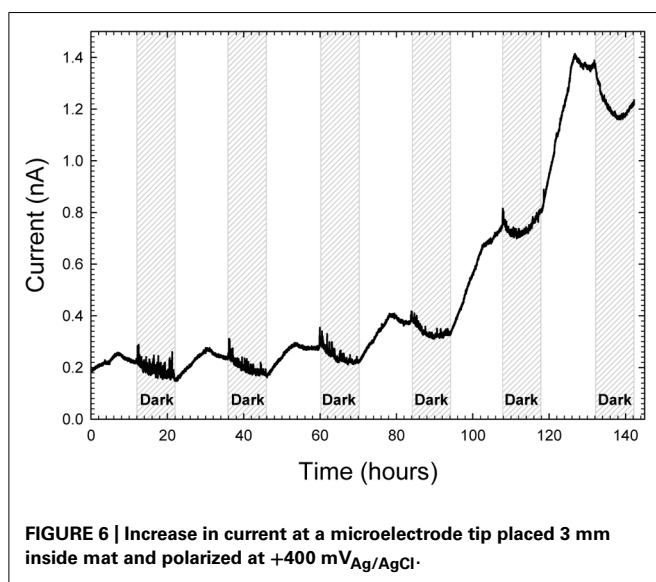
of 1.4 nA. We note that with a microelectrode tip diameter of 20 μm, assuming hemispherical shape, the surface area was 6.28×10^{-6} cm². Therefore, current densities were on the order of ~200 μA/cm². The observed range of current density at the microelectrode tip was of the same order of magnitude as macroscale investigations where we measured electrochemical gradients above polarized electrodes (Babauta et al., 2012b, 2013). Therefore, large and experimentally meaningful electrochemical gradients were imposed upon tightly localized regions in the mat. Broadly speaking, the baseline current increase reveals that more than one mechanism of energy transfer could be occurring at varying depths in the mat. Since the oscillations with the diel cycle are visible as a superposition on top of the increasing baseline, the increasing baseline itself is likely a result of a local change in the physicochemical environment in response to the electrochemical gradient formed around the microelectrode tip. Due to the complexity of the mat community, any number of explanations could explain the current increase. One possible explanation could be that the current increase is caused by a change in the local mat community.

MICROBIAL COMMUNITY ANALYSIS

To determine whether the microelectrode-associated community was substantially different than the neighboring community, we extracted gDNA from the microelectrode tip and the surrounding mat after 27-day in-mat incubation, performed 16S rRNA sequencing and compared the two communities. We recovered about 0.33 μg of genomic DNA from the microelectrode tip. PacBio sequencing of the V1–V3 region of the 16S gene yielded a total of 66,811 CCS from both samples. Approximately 85% of the raw reads passed the PacBio quality filtering standards with an average length of 400 bp. Post-sequencing quality filtration and subsampling yielded 20,846 sequences (10,423 from each sample).

Figure 7A shows that the microbial mat sample was dominated by members of the phyla *Proteobacteria* (30.1%), *Cyanobacteria-Chloroplast* (13.7%), *Bacteroidetes* (12%), *Chlorobi* (5.6%) and a large fraction of reads that could not be assigned to a phylum (29%). Except for *Chlorobi*, these phyla were also dominant in the Hot Lake mat throughout the seasonal cycle of 2011 (Lindemann et al., 2013). On the microelectrode tip, shown in **Figure 7B**, >95% of the reads were placed within only two phyla, *Proteobacteria* (61.4%) and *Chlorobi* (32.7%). In contrast

with the neighboring mat, reads attributed to *Cyanobacteria-Chloroplast* and *Bacteroidetes* represented only about 0.2 and 3.1% of the total microelectrode-associated population, respectively. The microelectrode-associated community exhibited reduced alpha diversity (inverse Simpson metric = 8.462) compared with the neighboring mat (inverse Simpson metric = 48.351), and both species richness and evenness were lower in the microelectrode-associated community than the neighboring mat. The distance between the two samples in Bray-Curtis beta diversity, which compares the relative abundances of species observed between communities, was 0.892. The divergence exhibited between the two communities suggests that a specific subpopulation of mat organisms was enriched on the microelectrode tip.

Sequences were clustered into OTUs and the dominant bacterial OTUs from the microelectrode-associated community are listed in **Table 1**. With one exception (OTU ME1), the other dominant bacterial OTUs in the table displayed low abundance in the neighboring mat community. OTU ME1, whose nearest cultured neighbor is *Prosthecochloris aestuarii* DSM 271 (97% identity across the sequenced region, GenBank accession: NR_074364.1), accounted for 4.6% of reads in the neighboring mat community, but more than 30% of the total reads from the microelectrode-associated community. OTU ME2 is very closely related (99%) to *Loktanella vestfoldensis* NBRC 102487 (GenBank accession: AB681826.1), known to be a strictly aerobic bacterium. Similarly, OTU ME4 and OTU ME6 were also identified as aerobic species which belong to genera *Catellibacterium* and *Thiomicrospira*, respectively. The remaining dominant OTUs in *Rhodobacteraceae* could not be classified below the family level.

The members of genus *Prosthecochloris* are known to be obligately anaerobic phototrophic green sulfur bacteria and can utilize sulfide and elemental sulfur as electron donors for photosynthesis (Gorlenko, 2001; Imhoff, 2003). The final oxidation product is soluble sulfate (Imhoff, 2003). The high relative abundance of OTU ME1 in the neighboring mat community and its prevalence on the microelectrode-associated community may be due to the abundance of sulfide in the Hot Lake microbial mat. Several previous studies have also reported that anaerobic phototrophic bacteria can generate electrical current in MFCs (Xing et al., 2008; Nishio et al., 2010). However, it remains unclear whether this dominant ME1 *Prothecochloris* bacterium is involved in current

Table 1 | Most abundant OTUs in the microelectrode-associated community.

Dominant OTUs	Family	Genus	Relative abundance (%)	
			Neighboring mat	Microelectrode
ME1	*Chlorobiaceae*	*Prosthecochloris*	4.59	30.34
ME2	*Rhodobacteraceae*	*Loktanella*	0.14	9.40
ME3	*Rhodobacteraceae*	Unclassified	0.09	9.17
ME4	*Rhodobacteraceae*	*Catellibacterium*	0.51	7.35
ME5	*Rhodobacteraceae*	Unclassified	0.12	4.88
ME6	*Piscirickettsiaceae*	*Thiomicrospira*	0.18	1.97

generation from our polarized microelectrode (**Figure 6**). In addition, OTU ME6 is classified within *Thiomicrospira,* which is also known to be chemolithoautotrophic with reduced sulfur compounds such as thiosulfate, elemental sulfur and sulfide as electron donors to reduce oxygen molecules (Brinkhoff et al., 1999; Takai et al., 2004). It is possible that these microorganisms were enriched because the microelectrode tip happened to be placed precisely at a depth in the mat that frequently contains both trace amounts of oxygen and sulfide or cycles between oxic and sulfidic conditions, which is underscored by phylogenetic evidence of a transition in metabolism from aerobic to microaerobic or anaerobic at equivalent depth in the mat (Lindemann et al., 2013). However, the microelectrode-associated community contains organisms with seemingly incompatible metabolisms. *Loktanella* members, for example, known to be strict aerobes (Van Trappen et al., 2004), were significant members of the community alongside *Prosthecochloris.* The curious co-enrichment of the strictly anaerobic *Prosthecochloris* and strictly aerobic *Loktanella, Catellibacterium,* and *Thiomicrospira* on the microelectrode tip requires further investigation integrating electrochemical function with activity of isolated members.

Recent studies have called into question the utility of PacBio sequencing for 16S rRNA analysis of microbial communities due to its error rate (Mosher et al., 2013); therefore, our quality filtration of raw sequences was extremely strict. As Mosher et al. (2013) did not report any quality filtration post-sequencing for PacBio reads, it is not possible to directly compare the stringency of our approach or our final error rate post-filtration to theirs. If the increased error rate of PacBio did proliferate clusters that are, in reality, derived from the same parent sequences in the communities, abnormally high alpha diversity is expected as each real OTU is, essentially, counted multiple times (Mosher et al., 2013). However, our inverse Simpson metric for the mat community was 48.351, which is similar to the Hot Lake mat at an equivalent depth as assayed using Itags (cf. **Figure 8**; Lindemann et al., 2013) and much lower for the microelectrode-associated community (8.462). Therefore, we did not detect abnormally high species richness in our communities, suggesting that our quality filtration was stringent enough to account for the error rates associated with PacBio in our comparative analysis.

COMBINING SULFIDE ELECTROCHEMISTRY WITH MICROBIAL COMMUNITY ANALYSIS

Because of its importance to environmental remediation of sulfide-contaminated industrial waste, sulfide electrochemical detection has been well-studied on carbon-based electrodes (Lawrence et al., 2004). According to Lawrence et al. (2004), HS^- is oxidized electrochemically to elemental sulfur on the electrode surface and acts as a passivation layer. Therefore, on unmodified carbon electrodes, sulfide oxidation to elemental sulfur is an unstable process with decreasing current as passivation proceeds and follows the half reaction listed in **Table 2**. Although the standard reduction potential is negative, anodic current is typically observed at more positive potentials on unmodified carbon electrodes due to a high overpotential and irreversibility of sulfur redox systems (Bard et al., 1985). Lawrence et al. (2004) observed sulfide oxidation to occur at approximately $+200$ $mV_{Ag/AgCl}$ in 50 mM phosphate buffer at pH 7.4 on bare glassy carbon surface. At pH 10, that value would shift to approximately $+274$ $mV_{Ag/AgCl}$. To a first approximation then, the polarization of the carbon fiber microelectrode tip at $+400$ $mV_{Ag/AgCl}$ would likely oxidize any HS^- diffusing to the microelectrode tip to elemental sulfur. Considering this, the microelectrode tip should have exhibited decreasing current due to sulfur passivation because Hot Lake mats are known to generate sulfide in the dark which is also verified with microelectrode measurements (data not shown). However, the increase in current during light periods cannot be explained electrochemically.

Elemental sulfur cannot be electrochemically oxidized by the microelectrode tip and, therefore, accumulates on the surface. The relative abundance of *Prosthecochloris* at the microelectrode tip suggests that the accumulated elemental sulfur may be oxidized to soluble thiosulfate ($S_2O_3^{2-}$) and sulfate (SO_4^{2-}) biotically if the region surrounding the microelectrode is persistently anoxic (Imhoff, 2003; Muyzer and Stam, 2008; Sun et al., 2010; Liang et al., 2013). If biotic oxidation of elemental sulfur did occur at the microelectrode tip, then a reasonable hypothesis for increasing current during light exposure can be formulated. The hypothesized sulfur cycle is shown in **Figure 8** where green phototrophic sulfur bacteria (*Prosthecochloris*) and aerobic sulfur-oxidizer (*Thiomicrospira*) oxidize elemental sulfur to sulfate in the presence of light. Sulfate is then reduced back to HS^- biotically by sulfate-reducing bacteria (e.g., *Desulfobacteraceae*) present in the mat. The sulfur cycle, completed by the electrochemical oxidation of HS^- at the microelectrode tip, can occur indefinitely until HS^- is depleted. Assuming that turnover of elemental sulfur by green phototrophic sulfur bacteria is the limiting step, enrichment of these members over time would increase the baseline current seen at the microelectrode tip. Furthermore, the inhibition of elemental sulfur turnover in the dark would also explain the oscillation of current. Therefore, the trend seen in **Figure 6** can be reasonably explained by the proposed bioelectrochemical cycling of sulfur.
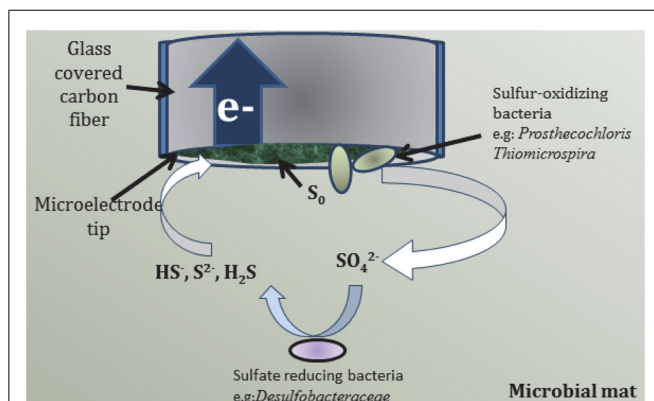

FIGURE 8 | Proposed pathway of bioelectrochemical cycling of sulfur species at the microelectrode tip.

Table 2 | Selected reduction reactions of relevant sulfur compounds (taken from Bard et al., 1985).

Type	Standard half reaction	Standard reduction potential V vs. Ag/AgCl(sat)
Sulfide to sulfur	$HS^- \rightarrow S + 2e^- + H^+$	−0.259
Polysulfide to sulfide*	$S_2^{2-} + 2e^- + 2H^+ \rightarrow 2HS^-$	0.090
	$S_3^{2-} + 4e^- + 3H^+ \rightarrow 3HS^-$	−0.100
Sulfate to sulfite	$SO_4^{2-} + 2e^- + 4H^+ \rightarrow H_2SO_3 + H_2O$	−0.039
Tetrathionate to thiosulfate	$S_4O_6^{2-} + 2e^- \rightarrow 2S_2O_3^{2-}$	−0.117
Thiosulfate to sulfide*	$S_2O_3^{2-} + 8e^- + 8H^+ \rightarrow 2HS^- + 3H_2O$	0.021

*Occurs through multiple electron transfer steps (Schippers, 2004).

The above discussion on sulfur cycling remains speculative because sulfur speciation in sediments is quite complex as the type of (metal) sulfide and solubility affects the mechanism of abiotic and biotic oxidation to sulfate (Schippers, 2004). Here, both the thiosulfate or polysulfide mechanisms of sulfide oxidation could account for the various sulfur species. Thiosulfate, sulfite, polythionates, polysulfides, and elemental sulfur are all possible intermediates that could play an important role in sulfur cycling in Hot Lake mats. Of equal importance is the availability of iron, oxygen, manganese, and nitrate (Jørgensen and Nelson, 2004) in the mat at depth. Finally, the role of calcium sulfates and magnesium sulfates in Hot Lake (Lindemann et al., 2013) on sulfur cycling is unknown. The possibility of targeting a particular sulfur compound redox potential on the microelectrode is interesting in terms of the ability to select for sub-populations of heterotrophs existing only along steep physicochemical gradients inside the mat. Carefully selecting the polarization potential of the microelectrode would tune into specific redox reactions (i.e., sulfur) that are utilized by targeted heterotrophs. Inspection of several standard reduction potentials of sulfur reactions listed in **Table 2** suggests that we could electrochemically differentiate sulfur reactions and thereby select those microorganisms that could metabolize the resulting products. We also note that sulfur oxidation is not the only exploitable redox reaction, as polarizing the microelectrode ca. -200 mV$_{Ag/AgCl}$ would tune into oxygen reduction and remove oxygen at the microelectrode surface. However, these electrochemical activities remain unexplored in Hot Lake microbial mats.

CONCLUSION

We quantified physicochemical gradients (redox potential, pH, and dissolved oxygen) and their diel variations within the Hot Lake mat. We further employed a mat fuel cell to demonstrate that phototrophic microbial mats can generate electricity, which increases upon exposure to light. We also found that a microelectrode with a carbon tip can be used to study local electron transfer processes in a microbial mat. Anodically polarizing a microelectrode tip showed increased current with time, reaching a maximum during the day period of a day–night cycle. The increased current at the microelectrode tip indicated presence of electroactive compounds near the microelectrode tip. 16S rRNA analysis revealed that the bacterial community attached to the polarized microelectrode tip was distinct from that of the neighboring microbial mat. The reduced alpha diversity of the microelectrode-associated community and the large distance in beta diversity between this community and the neighboring mat, driven by differences in the relative abundances of reads attributed to phyla *Proteobacteria* and *Chlorobi*, suggested that the polarized microelectrode tip may locally enhance the growth of certain bacterial phylotypes. Furthermore, the characteristics of the most abundant OTUs in the microelectrode-associated community suggested that the current cycle obtained from our polarized microelectrode may be related to bacterial sulfur cycling. It remains unclear whether electron transfer to the electrode surface during current generation was direct or indirect. Our results describe a new method for monitoring of local electron transfer rates within a microbial mat and subsequent assaying of the adherent community developed at the microscale. This method can contribute to future electron transfer studies and aid in our ability to enrich specific subpopulations *in situ* within microbial mats.

REFERENCES

Anderson, G. C. (1958). Some limnological features of a shallow saline meromictic lake. *Limnol. Oceanogr.* 3, 259–270. doi: 10.4319/lo.1958.3.3.0259

Babauta, J., Renslow, R., Lewandowski, Z., and Beyenal, H. (2012a). Electrochemically active biofilms: facts and fiction. A review. *Biofouling* 28, 789–812. doi: 10.1080/08927014.2012.710324

Babauta, J. T., Nguyen, H. D., Harrington, T. D., Renslow, R., and Beyenal, H. (2012b). pH, redox potential and local biofilm potential microenvironments within Geobacter sulfurreducens biofilms and their roles in electron transfer. *Biotechnol. Bioeng.* 109, 2651–2662. doi: 10.1002/Bit.24538.

Babauta, J. T., Nguyen, H. D., Istanbullu, O., and Beyenal, H. (2013). Microscale gradients of oxygen, hydrogen peroxide, and pH in freshwater cathodic biofilms. *Chemsuschem* 6, 1252–1261. doi: 10.1002/cssc.201300019

Badalamenti, J. P., Torres, C. I., and Krajmalnik-Brown, R. (2013). Light-responsive current generation by phototrophically enriched anode biofilms dominated by green sulfur bacteria. *Biotechnol. Bioeng.* 110, 1020–1027. doi: 10.1002/Bit.24779

Bard, A.J., Parsons, R., and Jordan, J. (eds). (1985). *Standard Potentials in Aqueous Solution.* New York, NY: Marcel Dekker, Inc.

Behrens, S., Lösekann, T., Pett-Ridge, J., Weber, P. K., Ng, W.-O., Stevenson, B. S., et al. (2008). Linking microbial phylogeny to metabolic activity at the single-cell level by using enhanced element labeling-catalyzed reporter deposition fluorescence in situ hybridization (EL-FISH) and NanoSIMS. *Appl. Environ. Microbiol.* 74, 3143–3150. doi: 10.1128/aem.00191-08.

Beyenal, H., Tanyolac, A., and Lewandowski, Z. (1998). Measurement of local effective diffusivity in heterogeneous biofilms. *Water Sci. Technol.* 38, 171–178. doi: 10.1016/S0273-1223(98)00691-X

Bishop, P. L., and Yu, T. (1999). A microelectrode study of redox potential change in biofilms. *Water Sci. Technol.* 39, 179–185.

Borole, A. P., Reguera, G., Ringeisen, B., Wang, Z. W., Feng, Y. J., and Kim, B. H. (2011). Electroactive biofilms: current status and future research needs. *Energy Environ. Sci.* 4, 4813–4834. doi: 10.1039/C1ee02511b

Bradley, R. W., Bombelli, P., Rowden, S. J. L., and Howe, C. J. (2012). Biological photovoltaics: intra- and extra-cellular electron transport by cyanobacteria. *Biochem. Soc. Trans.* 40, 1302–1307. doi: 10.1042/Bst20120118

Brinkhoff, T., Muyzer, G., Wirsen, C. O., and Kuever, J. (1999). *Thiomicrospira kuenenii* sp. nov. and *Thiomicrospira frisia* sp. nov., two mesophilic obligately chemolithoautotrophic sulfur-oxidizing bacteria isolated from an intertidal mud flat. *Int. J. Syst. Bacteriol.* 49, 385–392. doi: 10.1099/00207713-49-2-385

Chandra, R., Subhash, G. V., and Mohan, S. V. (2012). Mixotrophic operation of photo-bioelectrocatalytic fuel cell under anoxygenic microenvironment enhances the light dependent bioelectrogenic activity. *Bioresour. Technol.* 109, 46–56. doi: 10.1016/j.biortech.2011.12.135

de Beer, D., Stoodley, P., Roe, F., and Lewandowski, Z. (1994). Effects of biofilm structures on oxygen distribution and mass transport. *Biotechnol. Bioeng.* 43, 1131–1138. doi: 10.1002/bit.260431118

Dewan, A., Donovan, C., Heo, D., and Beyenal, H. (2010). Evaluating the performance of microbial fuel cells powering electronic devices. *J. Power Sources* 195, 90–96. doi: 10.1016/j.jpowsour.2009.07.001

Erable, B., Féron, D., and Bergel, A. (2012). Microbial catalysis of the oxygen reduction reaction for microbial fuel cells: a review. *Chemsuschem* 5, 975–987. doi: 10.1002/cssc.201100836

Franks, J., and Stolz, J. F. (2009). Flat laminated microbial mat communities. *Earth Sci. Rev.* 96, 163–172. doi: 10.1016/j.earscirev.2008.10.004

Gorlenko, V. M. (2001). "Genus V. Prosthecochloris," in *Bergey's Manual of Systematic Bacteriology*, eds D. R. Boone, R. W. Castenholz, and G. M. Garrity (New York: Springer), 617–620.

Guerrero, R., Piqueras, M., and Berlanga, M. (2002). Microbial mats and the search for minimal ecosystems. *Int. Microbiol.* 5, 177–188. doi: 10.1007/s10123-002-00948.

He, Z., Kan, J., Mansfeld, F., Angenent, L. T., and Nealson, K. H. (2009). Self-sustained phototrophic microbial fuel cells based on the synergistic cooperation between photosynthetic microorganisms and heterotrophic bacteria. *Environ. Sci. Technol.* 43, 1648–1654. doi: 10.1021/Es803084a

Hori, K., and Matsumoto, S. (2010). Bacterial adhesion: from mechanism to control. *Biochem. Eng. J.* 48, 424–434. doi: 10.1016/j.bej.2009.11.014

Imhoff, J. F. (2003). Phylogenetic taxonomy of the family Chlorobiaceae on the basis of 16S rRNA and fmo (Fenna–Matthews–Olson protein) gene sequences. *Int. J. Syst. Evol. Microbiol.* 53, 941–951. doi: 10.1099/ijs.0.02403-0

Jørgensen, B. B., and Nelson, D. C. (2004). Sulfide oxidation in marine sediments: geochemistry meets microbiology. *Geol. Soc. Am. Special Pap.* 379, 63–81. doi: 10.1130/0-8137-2379-5.63

Jørgensen, B. B., and Revsbech, N. P. (1983). Colorless sulfur bacteria, Beggiatoa spp. and Thiovulum spp., in O2 and H2S microgradients. *Appl. Environ. Microbiol.* 45, 1261–1270.

Kroes, I., Lepp, P. W., and Relman, D. A. (1999). Bacterial diversity within the human subgingival crevice. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14547–14552. doi: 10.1073/pnas.96.25.14547

Kühl, M., Lassen, C., and Jørgensen, B. (1994). "Optical properties of microbial mats: light measurements with fiber-optic microprobes," in *Microbial Mats*, eds L. Stal and P. Caumette (Berlin : Springer), 149–166.

Lan, J. C. W., Raman, K., Huang, C. M., and Chang, C. M. (2013). The impact of monochromatic blue and red LED light upon performance of photo microbial fuel cells (PMFCs) using *Chlamydomonas reinhardtii* transformation F5 as biocatalyst. *Biochem. Eng. J.* 78, 39–43. doi: 10.1016/j.bej.2013.02.007

Lawrence, N. S., Deo, R. P., and Wang, J. (2004). Electrochemical determination of hydrogen sulfide at carbon nanotube modified electrodes. *Anal. Chim. Acta* 517, 131–137. doi: 10.1016/j.aca.2004.03.101

Lewandowski, Z., and Beyenal, H. (2014). *Fundamentals of Biofilm Research.* Boca Raton, FL: Taylor & Francis Group, LLC.

Liang, F.-Y., Deng, H., and Zhao, F. (2013). Sulfur pollutants treatment using microbial fuel cells from perspectives of electrochemistry and microbiology. *Chin. J. Anal. Chem.* 41, 1133–1139. doi: 10.1016/S1872-2040(13)60669-6

Lin, C. C., Wei, C. H., Chen, C. I., Shieh, C. J., and Liu, Y. C. (2013). Characteristics of the photosynthesis microbial fuel cell with a *Spirulina platensis* biofilm. *Bioresour. Technol.* 135, 640–643. doi: 10.1016/j.biortech.2012.09.138

Lindemann, S. R., Moran, J. J., Stegen, J. C., Renslow, R. S., Hutchison, J. R., Cole, J. K., et al. (2013). The epsomitic phototrophic microbial mat of hot lake, Washington: community structural responses to seasonal cycling. *Front. Microbiol.* 4:323. doi: 10.3389/fmicb.2013.00323

Malvankar, N. S., and Lovley, D. R. (2012). Microbial nanowires: a new paradigm for biological electron transfer and bioelectronics. *Chemsuschem* 5, 1039–1046. doi: 10.1002/cssc.201100733

Mosher, J. J., Bernberg, E. L., Shevchenko, O., Kan, J., and Kaplan, L. A. (2013). Efficacy of a 3rd generation high-throughput sequencing platform for analyses of 16S rRNA genes from environmental samples. *J. Microbiol. Methods* 95, 175–181. doi: 10.1016/j.mimet.2013.08.009

Nguyen, H. D., Cao, B., Mishra, B., Boyanou, M. I., Kemner, K. M., Fredrickson, J. K., et al. (2012). Microscale geochemical gradients in Hanford 300 Area sediment biofilms and influence of uranium. *Water Res.* 46, 227–234. doi: 10.1016/j.watres.2011.10.054

Nishio, K., Hashimoto, K., and Watanabe, K. (2010). Light/electricity conversion by a self-organized photosynthetic biofilm in a single-chamber reactor. *Appl. Microbiol. Biotechnol.* 86, 957–964. doi: 10.1007/s00253-009-2400-2

Reimers, C. E., Girguis, P., Stecher, H. A., Tender, L. M., Ryckelynck, N., and Whaling, P. (2006). Microbial fuel cell energy from an ocean cold seep. *Geobiology* 4, 123–136. doi: 10.1111/j.1472-4669.2006.00071.x

Renslow, R. S., Babauta, J. T., Dohnalkova, A. C., Boyanov, M. I., Kemner, K. M., Majors, P. D., et al. (2013). Metabolic spatial variability in electrode-respiring *Geobacter sulfurreducens* biofilms. *Energy Environ. Sci.* 6, 1827–1836. doi: 10.1039/C3ee40203g

Revsbech, N. P., and Ward, D. M. (1984). Microelectrode studies of interstitial water chemistry and photosynthetic activity in a hot spring microbial mat. *Appl. Environ. Microbiol.* 48, 270–275.

Roeselers, G., Loosdrecht, M. C. M. V., and Muyzer, G. (2008). Phototrophic biofilms and their potential applications. *J. Appl. Phycol.* 20, 227–235. doi: 10.1007/s10811-007-9223-2

Schippers, A. (2004). Biogeochemistry of metal sulfide oxidation in mining environments, sediments, and soils. *Geol. Soc. Am. Special Pap.* 379, 49–62. doi: 10.1130/0-8137-2379-5.49

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Shi, L. A., Richardson, D. J., Wang, Z. M., Kerisit, S. N., Rosso, K. M., Zachara, J. M., et al. (2009). The roles of outer membrane cytochromes of Shewanella and Geobacter in extracellular electron transfer. *Environ. Microbiol. Rep.* 1, 220–227. doi: 10.1111/j.1758-2229.2009.00035.x

Snider, R. M., Strycharz-Glaven, S. M., Tsoi, S. D., Erickson, J. S., and Tender, L. M. (2012). Long-range electron transport in *Geobacter sulfurreducens* biofilms is redox gradient-driven. *Proc. Natl. Acad. Sci. U.S.A.* 109, 15467–15472. doi: 10.1073/pnas.1209829109

Strycharz-Glaven, S. M., Glaven, R. H., Wang, Z., Zhou, J., Vora, G. J., and Tender, L. M. (2013). Electrochemical investigation of a microbial solar cell reveals a nonphotosynthetic biocathode catalyst. *Appl. Environ. Microbiol.* 79, 3933–3942. doi: 10.1128/Aem.00431-13

Sun, M., Tong, Z.-H., Sheng, G.-P., Chen, Y.-Z., Zhang, F., Mu, Z.-X., et al. (2010). Microbial communities involved in electricity generation from sulfide oxidation in a microbial fuel cell. *Biosens. Bioelectron.* 26, 470–476. doi: 10.1016/j.bios.2010.07.074

Takai, K., Hirayama, H., Nakagawa, T., Suzuki, Y., Nealson, K. H., and Horikoshi, K. (2004). *Thiomicrospira thermophila* sp. nov., a novel microaerobic, thermotolerant, sulfur-oxidizing chemolithomixotroph isolated from a deep-sea hydrothermal fumarole in the TOTO caldera, Mariana Arc, Western Pacific. *Int. J. Syst. Evol. Microbiol.* 54, 2325–2333. doi: 10.1099/ijs.0.63284-0

Timothy, J. G., Stefan, H., James, T. N., and Paul, G. T. (2011). "Introduction to aquatic redox chemistry," in *Aquatic Redox Chemistry*, eds P. G. Tratnyek, T. J. Grundl, and S. B. Haderlein (Washington, DC: ACS Symposium Series) 1071, 1–14.

van der Meer, M. T. J., Schouten, S., Bateson, M. M., Nübel, U., Wieland, A., Kühl, M., et al. (2005). Diel variations in carbon metabolism by green nonsulfur-like bacteria in alkaline siliceous hot spring microbial mats from Yellowstone National Park. *Appl. Environ. Microbiol.* 71, 3978–3986. doi: 10.1128/aem.71.7.3978-3986.2005

Van Trappen, S., Mergaert, J., and Swings, J. (2004). *Loktanella salsilacus* gen. nov., sp. nov., *Loktanella fryxellensis* sp. nov. and *Loktanella vestfoldensis* sp. nov., new members of the Rhodobacter group, isolated from microbial mats in Antarctic lakes. *Int. J. Syst. Evol. Microbiol.* 54, 1263–1269. doi: 10.1099/ijs.0.3006-0

Vargas, M., Malvankar, N. S., Tremblay, P. L., Leang, C., Smith, J. A., Patel, P., et al. (2013). Aromatic amino acids required for pili conductivity and long-range extracellular electron transport in *Geobacter sulfurreducens*. *Mbio* 4, e00105–e00113. doi: 10.1128/mBio.00105–13

Xing, D., Zuo, Y., Cheng, S., Regan, J. M., and Logan, B. E. (2008). Electricity generation by *Rhodopseudomonas palustris* DX-1. *Environ. Sci. Technol.* 42, 4146–4151. doi: 10.1021/es800312v

Microsensor measurements of hydrogen gas dynamics in cyanobacterial microbial mats

Michael Nielsen[1], Niels P. Revsbech[1] and Michael Kühl[2,3]*

[1] Section of Microbiology, Department of Bioscience, Aarhus University, Aarhus, Denmark, [2] Marine Biological Section, Department of Biology, University of Copenhagen, Helsingør, Denmark, [3] Plant Functional Biology and Climate Change Cluster, University of Technology, Sydney, Ultimo, NSW, Australia

We used a novel amperometric microsensor for measuring hydrogen gas production and consumption at high spatio-temporal resolution in cyanobacterial biofilms and mats dominated by non-heterocystous filamentous cyanobacteria (*Microcoleus chtonoplastes and Oscillatoria sp.*). The new microsensor is based on the use of an organic electrolyte and a stable internal reference system and can be equipped with a chemical sulfide trap in the measuring tip; it exhibits very stable and sulfide-insensitive measuring signals and a high sensitivity (1.5–5 pA per μmol L^{-1} H_2). Hydrogen gas measurements were done in combination with microsensor measurements of scalar irradiance, O_2, pH, and H_2S and showed a pronounced H_2 accumulation (of up to 8–10% H_2 saturation) within the upper mm of cyanobacterial mats after onset of darkness and O_2 depletion. The peak concentration of H_2 increased with the irradiance level prior to darkening. After an initial build-up over the first 1–2 h in darkness, H_2 was depleted over several hours due to efflux to the overlaying water, and due to biogeochemical processes in the uppermost oxic layers and the anoxic layers of the mats. Depletion could be prevented by addition of molybdate pointing to sulfate reduction as a major sink for H_2. Immediately after onset of illumination, a short burst of presumably photo-produced H_2 due to direct biophotolysis was observed in the illuminated but anoxic mat layers. As soon as O_2 from photosynthesis started to accumulate, the H_2 was consumed rapidly and production ceased. Our data give detailed insights into the microscale distribution and dynamics of H_2 in cyanobacterial biofilms and mats, and further support that cyanobacterial H_2 production can play a significant role in fueling anaerobic processes like e.g., sulfate reduction or anoxygenic photosynthesis in microbial mats.

Keywords: hydrogen, oxygen, sulfide, pH, irradiance, microsensor, microbial mat, cyanobacteria

Introduction

Molecular hydrogen (H_2) is produced during fermentative anaerobic degradation of organic matter (Conrad, 1988). Formation of H_2 as a by-product of nitrogenase activity is also described as a major H_2 producing process (Bothe et al., 2010). In excess of energy and reducing power, photosynthetic bacteria can also photo-produce H_2 (Gest and Kamen, 1949; Warthmann et al., 1992). Hydrogen is a very good energy source that is readily reacting with O_2

(chemically or catalyzed by "Knallgas"-bacteria) or is consumed by anaerobic mineralization processes, e.g., as an electron donor in sulfate reduction and in methanogenesis (Hoehler et al., 2002). Anoxygenic photosynthetic bacteria can also utilize hydrogen as an electron donor (Overmann and Garcia-Pichel, 2000).

Efficient inter-species hydrogen transfer in consortia of microorganisms allows for syntrophic processes, which separately would otherwise be energetically unfavorable (Wolin, 1982; Hoehler et al., 2001). Consequently, only very low H_2 concentrations are detected in most natural environments. Higher H_2 levels can, however, be found in special environments like the digestive tracts of termites (Ebert and Brune, 1997) or in legume nodules harboring N_2-fixing bacteria (Witty, 1991). Geothermal features can also exhibit high levels of H_2, and hydrogenotrophs are widespread in many hot springs, where H_2 metabolism can be predominant (Spear et al., 2005).

Hydrogen production in cyanobacteria has been known for a long time (Jackson and Ellms, 1896; Benemann and Weare, 1974; Oschchepkov et al., 1974) and has been studied for a large number of strains and in different environments (Lambert and Smith, 1981; Houchins, 1984; Kothari et al., 2012; Otaki et al., 2012). In recent years, cyanobacterial H_2 production has also become a major research topic in connection with the search for new clean energy generating processes (e.g., Lee et al., 2010; Hallenbeck, 2012). Production of H_2 in cyanobacteria is primarily associated with N_2 fixation, where H_2 is a major by-product (Bothe et al., 2010), or due to dark fermentation of storage products accumulated during daytime photosynthesis (Moezelaar et al., 1996; Stal and Moezelaar, 1997). In a survey of different cyanobacterial strains, Kothari et al. (2014) demonstrated that fermentative pathways and bidirectional NADH-linked [Ni-Fe] hydrogenases are of prime importance for H_2 production under dark anoxic conditions in the filamentous non-heterocystous cyanobacteria *Microcoleus chtonoplastes* and *Lyngbya aestuarii* that form dense microbial mats in coastal and hypersaline environments; these species exhibited higher production rates and reached much higher steady state H_2 concentrations than many other cyanobacteria. Bidirectional hydrogenases are also involved in light-driven H_2 formation under anoxia via direct biophotolysis, i.e., the light-driven splitting of water (Appel et al., 2000).

Earlier reports of significant H_2 production in cyanobacterial mats were based on gas chromatographic analysis of intact mats (Skyring et al., 1988, 1989) and of gas bubbles carefully sampled from the surface of intact hypersaline cyanobacterial mats (Hoehler et al., 2001). The latter study, and more recently Burow et al. (2012) studying a coastal microbial mat also obtained a coarse depth distribution of H_2 production by incubating 2 mm thick slices of mats from different depth horizons below the surface, showing maximal H_2 production in the top 2 mm of the mat during night time. These findings were used to hypothesize that H_2 production by ancient microbial mats and subsequent escape of the H_2 to space was a major mechanism for facilitating the oxidation of the primitive Earth (Hoehler et al., 2001; Jørgensen, 2001). A series

of elegant follow up studies combining biogeochemical process measurements with modern molecular tools, have (i) identified filamentous non-heterocystous cyanobacteria as the major H_2 producers in such mats (Burow et al., 2012; Marschall et al., 2012), (ii) demonstrated cyanobacterial fermentation as the major H_2 producing process (Burow et al., 2012; Lee et al., 2014), and (iii) demonstrated that sulfate reducing bacteria (SRB) are predominant hydrogenotrophs in cyanobacterial mats (Burow et al., 2014). There is thus increasing evidence that fermentative H_2 and organic acid production is a key component in microbial mat biogeochemistry facilitating close interactions between cyanobacteria, anoxygenic phototrophs and heterotrophic bacteria (Otaki et al., 2012; Lee et al., 2014).

Despite an increasing interest in understanding the production and consumption of H_2 in the environment, very few studies have described the fine scale distribution and dynamics of H_2 (Witty, 1991; Ebert and Brune, 1997). Part of the reason has been lack of suitable technology (Hübert et al., 2011). Conventional Clark-type electrochemical H_2 microsensors, which are based on the oxidation of H_2 at a positively charged platinum electrode in an acidic KCl containing electrolyte (Wang et al., 1971; Witty, 1991), often suffer from unstable signals and calibration drift when used in natural systems. This has limited their applicability, especially in environments like sediments and microbial mats, where H_2 measurements were hampered by sulfide interference on the measuring signal.

A sulfide-sensitive amperometric H_2 microsensor based on the use of non-aqueous electrolyte is commercially available (Unisense A/S, Denmark), and a robust version of the sensor has proven useful for quantifying H_2 production in vials with cyanobacterial cultures (Kothari et al., 2012, 2014). This sensor has very recently been employed for first microscale H_2 measurement in intertidal microbial mats (Hoffmann et al., 2015) showing pronounced accumulation and efflux of H_2 in darkness driven by cyanobacterial fermentation in the upper mm's of the mat. The microenvironmental dynamics of H_2 in hypersaline water covered mats remain to be studied in more detail, as these mat types often exhibit high sulfide levels causing interference on the commercial H_2 microsensor The sensor design has now been further improved and a sulfide-insensitive H_2 microsensor was recently developed (Nielsen et al., 2015). In the present study we use these microsensors for studying the H_2 microenvironment and its relation to light, O_2, pH and H_2S micro gradients in coastal and hypersaline microbial mats. We demonstrate pronounced H_2 dynamics during experimental light-dark shifts, and discuss the role of H_2 for biogeochemical processes in microbial mats.

Materials and Methods

Experimental Setup

We studied H_2 dynamics in two different microbial mats, both harboring a 1–3 mm thick dark-green surface layer with dense populations of filamentous non-heterocystous cyanobacteria, and anoxygenic *Chloroflexi*-like phototrophs.

Hypersaline Mat

Dense biofilms of filamentous cyanobacteria were retrieved from the top layer of a hypersaline microbial mat sampled in a salt evaporation pond of Saline de Giraud, Camargue, France. The mat locality and a detailed description of the biogeochemistry and microbial composition of the mat are presented elsewhere (Fourcans et al., 2004; Wieland et al., 2005). Microbial mat samples were transported to our laboratory and were kept in trays with aerated brine at *in situ* salinity ($\sim$80–100 ppt) under a 12 h light-12 h dark period in a thermostated room at 16°C. The mat was covered by a 2-3 mm thick deep-green biofilm of motile filamentous cyanobacteria. Microscopic investigations of the biofilm showed a dominance of morphotypes similar to *M. chtonoplastes* mixed with other motile filaments of *Oscillatoria* sp. and *Spirulina* sp. The upper millimeters of the mat remained non-sulfidic due to a conspicuous layer of oxidized iron below the cyanobacterial layer that buffered against sulfide formation in the uppermost mat layer during night-time (see details in Wieland et al., 2005).

Prior to experiments, a small piece of the surface biofilm was transferred to a small 3–4 mm high and 8 mm wide glass beaker with a thin layer of semisolid agar at $\sim$38–40°C. During subsequent cooling the biofilm bottom and side adhered to the solidifying agar leaving the upper biofilm surface uncovered. The small beaker was mounted in a flow chamber (Lorenzen et al., 1995) with the biofilm surface flush with a larger agar slab, and aerated brine (90 ppt, 25°C, pH 8) was constantly circulated over the biofilm surface. The air-saturated brine contained 154 μmol O_2 l^{-1} according to a table compiled from published empirical solubility equations[1]. Illumination was provided with a fiber-optic halogen lamp equipped with a collimating lens (KL-2500, Schott, Germany), where the irradiance was regulated with a built-in neutral density screen. During long-term incubations the lamp was turned on and off at defined times by an electrical switch with a timer. Irradiance levels at defined lamp settings were determined with a quantum irradiance meter (LI-250, LiCor, USA) equipped with a small spherical irradiance sensor (Walz GmbH, Germany).

Coastal Mat

Coastal microbial mat samples were collected in small acrylate coring tubes from the upper air-exposed yet moist part of a sandbar in Limfjorden near Aggersund, Denmark (57°00′02.15N; 9°17′12.89E). The mat undergoes irregular cycles of inundation and air exposure depending on prevailing wind directions and consisted of well sorted fine grained sand bound together by a dense 2–3 mm top layer of motile filamentous cyanobacteria (*M. chtonoplastes* and *Oscillatoria* sp.) and some green filamentous anoxygenic phototrophs *Chloroflexus* sp. and exopolymers on top of a black sulfidic layer, which also contained filamentous sulfide oxidizing bacteria (*Beggiatoa* sp.; Lassen et al., 1992). The mat samples were incubated for 1–2 days in aerated seawater under moderate illumination by halogen lamps ($\sim$100–200 μmol photons m^{-2} s^{-1}) prior to experiments.

During this time, the surface became densely covered by a dense layer of filamentous cyanobacteria (Supplementary Figure S1A). For comparison, we also sampled and investigated permanently submerged sediment samples from the same locality that were predominated by a dense benthic diatom film (Supplementary Figure S1B); these samples were obtained at the same location but from a sandier and less sulfidic sediment that was permanently water covered.

Experiments were conducted with the core samples mounted in an aquarium 1 cm below the surface of continuously aerated seawater (25 ppt, 21–22°C), which was circulated over the mat by a gentle airstream from a Pasteur pipette. The air-saturated seawater contained 240 μmol O_2 l^{-1} according to a table compiled from published empirical solubility equations[1]. Illumination was provided by a halogen lamp bulb, where the irradiance was regulated by varying the distance to the mat surface. During long-term incubations, the lamp was turned on and off at defined times by an electrical switch with a timer. Irradiance levels at defined lamp distance were determined with a quantum irradiance meter (LI-250, LiCor, USA) equipped with a small spherical irradiance sensor (Walz GmbH, Germany).

Inhibition of sulfate reduction in the mat was done by incubating a mat sample in aerated and stirred sea water with 2.5 mM sodium molybdate for 6 h prior to measurements. Using published diffusion coefficients (D) of molybdate in water ($9.91 \cdot 10^{-6}$ cm^2 s^{-1}; Li and Gregory, 1974) and gel ($6.48 \cdot 10^{-6}$ cm^2 s^{-1}; Mason et al., 2005), we estimated the penetration depth of molybdate after $t = 6$ h of incubation as $L = \sqrt{2Dt}$ assuming a one-dimensional diffusion geometry (Berg, 1983). This showed that after 6 h of incubation, molybdate penetrated about 5.3–6.5 mm into the microbial mat ensuring sufficient exposure of SRB's in the complete photic zone as well as several mm's of the underlaying aphotic zone of the mat.

Microsensor Measurements of O_2, pH, H_2S, and H_2

Chemical microprofiles were measured with electrochemical microsensors for O_2, H_2S, pH and H_2 with tip diameters of 10–70 μm (Unisense A/S, Denmark). Construction of electrochemical O_2, pH, and H_2S microsensors, their calibration and application have been described in previous publications (Revsbech and Jørgensen, 1986; Revsbech, 1989; Kühl et al., 1996, 1998; Kühl and Revsbech, 2001).

The H_2 microsensor is constructed like a Clark-type O_2 microsensor (Revsbech, 1989) and consists of an outer casing sealed by a thin silicon rubber membrane, and an internal measuring microanode polarized at +0.6 to +1.0 V relative to an internal reference electrode. The casing is filled with an organic electrolyte and this configuration facilitates a stable measurement of H_2 via its oxidation at the measuring anode. The H_2 microsensor is commercially available and further details on the sensor and its calibrations can be obtained from the manufacturer's website[2]. We tested the interference of several compounds, which can pass the silicon membrane of the H_2 microsensor and react at the measuring anode. Compounds like

---

[1]http://www.unisense.com/files/PDF/Diverse/Seawater%20&%20Gases%20 table.pdf

[2]www.unisense.com

dimethyl sulfide (DMS) and methyl mercaptan can strongly affect the microsensor performance and seem to have a poisoning effect, but levels of these potential interfering agents in microbial mats have been found to be in the lower nM range (Visscher et al., 2003) and should thus not affect our measurements significantly. We found no interference from carbon monoxide, which has been shown to be present in hypersaline microbial mats during daytime (Hoehler et al., 2001) and is a known interfering agent of amperometric H_2 sensors with aquatic electrolytes (Hübert et al., 2011). No sensitivity to light was observed. In the absence of a H_2S shield, dissolved H_2S is a major interfering substance and for a given concentration gives rise to a signal of $\sim$20–30% of the signal measured for the same concentration of H_2 (Nielsen et al., 2015). However, by mounting a thin outer capillary containing $ZnCl_2$ in propylene carbonate and sealed with a thin silicone membrane, sulfide insensitive H_2 microsensors can be constructed (Nielsen et al., 2015). We did not repeat tests of CO, DMS, and methyl mercaptan interference on the sulfide insensitive H_2 sensors, but such interference will not be larger than on the unshielded sensors.

The new H_2S insensitive H_2 microsensor (see Nielsen et al., 2015 for details on construction and sensor design) exhibits a linear response from 0 to 100% H_2. By varying the thickness and diameter of the silicone membrane sealing the microsensor tip as well as the distance from the membrane to the internal measuring anode, we could manufacture H_2 microsensors with various measuring characteristics. Sensors without a H_2S shield could be constructed with a very fast response time of <0.2 s, but they also exhibited a relatively large stirring sensitivity, which can cause severe measuring artifacts, especially when measuring concentration gradients within a gradient of flow, e.g., in the diffusive boundary layer above sediments and biofilms (Revsbech, 1989; Klimant et al., 1995). Hydrogen gas microsensors with a lower stirring sensitivity can be constructed by using smaller silicone membrane diameters and a longer internal diffusion path, and here the presence of a H_2S shield contributes to the latter. The sulfide insensitive H_2 microsensors used in this study exhibited a sensitivity of 1.5–5 pA μM^{-1} H_2, with a negligible stirring sensitivity and a 90% response time of $\sim$20–40 s. In all cases, the new H_2 microsensors exhibited low and stable zero currents (1–10 pA) and a temperature sensitivity similar to other amperometric microsensors, i.e., an increase in sensor signal of 2–3% per °C.

We calibrated the sensor in salt water flushed with various defined amounts of H_2, either by help of a gas-mixing unit or by using commercially available defined mixtures of H_2 and N_2. Hydrogen data were either expressed as partial pressures (%H_2 saturation) or in molar concentration units. The H_2 concentration in saturated water at experimental salinity and temperature was calculated from tabulated values of H_2 solubility according to Wiesenburg and Guinasso (1979).

The amperometric microsensors were used in connection with a pA-meter (PA2000 and Microsensor Multimeter, Unisense A/S, Denmark), while the pH microsensors were used with a standard calomel reference electrode both connected to a high impedance mV meter (Microsensor Multimeter, Unisense A/S, Denmark). Measuring signals were either recorded on a stripchart recorder (BD-25, Kipp&Zonen, Netherlands) or via an A/D converter (Unisense A/S, Denmark) connected to a PC. Microsensors were mounted in a motorized micromanipulator that was mounted on a heavy stand and was remotely controlled by a PC-interfaced motor controller (Unisense A/S, Denmark). Automated data acquisition and positioning of microsensors was done with commercial software (Profix and Sensor TracePro, Unisense A/S, Denmark). The microsensors were inserted into the biofilm vertically from above in defined steps of 100–200 μm.

The efflux of H_2 from the microbial mats, $J(H_2)$ quantified net H_2 production and was calculated from measured concentration gradients using Fick's firs law, $J(H_2) = D^*$ (dC/dz), where D is the molecular diffusion coefficient of H_2 at experimental temperature and salinity and dC/dz is the linear H_2 concentration gradient in the diffusive boundary layer above the mat. Similar flux calculations, $J(O_2)$ were done with O_2 concentration profiles to quantify net photosynthesis in the light and dark O_2 uptake rates, using the molecular diffusion coefficient of O_2 at experimental temperature and salinity. Diffusion coefficients were taken from Broecker and Peng (1974) and corrected for temperature and salinity according to Li and Gregory (1974): $D(O_2) = 2.05 \cdot 10^{-5}$ cm^2 s^{-1} and $D(H_2) = 3.93 \cdot 10^{-5}$ cm^2 s^{-1} at 21°C and 25 ppt; $D(O_2) = 2.04 \cdot 10^{-5}$ cm^2 s^{-1} and $D(H_2) = 3.90 \cdot 10^{-5}$ cm^2 s^{-1} at 25°C and 90 ppt.

Microsensor Measurements of Scalar Irradiance

Light penetration in the coastal microbial mat was measured with a scalar irradiance microsensor (Lassen et al., 1992; Kühl et al., 1997; Kühl, 2005) connected to a sensitive fiber-optic spectrometer (QE65000, Ocean Optics, USA) that was interfaced to a PC running dedicated spectral acquisition software (Spectrasuite, Ocean Optics, USA). Mat samples were illuminated vertically from above with a fiber-optic halogen lamp equipped with a collimating lens (KL-2500, Schott, Germany), where the downwelling photon irradiance was regulated with a built-in neutral density screen to 500 μmol photons m^{-2} s^{-1}. A scalar irradiance microsensor was mounted in a manually operated micromanipulator (MM33, Märtzhäuser GmbH, Germany) and inserted into the mat at a 45° angle relative to the vertically incident light. Measurements were corrected for the measuring angle, and depths are given as vertical depth below the mat surface. Data were normalized to the incident downwelling irradiance as measured with the scalar irradiance microsensor positioned in the light path at similar distance as the mat surface but over a black light absorbing well.

Results and Discussion

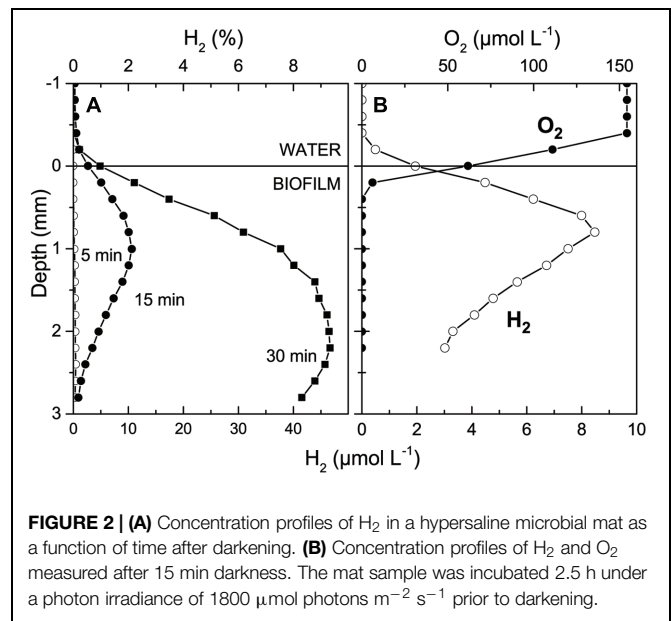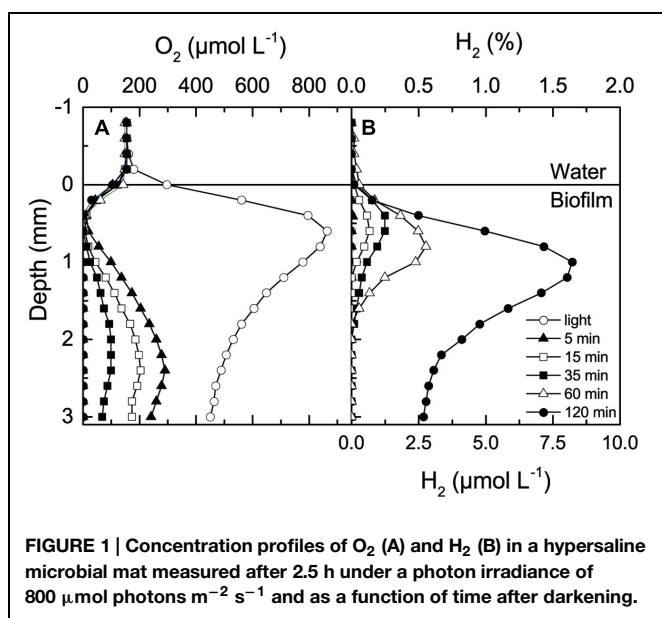We measured H_2 dynamics in two different cyanobacterial mats: (i) a hypersaline mat with a pronounced layer of oxidized iron buffering the cyanobacterial top layer against sulfide exposure (Wieland et al., 2005), and (ii) a highly sulfidic coastal cyanobacterial mat (Lassen et al., 1992). Data from the hypersaline mat were measured with H_2 microsensors without a sulfide trap and in the absence of H_2S, as checked with a

H_2S microsensor (data not shown). Data in the highly sulfidic coastal mat were measured with H_2 microsensors equipped with a chemical sulfide trap in front of the measuring tip (Nielsen et al., 2015).

Hydrogen Production in the Hypersaline Cyanobacterial Mat

When incubated under an irradiance of 800 μmol photons m^{-2} s^{-1} for 2.5 h, intense photosynthesis in the dense 2–3 mm thick hypersaline cyanobacterial biofilm lead to hyperoxic conditions reaching 4–5 times air saturation in the upper mm and supersaturating O_2 levels throughout the whole sample, which was contained in a small glass container (**Figure 1A**). Upon darkening, O_2 was most rapidly depleted in the region showing highest O_2 production activity in light, and H_2 was first detected in this zone after 15 min. As O_2 became further depleted, H_2 accumulated to higher concentrations and over a wider zone in the biofilm reaching a maximum of 8 μmol H_2 L^{-1} (~1.6% H_2) at 1 mm depth after 2 h in the dark. Hydrogen was consumed in the lowermost parts of the biofilm sample, which was constrained by the bottom of the small glass incubation container. The apparent migration of the H_2 peak into slightly deeper layers probably reflects a shift in the relative balance between H_2 production, consumption and transport, especially as the mat sample was confined in a small glass vial presenting a diffusion barrier ~4 mm below the mat surface.

We note that the absolute amount of H_2 produced in the mats after darkening apparently depended on the irradiance level during the previous light incubation. When we increased the irradiance to 1800 μmol photons m^{-2} s^{-1} for 2.5 h we again saw a very strong O_2 accumulation in the sample but observed a much higher H_2 production reaching ~2% H_2 15 min after onset of darkness. Maximal levels of 40–50 μM (8–9% H_2) were reached in the upper millimeters of the mat within 30 min after darkening (**Figure 2A**). These H_2 levels are higher than most other findings

FIGURE 2 | (A) Concentration profiles of H_2 in a hypersaline microbial mat as a function of time after darkening. **(B)** Concentration profiles of H_2 and O_2 measured after 15 min darkness. The mat sample was incubated 2.5 under a photon irradiance of 1800 μmol photons m^{-2} s^{-1} prior to darkening.

FIGURE 1 | Concentration profiles of O_2 (A) and H_2 (B) in a hypersaline microbial mat measured after 2.5 h under a photon irradiance of 800 μmol photons m^{-2} s^{-1} and as a function of time after darkening.

in more permanently submerged hypersaline mats that generally exhibit lower H_2 accumulation than intertidal mats (**Table 1**).
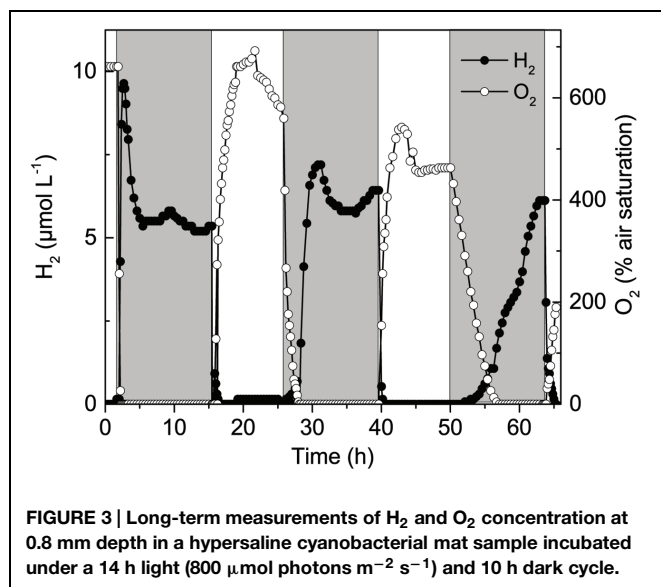
We speculate that the apparent enhancement in H_2 production with irradiance reflects a higher accumulation of storage products in the cyanobacteria enhancing subsequent dark fermentation. A similar explanation was proposed by Hoffmann et al. (2015), who measured sustained H_2 production during darkness in intertidal microbial mats kept in a greenhouse for 1.5 years, and with an apparent positive correlation between the solar radiative flux during daytime and the night-time H_2 production in the mats. A rigorous test of this hypothesis would require measurements of cyanobacterial photosynthate accumulation as well as H_2 and fermentation products as a function of light incubation time and may be complicated by the cross-feeding of cyanobacterial fermentation products to other mat members such as SRB and Chloroflexi (Burow et al., 2014; Lee et al., 2014). Measurements on cyanobacterial cultures using methodology described by Kothari et al. (2014) may thus be more straightforward.

Alignment of H_2 and O_2 microsensor measurements showed that H_2 diffused out of the mat in the dark (**Figure 2B**) with an estimated maximal H_2 efflux of ~23.8 nmol H_2 cm^{-2} h^{-1} amounting to about 13% of the diffusive O_2 uptake of 182.2 nmol O_2 cm^{-2} h^{-1} in the dark and 2% of the net photosynthetic O_2 production in light of 1420 nmol O_2 cm^{-2} h^{-1}. Hoffmann et al. (2015) did similar measurements in different intertidal mats and found that H_2 production in the dark amounted to 0.2–5% of net photosynthesis and 0.4–28% of O_2 respiration.

Long-term experiments with simultaneous O_2 and H_2 measurements over several days (14 h dark: 10 h light) with the microsensor tips positioned 0.8 mm below the mat surface, i.e., in the zone of maximal O_2 production in the light, showed recurrent H_2 production that persisted in the anoxic mat throughout the 14 h dark incubation period (**Figure 3**). Quantification of H_2 production in *Lyngbya aestuarii* and

TABLE 1 | Comparison of maximal concentrations and fluxes of H_2 reported in microbial mats. Listed in chronological order.

Sample origin and mat type	Salinity (ppt)	Maximal H_2 concentration (μmol L^{-1})		Maximal H_2 production (nmol H_2 cm^{-2} h^{-1})		Source
		−molybdate	+molybdate	−molybdate	+ molybdate	
Hamelin Pool, Australia						Skyring et al. (1989), estimated from Figure 1
Microcoleus mat (smooth)	50–70	–	–	3-	17.9	
Lyngbya mat (tufted)	50–70	–	–	<0.1	0.58	
Entophysalis mat (pustular)	50–70	–	–	5.8	4.8	
Spencer Gulf, Australia						Skyring et al. (1989), Table 1
Microcoleus mat	45	–	–	0.125	36.3	
Guerrero Negro, Mexico						Hoehler et al. (2001), Figures 1 and 2, Table 1
Lyngbya mat (intertidal)	95	50	–	62.9	–	
Microcoleus mat (subtidal)	95	<0.5	–	0.35	–	
Elkhorn Slough, USA						Burow et al. (2012), estimated from Supplementary Figure S1
Microcoleus mat	35	–	–	49.5	–	
Guerrero Negro, Mexico						Hoffmann et al. (2015), estimated from Figures 3 and 4B
Leptolyngbya / Microcoleus mat (low tide)	40	∼5	∼20	11	–	
Microcoleus mat (mid tide)	40	∼15	–	6.8	–	
Microcoleus / Calothrix mat (high tide)	40	∼20	–	0.3	–	
Aggersund, Denmark						This study
Microcoleus mat	25	25–40	70	16.3	86.5	
Saline des Giraud, France						This study
Microcoleus mat	90	47	–	23.8	–	

FIGURE 3 | Long-term measurements of H_2 and O_2 concentration at 0.8 mm depth in a hypersaline cyanobacterial mat sample incubated under a 14 h light (800 μmol photons m^{-2} s^{-1}) and 10 h dark cycle.

M. chtonoplastes cultures showed similar long term persistence for >24 h (Kothari et al., 2014). The build-up of H_2 was highest immediately after darkening and then leveled off after 1–2 h. Strong H_2 depletion occurred rapidly after onset of the illumination leading to O_2 accumulation from photosynthesis. Such depletion can be explained by several mechanisms such as (i) O_2 inhibition of H_2 production coupled with diffusive losses, (ii) H_2 consumption with O_2, by e.g., Knallgas bacteria, and/or intermittent anoxygenic photosynthesis. However, our limited experimental data do not allow us to discriminate between the relative importance of these H_2 consuming processes.

In the second light period, the apparently constant O_2 level at this particular depth over many hours indicated formation of a gas bubble in the mat that acted as an O_2 and H_2 reservoir slowing the O_2 depletion and the build-up of H_2 after the subsequent light-dark shift. However, the slower O_2 and H_2 dynamic in subsequent light–dark shifts could also indicate an increasing substrate limitation as the light period may not have been sufficient to restock cyanobacterial storage products.

Overall, the observed H_2 dynamics in the hypersaline mat is very similar to patterns recently observed in cyanobacterial cultures (Kothari et al., 2014) and intertidal microbial mats (Hoffmann et al., 2015). In comparison to other studies of H_2 in hypersaline mats (**Table 1**), we found much higher H_2 levels after darkening. The reasons for such high H_2 accumulation remain to be studied in detail, but we speculate that the high content of oxidized iron in the upper layers of the Saline des Giraud mat

(Wieland et al., 2005) may lead to less sulfate reduction and thus less consumption of H_2, in contrast to most other hypersaline mat systems that often become highly sulfidic in darkness due to intense sulfate reduction in the top layers. However, the higher H_2 levels may also simply reflect that the small glass incubation vial confined the sample to only a 3–4 mm thick top layer and thus did not allow a diffusive exchange and consumption in deeper more sulfidic mat layers. There is thus clearly a need for more detailed H_2 and H_2S measurements on deeper mat cores from Saline des Giraud.

Hydrogen Dynamics in Coastal Cyanobacterial Mats

More detailed microenvironmental analyses of H_2 dynamics were done in a sulfidic coastal cyanobacterial mat (Supplementary Figure S1A), whereas measurements in coastal sediment with a surface biofilm of diatoms (Supplementary Figure S1B) showed no accumulation of H_2 (data not shown). Spectral scalar irradiance measurements showed strong light attenuation with depth in the dense 1–2 mm thick top layer of the coastal mat, where distinct throughs in the transmission spectra indicated a high density of cyanobacteria with Chl a and phycobilins, as well as anoxygenic phototrophs with Bchl a and Bchl c (Supplementary Figure S2), with the latter being indicative of the presence of Chloroflexi. The euphotic zone for oxygenic photosynthesis was limited to the uppermost mm of the mat, wherein visible light (PAR, 400–700 nm) was attenuated to ~ 0.1 μmol photons m^{-2} s^{-1} (**Figure 4A**). Similar optical characteristics were found in samples from the same site by Lassen et al. (1992) >20 years ago.
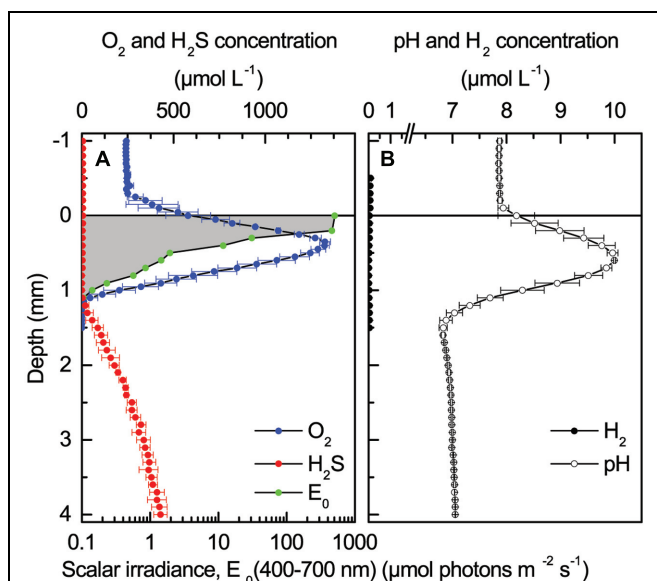

FIGURE 4 | Light and chemical gradients in a coastal cyanobacterial mat under an incident photon irradiance of 500 μmol photons m^{-2} s^{-1}. (A) Depth profiles of photon scalar irradiance, O_2 and H_2S concentrations. **(B)** Depth profiles of pH and H_2 concentration. Symbols with error bars for chemical parameters represent the mean ± SD ($n = 3$).

Chemical Microenvironment in Light

The chemical conditions in the mat exhibited steep concentration gradients (**Figure 4**). Under high irradiance of PAR (500 μmol photons m^{-2} s^{-1}), intense photosynthesis led to peak O_2 concentrations of ~ 4.5 times air-saturation 0.5 mm below the mat surface. However, O_2 only penetrated to ~ 1.2 mm in the light due to intense respiration and re-oxidation of reduced chemical species. Sulfide was produced by SRB in deeper mat layers, where H_2S levels reached about 0.5 mM at 3 mm depth. Sulfide was re-oxidized by O_2 in a thin zone around 1.2-1.4 mm depth. Strong photosynthesis caused a strong pH increase in the photic zone peaking at pH 10 around 0.5–0.7 mm below the mat surface, i.e., >2 pH units above the overlaying water pH of 7.9. With increasing depth, pH dropped by ~ 3 units reaching pH <7 in the sulfide oxidation zone before stabilizing around pH 7 in deeper mat layers. No H_2 was detected in the upper millimeters of the illuminated mat. Measurements in another mat sample from the same habitat showed the same chemical zonations and extremes, albeit with a somewhat deeper O_2 penetration depth in light of ~ 1.5 mm and a more heterogeneous distribution of H_2S in deeper mat layers (Supplementary Figure S3).

Hydrogen Gas Production and Chemical Dynamics after Darkening

The chemical conditions in the coastal mat changed dramatically after darkening (**Figure 5**). Within 5 min after darkening, O_2 became strongly depleted and only penetrated ~ 0.3 mm into the mat, with a further decrease in the O_2 penetration depth to 0.2 mm over the following 85 min. The O_2 and H_2S concentration profiles were initially separated by a ~ 0.7 mm wide zone, wherein H_2 accumulated rapidly after darkening. Peak concentrations were measured 0.5–0.7 mm below the mat surface increasing from ~ 13 μM H_2 after 5 min to ~ 22 μM H_2 after 45 min dark incubation. Over this time interval, pH in the H_2 production zone decreased to a stable value of pH 7.2–7.8. Produced H_2 diffused both toward the mat surface and toward the sulfidic zone, where it was consumed. After 45 min, H_2 levels in the mat started to decrease, while sulfide levels continued to increase in the upper mat layers. Sulfide started to overlap with the O_2 concentration profile after 90 min. Hydrogen levels in the mat continued to decrease slowly and in a second experiment complete H_2 depletion was only found after about 7 h dark incubation as seen in **Figure 6**, where data from continuous measurements at 0.6 mm depth are shown. The observed pattern of rapid build-up followed by a slow decline in H_2 concentration follows a similar pattern observed in studies of H_2 evolution in cyanobacterial isolates (Kothari et al., 2012).

Hydrogen Accumulation in the Presence of Molybdate

Additional measurements of H_2 production after onset of darkness were done in another coastal mat sample from the same habitat (**Figure 7**) that was incubated 8 h under a photon irradiance of 500 μmol photons m^{-2} s^{-1} prior to darkening. When incubated in normal seawater, the mat reached maximal concentrations of ~ 18 μM H_2 at 0.5–0.7 mm depth within
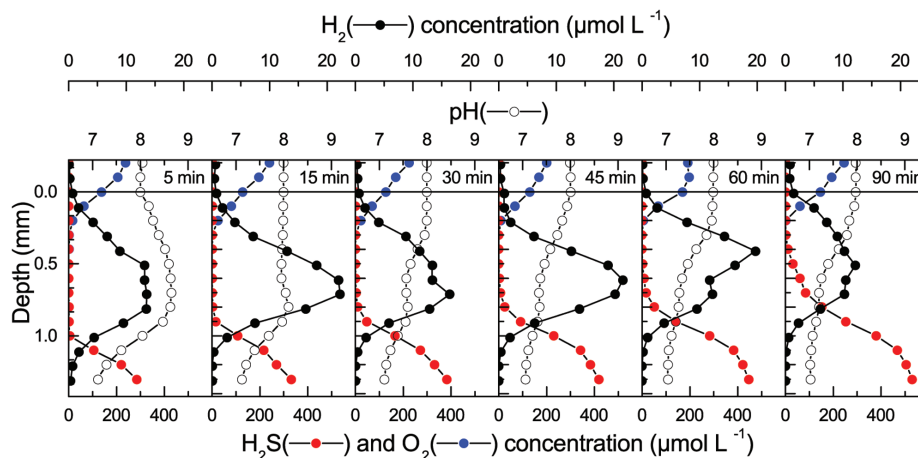
FIGURE 5 | Chemical gradients in a coastal cyanobacterial mat (same sample as in Figure 4) as a function of time after darkening. Prior to darkening, the mat was incubated in aerated seawater under a photon irradiance of 500 μmol photons m^{-2} s^{-1}.
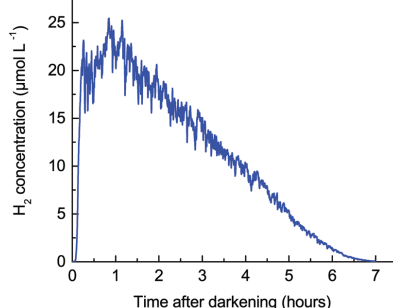

FIGURE 6 | Accumulation and depletion of H$_2$ in a coastal cyanobacterial mat (same sample as in Figure 4) as a function of time after darkening. Measurements were performed 0.6 mm below the mat surface in the zone of maximal H$_2$ production (see **Figure 5**).
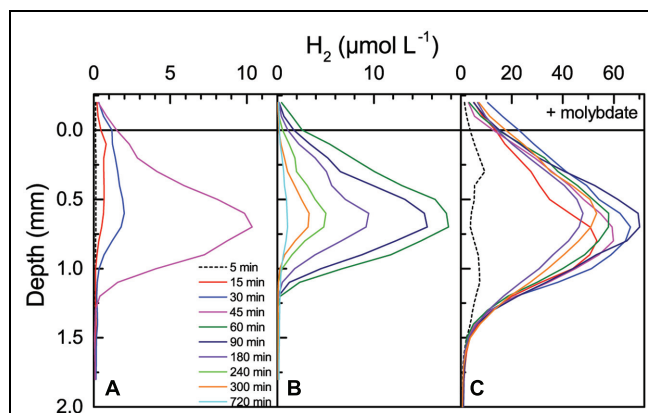

FIGURE 7 | Accumulation and depletion of H$_2$ in a coastal cyanobacterial mat as a function of time after darkening (see numbers and color legends in the graph). Prior to darkening, the mat was incubated under a photon irradiance of 500 μmol photons m^{-2} s^{-1} in aerated seawater for 8 h without **(A,B)** and for 6 h with 2.5 mM molybdate **(C)**. Note the use of different concentration ranges in the different panels.

60 min after darkening, where after H$_2$ levels in the mat declined gradually to ~1.5 μM H$_2$ after 720 min in darkness (**Figures 7A,B**). Thereafter, the same mat was incubated 6 h under a photon irradiance of 500 μmol photons m^{-2} s^{-1} in seawater with 2.5 mM molybdate, an inhibitor of sulfate reduction. Measurements in light at the end of this incubation showed a similar O$_2$ and pH distribution, whereas H$_2$S levels in the mat were much lower below the photic zone than in the absence of molybdate (Supplementary Figures S3 and S5); interestingly, a slight accumulation of H$_2$ (reaching 1–1.5 μM) was also observed in deeper mat layers around 3 mm depth after the molybdate treatment, i.e., in the aphotic zone that exhibited high H$_2$S levels in the absence of molybdate.

In presence of molybdate, the H$_2$ accumulation in the mat after darkening was much stronger and H$_2$ penetrated deeper into the mat (**Figure 7C**). Within 30 min after darkening, H$_2$ concentrations reached peak values of >60 μM H$_2$ 0.6–0.8 mm below the mat surface. The H$_2$ concentrations remained high in the dark incubated mat for about 5 h and showed

much slower H$_2$ depletion than in the absence of molybdate. Without molybdate, the produced H$_2$ penetrated to a depth of 1.1–1.3 mm where it became fully depleted in a relatively narrow zone (**Figures 7A,B**). In the presence of molybdate, H$_2$ penetrated to a depth of almost 2 mm and the H$_2$ concentration profile showed a much more gradual depletion with depth (**Figure 7B**). Both with and without molybdate, there was no indication of strong H$_2$ depletion in the uppermost mat layer and the microprofiles showed an efflux of H$_2$ into the overlaying water. The H$_2$ efflux was strongly stimulated in the presence of molybdate (**Table 1**). In the absence of molybdate, the maximal H$_2$ efflux 60 min after onset of darkness reached 16.3 nmol H$_2$ cm^{-2} h^{-1} amounting to 5.5% of the dark O$_2$ consumption (296.2 nmol O$_2$ cm^{-2} h^{-1}) and 1% of the net photosynthetic O$_2$ production prior to darkening (1500 nmol O$_2$ cm^{-2} h^{-1}). In the

presence of molybdate, the maximal H_2 efflux reached 86.5 nmol H_2 cm^{-2} h^{-1} amounting to 29% of the dark O_2 consumption and 6% of the net photosynthesis.
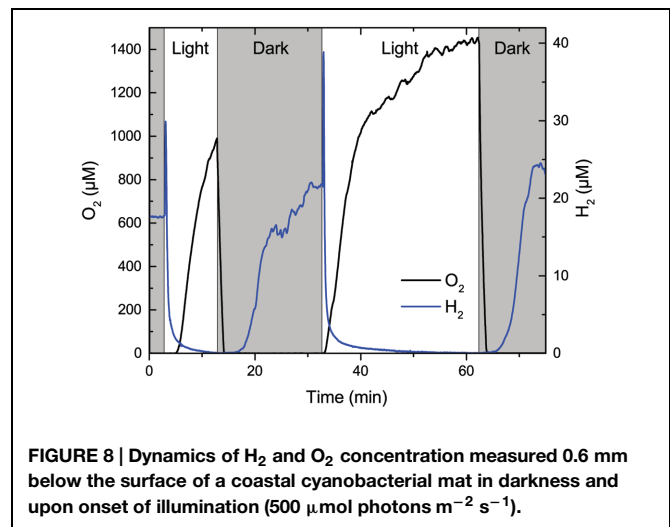
Hoffmann et al. (2015) did a similar experiment in intertidal microbial mats showing stimulated H_2 production reaching 4 times higher peak concentrations of H_2 in the presence of molybdate reaching up to ∼20 μM H_2 in the upper mm of the mat (**Table 1**). However, their measurements with a sulfide sensitive H_2 microsensor showed a H_2 concentration peak on top of an apparent gradually increasing H_2 concentration with depth, which can be interpreted as an sulfide interference on the microsensor signal (see below).

Bulk measurements of the H_2 production of coastal and hypersaline mats (Skyring et al., 1989) and in the upper 2 mm of a coastal cyanobacterial mat (Burow et al., 2014) and two hypersaline cyanobacterial mats (Lee et al., 2014) all showed a strong stimulation of H_2 production upon inhibition of sulfate reduction activity by addition of molybdate to the incubation vials. Molecular analyses of the microbial diversity and gene expression in such mats identified SRB as major hydrogenotrophs in the mat along with filamentous anoxygenic phototrophs belonging to the Chloroflexi (Burow et al., 2014; Lee et al., 2014). In the present study, we did not investigate the distribution and identity of SRB or Chloroflexi or their hydrogenase gene expression in the mat samples, but our H_2 microsensor data strongly support the findings in other cyanobacterial mats identifying SRB as primary hydrogenotrophs in the upper mat layers.

A close spatial co-occurrence of SRB and cyanobacteria has been demonstrated in the upper millimeters of several microbial mat environments (e.g., Baumgartner et al., 2006; Fike et al., 2008), including observation of migratory behavior of motile SRB toward the photic zone (Krekeler et al., 1998). This has been ascribed to aerobic sulfate reduction (Canfield and Des Marais, 1991) and/or a possible aerotaxis combined with aggregation and high O_2 respiration as a survival mechanism for SRB in the photic zone (Cypionka, 2000; Baumgartner et al., 2006). But the presence of SRB within the photic zone also reflects the easy access of SRB to both electron donor and acceptor immediately after onset of darkness and onset of fermentative H_2 production (Lee et al., 2014), and we note that some SRB can also catalyze the oxidation of H_2, organic substrates and inorganic sulfur species with O_2 as an electron acceptor (Dannenberg et al., 1992). Metabolic flexibility including a versatile H_2 metabolism and chemotaxis of SRB may thus enable them to thrive in the highly variable chemical microenvironment of the photic zone in microbial mats.

Photo-Stimulation of H_2 Production

Simultaneous measurements of O_2 and H_2 concentration in the coastal mat at a depth of 0.6 mm, i.e., within the zone of maximal H_2 production in the dark and photosynthetic O_2 production in the light, showed pronounced dynamics (**Figure 8**). In the dark, O_2 was depleted completely and H_2 concentrations reached levels of 20–25 μM H_2 after 15–20 min. However, immediately after onset of illumination we observed a burst in H_2 production driving local concentrations up to 30–40 μM H_2. This burst

FIGURE 8 | Dynamics of H_2 and O_2 concentration measured 0.6 mm below the surface of a coastal cyanobacterial mat in darkness and upon onset of illumination (500 μmol photons m^{-2} s^{-1}).

only lasted for 20–30 s, where after H_2 became rapidly depleted as O_2 from photosynthesis accumulated to super saturating concentration levels in the mat. Similar measurements (data not shown) in depth horizons closer to the mat surface showed a shorter and less intense burst due to more rapid O_2 accumulation and thus faster H_2 depletion, whereas measurement in deeper mat layers showed a less intense build-up of H_2 upon onset of illumination due to strong light limitation; at 0.8 mm depth we observed no photo stimulation of H_2 production. Such intermittent pulses of H_2 upon illumination have been ascribed to direct biophotolysis in cyanobacteria involving a bidirectional Ni–Fe hydrogenase (Appel et al., 2000). While our data give first evidence that such biophotolysis can occur in the uppermost parts of cyanobacterial mats, the process is limited to <1 min after onset of illumination and thus plays a very minor role for the total H_2 production.

Sulfide Interference on H_2 Microsensor Measurements

Our measurements in the hypersaline mat were done under absence of sulfide (checked by H_2S microsensor measurements) due to a pronounced layer of oxidized iron buffering against accumulation of free sulfide in the photic zone during darkness (Wieland et al., 2005). Under such conditions, the commercially available H_2 microsensor from Unisense performs well and gives accurate quantifications of H_2 concentrations. However, this sensor is also sensitive to hydrogen sulfide giving rise to about 20–30% of the signal for a given H_2S concentration as compared to the same H_2 concentration, and exposure to high H_2S levels can also affect sensor calibration (cf. H_2 microsensor manual available at: www.unisense.com/manuals/). Typical H_2S concentrations in the upper millimeters of hypersaline and coastal mats can reach up to 500 μM just below the photic zone in light or in the uppermost mat layers in darkness (Wieland and Kühl, 2000). A strong gradient of increasing H_2S concentration with depth would thus be detected in H_2 microsensor measurements giving rise to false H_2 signals of up to >50 μM and typically showing a continuously increasing H_2 level with depth in the mat

following the increasing H_2S concentration. Hoffmann et al. (2015) measured with the sulfide-sensitive H_2 microsensor in three types of intertidal mats. In the upper tidal apparently less sulfidic mat they found a clear peak of H_2 production overlapping with the photic zone, while measurements in the other mats showed that such a H_2 production peak was overlayed by a strong continuously increasing H_2 concentration with depth (cf. Figure 3 in Hoffmann et al., 2015). While the authors did not report on H_2S measurements in their mats, we suggest that their measurements in deeper mat layers may include H_2S interference. We have observed similar patterns when using sulfide sensitive H_2 microsensors in the coastal mat from Aggersund (data not shown). Such interference would also affect subsequent rate calculations on measured H_2 concentration profiles, especially in deeper zones. With the new sulfide-insensitive microsensor it is now possible to measure in strongly sulfidic microbial mats without such potential artifacts.

Conclusion

Our measurements demonstrated distinct microscale dynamics of H_2 in hypersaline and coastal microbial mats that are densely populated by filamentous cyanobacteria. We found a pronounced build-up of H_2 in the upper millimeters of such mats upon darkening, while more oxidized coastal sediments with a surface biofilm of diatoms did not show any H_2 accumulation. In general, our results support other recent demonstrations of strong H_2 production in mat-forming cyanobacteria (Kothari et al., 2012, 2014) and intact microbial mats (Hoehler et al., 2001; Burow et al., 2012; Hoffmann et al., 2015), where cyanobacterial fermentation of photosynthate in darkness is the major H_2 source. Such H_2 formation upon light–dark shifts in the upper photic zone thus seems inherent in many coastal and hypersaline microbial mats and presents an important, yet intermittent energy source that together with photosynthate fermentation products can fuel anaerobic respiration processes such as sulfate reduction (Lee et al., 2014).

In conclusion, this first application of a new H_2 microsensor (Nielsen et al., 2015) in concert with microsensors for O_2, pH, H_2S, and scalar irradiance demonstrated a pronounced potential for H_2 production in the photic zone of microbial mats. Strong intermittent H_2 accumulation (up to 30–40 μM H_2) and efflux of H_2 to the overlaying water originated in the uppermost cyanobacterial layers, with the most intense H_2 formation in the depth horizons exhibiting maximal photosynthesis in the light. In the dark, H_2 first accumulates in the uppermost mm of the mat and is then released to the overlaying water and

consumed over 6–7 h by anaerobic respiration. Depletion of H_2 within the mat was strongly inhibited by molybdate addition pointing to SRB as major hydrogenotrophs in the mat. Strong H_2 production by biophotolysis was observed in the uppermost anoxic mat layers (0.2–0.8 mm) immediately (for <1 min) after onset of illumination but was quickly inhibited by oxygenic photosynthesis and did not contribute significantly to the H_2 production, which was primarily observed in the dark. The new microsensors allow detailed studies of H_2 dynamics in sulfidic environments at high spatio-temporal resolution. Such studies have until recently been limited to mm-scale measurements of net H_2 production from incubated mat samples using gas chromatography or other bulk phase measurements. Here we have focused on costal and hypersaline cyanobacterial mats, but the new H_2 sensor is also suitable for measurements at higher temperatures (up to 60°C; Nielsen et al., 2015) and we are currently investigating H_2 dynamics in hot spring microbial mats, where H_2 metabolism often plays a major role (Spear et al., 2005) although the relative importance of H_2 and H_2S as an energy source remains debated (D'Imperio et al., 2008).

Author Contributions

Planned and designed experiments (MN, NR, MK). Performed experiments and analyzed data (MN, NR, MK). Wrote the article (MK with editorial help by MN and NR).

Acknowledgments

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.00726

References

Appel, J., Phunruch, S., Steinmüller, K., and Schulz, R. (2000). The bidirectional hydrogenase of *Synechocystis* sp. *PCC* 6803 works as an electron valve during photosynthesis. *Arch. Microbiol.* 173, 333–338. doi: 10.1007/s002030000139

Baumgartner, L. K., Reid, R. P., Dupraz, C., Decho, A. W., Buckley, D. H., Spear, J. R., et al. (2006). Sulfate reducing bacteria in microbial

mats: changing paradigms, new discoveries. *Sed. Geol.* 185, 131–145. doi: 10.1016/j.sedgeo.2005.12.008

Benemann, J. R., and Weare, N. M. (1974). Hydrogen evolution by nitrogen-fixing *Anabaena cylindrica* cultures. *Science* 184, 174–175. doi: 10.1126/science.184.4133.174

Berg, H. C. (1983). *Random Walks in Biology*. Princeton, NJ: Princeton University Press.

Bothe, H., Schmitz, O., Yates, M. G., and Newton, W. E. (2010). Nitrogen fixation and hydrogen metabolism in cyanobacteria. *Microbiol. Mol. Biol. Rev.* 74, 529–551. doi: 10.1128/MMBR.00033-10

Broecker, W. S., and Peng, T. H. (1974). Gas exchange rates between air and sea. *Tellus* 26, 21–35. doi: 10.1111/j.2153-3490.1974.tb01948.x

Burow, L. C., Woebken, D., Bebout, B. M., McMurdie, P. J., Singer, S. W., Pett-Ridge, J., et al. (2012). Hydrogen production in photosynthetic microbial mats in the Elkhorn Slough estuary, Monterey Bay. *ISME J.* 6, 863–874. doi: 10.1038/ismej.2011.142

Burow, L. C., Woebken, D., Marshall, I. P. G., Singer, S. W., Pett-Ridge, J., Prufert-Bebout, L., et al. (2014). Identification of desulfobacterales as primary hydrogenotrophs in a complex microbial mat community. *Geobiology* 12, 221–230. doi: 10.1111/gbi.12080

Canfield, D. E., and Des Marais, D. J. (1991). Aerobic sulfate reduction in microbial mats. *Science* 251, 1471–1473. doi: 10.1126/science.11538266

Conrad, R. (1988). Biogeochemistry and ecophysiology of atmospheric CO and H_2. *Adv. Microb. Ecol.* 10, 231–283. doi: 10.1007/978-1-4684-5409-3_7

Cypionka, H. (2000). Oxygen respiration by *Desulfovibrio* species. *Annu. Rev. Microbiol.* 54, 827–848. doi: 10.1146/annurev.micro.54.1.827

Dannenberg, S., Kroder, M., Dilling, W., and Cypionka, H. (1992). Oxidation of H_2, organic compounds and inorganic sulfur compounds coupled to reduction of O_2 or nitrate by sulfate-reducing bacteria. *Arch. Microbiol.* 158, 93–99. doi: 10.1007/BF00245211

D'Imperio, S., Lehr, C. R., Oduro, H., Druschel, G., Kühl, M., and McDermott, T. R. (2008). The relative importance of H_2 and H_2S as energy sources for primary production in geothermal springs. *Appl. Environ. Microbiol.* 74, 5802–5808. doi: 10.1128/AEM.00852-08

Ebert, A., and Brune, A. (1997). Hydrogen concentration profiles at the oxic-anoxic interface: a microsensor study of the hindgut of the wood-feeding lower termite *Reticulitermes flavipes* (Kollar). *Appl. Environ. Microbiol.* 63, 4039–4046.

Fike, D. A., Gammon, C. L., Ziebis, W., and Orphan, V. J. (2008). Micron-scale mapping of sulfur cycling across the oxycline of a cyanobacterial mat: a paired nanoSIMS and CARD-FISH approach. *ISME J.* 2, 749–759. doi: 10.1038/ismej.2008.39

Fourcans, A., Garcia de Oteyza, T., Wieland, A., Sole, A., Diestra, E., van Bleijswijk, J., et al. (2004). Characterization of functional bacterial groups in a hypersaline microbial mat community (Salins-de-Giraud, Camargue, France). *FEMS Microbiol. Ecol.* 51, 55–70. doi: 10.1016/j.femsec.2004.07.012

Gest, H., and Kamen, M. D. (1949). Photoproduction of molecular hydrogen by *Rhodospirillum rubrum*. *Science* 109, 558–559. doi: 10.1126/science.109.2840.558

Hallenbeck, P.C. (2012). "Hydrogen production by cyanobacteria," in *Microbial Techniques in Advanced Biofuels Production*, ed. P. C. Hallenbeck (Berlin: Springer).

Hoehler, T., Albert, D. B., Alperin, M. J., Bebout, B. M., Martens, C. S., and Des Marais, D. J. (2002). Comparative ecology of H_2 cycling in sedimentary and phototrophic ecosystems. *Antonie Van Leeuwenhoek* 81, 575–585. doi: 10.1023/A:1020517924466

Hoehler, T., Bebout, B. M., and Des Marais, D. J. (2001). The role of microbial mats in the production of reduced gases on the early Earth. *Nature* 412, 324–327. doi: 10.1038/35085554

Hoffmann, D., Maldonado, J., Wojciechowski, M. F., and Garcia-Pichel, F. (2015). Hydrogen export from intertidal cyanobacterial mats: sources, fluxes, and the influence of community composition. *Environ. Microbiol.* doi: 10.1111/1462-2920.12769 [Epub ahead of print].

Houchins, J. P. (1984). The physiology and biochemistry of hydrogen metabolism in cyanobacteria. *Biochim. Biophys. Acta* 768, 227–255. doi: 10.1016/0304-4173(84)90018-1

Hübert, T., Boon-Brett, L., Black, G., and Banach, U. (2011). Hydrogen sensors – a review. *Sens. Act. B* 157, 329–352. doi: 10.1016/j.snb.2011.04.070

Jackson, D. D., and Ellms, J. W. (1896). On odors and tastes of surface waters with special reference to *Anabaena*, a microscopical organism found in certain water supplies of Massachusetts. *Rep. Mass. State Board Health* 1896:410–420.

Jørgensen, B. B. (2001). Space for hydrogen. *Nature* 412, 286–289. doi: 10.1038/35085676

Klimant, I., Meyer, V., and Kühl, M. (1995). Fiber-optic oxygen microsensors, a new tool in aquatic biology. *Limnol. Oceanogr.* 40, 1159–1165. doi: 10.4319/lo.1995.40.6.1159

Kothari, A., Parameswaran, P., and Garcia-Pichel, F. (2014). Powerful fermentative hydrogen evolution of photosynthate in the cyanobacterium *Lyngbya aestuarii* BLJ mediated by a bidirectional hydrogenase. *Front. Microbiol.* 5:680. doi: 10.3389/fmicb.2014.00680

Kothari, A., Potrofka, R., and Garcia-Pichel, F. (2012). Diversity in hydrogen evolution from bidirectional hydrogenases in cyanobacteria from terrestrial, freshwater and marine intertidal environments. *J. Biotechnol.* 162, 105–114. doi: 10.1016/j.jbiotec.2012.04.017

Krekeler, D., Teske, A., and Cypionka, H. (1998). Strategies of sulfate-reducing bacteria to escape oxygen stress in a cyanobacterial mat. *FEMS Microbiol. Ecol.* 25, 89–96. doi: 10.1111/j.1574-6941.1998.tb00462.x

Kühl, M. (2005). Optical microsensors for analysis of microbial communities. *Methods Enzymol.* 397, 166–199. doi: 10.1016/S0076-6879(05)97010-9

Kühl, M., Glud, R. N., Ploug, H., and Ramsing, N. B. (1996). Microenvironmental control of photosynthesis and photosynthesis-coupled respiration in an epilithic cyanobacterial biofilm. *J. Phycol.* 32, 799–812. doi: 10.1111/j.0022-3646.1996.00799.x

Kühl, M., Lassen, C., and Revsbech, N. P. (1997). A simple light meter for measurements of PAR (400-700 nm) with fiber-optic microprobes: application for P vs. I measurements in microbenthic communities. *Aq. Microb. Ecol.* 13, 197–207. doi: 10.3354/ame013197

Kühl, M., and Revsbech, N. P. (2001). "Biogeochemical microsensors for boundary layer studies," in *The Benthic Boundary Layer*, eds B. P. Boudreau and B. B. Jørgensen (Oxford: Oxford University Press), 180–210.

Kühl, M., Steuckart, C., Eickert, G., and Jeroschewski, P. (1998). A H_2S microsensor for profiling sediments and biofilms: application in acidic sediment. *Aq. Microb. Ecol.* 15, 201–209. doi: 10.3354/ame015201

Lambert, G. R., and Smith, G. D. (1981). The hydrogen metabolism of cyanobacteria (blue-green algae). *Biol. Rev.* 56, 589–660. doi: 10.1111/j.1469-185X.1981.tb00360.x

Lassen, C., Ploug, H., and Jørgensen, B. B. (1992). Microalgal photosynthesis and spectral scalar irradiance in coastal marine sediments of Limfjorden, Denmark. *Limnol. Oceanogr.* 37, 760–772. doi: 10.4319/lo.1992.37.4.0760

Lee, H., Vermaas, W. F. J., and Rittman, B. E. (2010). Biological hydrogen production: prospects and challenges. *Trends Biotechnol.* 28, 262–271. doi: 10.1016/j.tibtech.2010.01.007

Lee, J. Z., Burow, L. C., Woebken, D., Everroad, R. C., Kubo, M. D., Spormann, A. M., et al. (2014). Fermentation couples Chloroflexi and sulfate-reducing bacteria to cyanobacteria in hypersaline microbial mats. *Front. Microbiol.* 5:61. doi: 10.3389/fmicb.2014.00061

Li, Y. H., and Gregory, S. (1974). Diffusion of ions in seawater and in deep-sea sediments. *Geochim. Cosmochim. Acta* 38, 703–714. doi: 10.1016/0016-7037(74)90145-8

Lorenzen, J., Glud, R. N., and Revsbech, N. P. (1995). Impact of microsensor caused changes in diffusive boundary layer thickness on O_2 profiles and photosynthetic rates in benthic communities of microorganisms. *Mar. Ecol. Prog. Ser.* 119, 237–241. doi: 10.3354/meps119237

Marschall, I. P. G., Berggren, D. R. V., Azizian, F., Burow, L., Semprini, L., and Spormann, A. M. (2012). The hydrogenase chip: a tiling oligonucleotide DNA microarray technique for characterizing hydrogen-producing and – consuming microbes in microbial communities. *ISME J.* 6, 814–826. doi: 10.1038/ismej.2011.136

Mason, S., Hamon, R., Nolan, A., Zhang, H., and Davison, W. (2005). Performance of a mixed binding layer for measuring anions and cations in a single assay using the diffusive gradients in thin films technique. *Anal. Chem.* 77, 6339–6346. doi: 10.1021/ac0507183

Moezelaar, R., Bijvank, S. M., and Stal, L. J. (1996). Fermentation and sulfur reduction in the mat-building cyanobacterium *Microcoleus chthonoplastes*. *Appl. Environ. Microbiol.* 62, 1752–1758.

Nielsen, M., Larsen, L. H., Ottosen, L. D. M., and Revsbech, N. P. (2015). Hydrogen microsensors with hydrogen sulfide traps. *Sens. Act. B* 215, 1–8. doi: 10.1016/j.snb.2015.03.035

Oschchepkov, V. P., Nikitina, A. A., Gusev, M. V., and Krasnovskii, A. A. (1974). Evolution of molecular hydrogen by cultures of blue-green algae. *Doklady Akademii Nauk SSSR* 213, 557–560.

Otaki, H., Everroad, R. C., Matsuura, K., and Haruta, S. (2012). Production and consumption of hydrogen in hot spring microbial mats dominated by filamentous anoxygenic photosynthetic bacterium. *Microbes Environ.* 27, 293–299. doi: 10.1264/jsme2.ME11348

Overmann, J., and Garcia-Pichel, F. (2000). "The phototrophic way of life," in *The Prokaryotes: An Evolving Electronic Resource for the Microbiologicial Community,* 3rd Edn. ed. M. Dworkin (New York: Springer).

Revsbech, N. P. (1989). An oxygen microelectrode with a guard cathode. *Limnol. Oceanogr.* 34, 474–478. doi: 10.4319/lo.1989.34.2.0474

Revsbech, N. P., and Jørgensen, B. B. (1986). Microelectrodes: their use in microbial ecology. *Adv. Microb. Ecol.* 9, 293–352. doi: 10.1007/978-1-4757-0611-6_7

Skyring, G. W., Lynch, R. M., and Smith, G. D. (1988). Acetylene reduction and hydrogen metabolism by a cyanobacterial/sulfate-reducing bacterial mat ecosystem. *Geomicrobiol. J.* 6, 25–31. doi: 10.1080/01490458809377819

Skyring, G. W., Lynch, R. M., and Smith, G. D. (1989). "Quantitative relationships between carbon, hydrogen, and sulfur metabolism in cyanobacterial mats," in *Microbial Mats, Physiology Ecology of Benthic Microbial Communities,* eds Y. Cohen and E. Rosenberg (Washinton, DC: American Society for Microbiology Press), 170–179.

Spear, J. R., Walker, J. J., McCollom, T. M., and Pace, N. R. (2005). Hydrogen and bioenergetics in the *Yellowstone geothermal* ecosystem. *Proc. Nat. Acad. Sci. U.S.A.* 102, 2555–2560. doi: 10.1073/pnas.0409574102

Stal, L. J., and Moezelaar, R. (1997). Fermentation in cyanobacteria. *FEMS Microbiol. Rev.* 21, 179–211. doi: 10.1016/S0168-6445(97)00056-9

Visscher, P. T., Baumgartner, L. K., Buckley, D. H., Rogers, D. R., Hogan, M. E., Raleigh, C. D., et al. (2003). Dimethyl sulphide and methanethiol formation in microbial mats: potential pathways for biogenic signatures. *Environ. Microbiol.* 5, 296–308. doi: 10.1046/j.1462-2920.2003.00418.x

Wang, R., Healey, F. P., and Myers, J. (1971). Amperometric measurements of hydrogen evolution in *Chlamydomonas. Plant Physiol.* 48, 108–110. doi: 10.1104/pp.48.1.108

Warthmann, R., Cypionka, H., and Pfennig, N. (1992). Photoproduction of H_2 from acetate by syntrophic co-cultures of green sulfur bacteria and sulfur-reducing bacteria. *Arch. Microbiol.* 157, 343–348. doi: 10.1007/BF00248679

Wieland, A., and Kühl, M. (2000). Short term temperature effects on oxygen and sulfide cycling in a hyper-saline cyanobacterial mat (Solar Lake, Egypt). *Mar. Ecol. Progr. Ser.* 196, 87–102. doi: 10.3354/meps196087

Wieland, A., Zopfi, J., Benthien, M., and Kühl, M. (2005). Biogeochemistry of an iron-rich hypersaline microbial mat (Camargue, France): oxygen, sulfur and carbon cycling. *Microb. Ecol.* 49, 34–49. doi: 10.1007/s00248-003-2033-4

Wiesenburg, D. A., and Guinasso, N. L. Jr. (1979). Equilibrium solubilities of methane, carbon monoxide and hydrogen in water and sea water. *J. Chem. Engin. Data* 24, 359–360. doi: 10.1021/je60083a006

Witty, J. F. (1991). Microelectrode measurements of hydrogen concentrations and gradients in legume nodules. *J. Exp. Bot.* 42, 765–771. doi: 10.1093/jxb/42.6.765

Wolin, M. J. (1982). "Hydrogen transfer in microbial communities," in *Microbial Interactions and Communities,* eds A. T. Bull and J. H. Slater (New York: Academic Press, Inc.), 323–356.

Fermentation couples *Chloroflexi* and sulfate-reducing bacteria to *Cyanobacteria* in hypersaline microbial mats

Jackson Z. Lee[1,2]*, Luke C. Burow[1,3], Dagmar Woebken[1,3†], R. Craig Everroad[1], Mike D. Kubo[1,4], Alfred M. Spormann[3], Peter K. Weber[5], Jennifer Pett-Ridge[5], Brad M. Bebout[1] and Tori M. Hoehler[1]

[1] Exobiology Branch, NASA Ames Research Center, Moffett Field, CA, USA
[2] Bay Area Environmental Research Institute, Sonoma, CA, USA
[3] Departments of Civil and Environmental Engineering, and Chemical Engineering, Stanford University, Stanford, CA, USA
[4] The SETI Institute, Mountain View, CA, USA
[5] Lawrence Livermore National Lab, Chemical Sciences Division, Livermore, CA, USA

Past studies of hydrogen cycling in hypersaline microbial mats have shown an active nighttime cycle, with production largely from *Cyanobacteria* and consumption from sulfate-reducing bacteria (SRB). However, the mechanisms and magnitude of hydrogen cycling have not been extensively studied. Two mats types near Guerrero Negro, Mexico—permanently submerged *Microcoleus* microbial mat (GN-S), and intertidal *Lyngbya* microbial mat (GN-I)—were used in microcosm diel manipulation experiments with 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU), molybdate, ammonium addition, and physical disruption to understand the processes responsible for hydrogen cycling between mat microbes. Across microcosms, H_2 production occurred under dark anoxic conditions with simultaneous production of a suite of organic acids. H_2 production was not significantly affected by inhibition of nitrogen fixation, but rather appears to result from constitutive fermentation of photosynthetic storage products by oxygenic phototrophs. Comparison to accumulated glycogen and to CO_2 flux indicated that, in the GN-I mat, fermentation released almost all of the carbon fixed via photosynthesis during the preceding day, primarily as organic acids. Across mats, although oxygenic and anoxygenic phototrophs were detected, cyanobacterial [NiFe]-hydrogenase transcripts predominated. Molybdate inhibition experiments indicated that SRBs from a wide distribution of DsrA phylotypes were responsible for H_2 consumption. Incubation with ^{13}C-acetate and NanoSIMS (secondary ion mass-spectrometry) indicated higher uptake in both *Chloroflexi* and SRBs relative to other filamentous bacteria. These manipulations and diel incubations confirm that *Cyanobacteria* were the main fermenters in Guerrero Negro mats and that the net flux of nighttime fermentation byproducts (not only hydrogen) was largely regulated by the interplay between *Cyanobacteria*, SRBs, and *Chloroflexi*.

Keywords: microbial mats, hydrogen, fermentation, Guerrero Negro, NanoSIMS

INTRODUCTION

Hypersaline microbial mats, living analogs of early life on Earth (Des Marais, 2003), are compact and structured laminations of highly diverse microbial communities that undergo significant redox changes over the diel (day-night) cycle, alternating between oxic and anoxic states. The metabolic diversity of microbial mats is reflected in a diverse potential for H_2 metabolism (Hoehler, 2005), and nitrogen fixation (Omoregie et al., 2004a,b), and previous work has documented significant efflux of H_2 from hypersaline microbial mats (Hoehler et al., 2001; Burow et al., 2012) under dark anoxic conditions. Hence, these systems are of interest not only in an ecological frame of reference, but also for bioenergy science. In previous work we showed that nighttime production of hydrogen gas from hypersaline mats of Elkhorn Slough, CA, USA originated within the photic layer of the mats and was primarily attributable to the fermentation activity of *Cyanobacteria*,

especially the dominant filamentous cyanobacterium *Microcoleus chthonoplastes*, and was largely insensitive to nitrogen fixation (Burow et al., 2012). In the Elkhorn Slough mats, hydrogen consumers [sulfate-reducing bacteria (SRB)] were present in close physical association with hydrogen producers, and significantly reduced hydrogen efflux (Burow et al., 2013, in press). Understanding the ecological and environmental factors that control net H_2 production is thus critical to understanding the role of H_2 cycling in mat structure and ecology.

In the present work we have characterized and quantified fermentative activity and consumption of fermentation products in two mat types from the previously documented site at Guerrero Negro, B.C.S., Mexico, that exhibit a range of net H_2 production rates. "GN-S" are well-developed subtidal mats located in pond 4 near pond 5 of the salt works and constructed primarily by the cyanobacterium *Microcoleus chthonoplastes* that have been

described extensively in previous reports (Spear et al., 2003; Ley et al., 2006; Feazel et al., 2008; and Robertson et al., 2009). "GN-I" are intertidal mat communities constructed largely by *Lyngbya* sp. The difference in dominant cyanobacterium and extent of development of accessory populations are reflected in differing chemical behavior of the two mat types, including in H_2 efflux. Previous work on the GN-I mats documented that the integrated H_2 production rate is equivalent to 16% of net daytime carbon fixation (on a per-electron basis), and individual bubbles at the mat surface may contain up to 10% hydrogen in the pre-dawn hours (Hoehler et al., 2001; Hoehler, 2005). Moreover, the amount of H_2 efflux has been found to vary over more than four orders of magnitude, as a function of environmental forcing (Bebout et al., 2002, 2004; Hoehler, 2005; Burow et al., 2012). In comparison, the submerged mats experience relatively stable conditions and release less H_2.

The extensive diversity, dominance patterns, and dynamic microbial response over space and time found within microbial mats are well suited for the development and deployment of advanced sequencing and isotope probing techniques to identify the complex biological interactions as well as energy and nutrient cycles of these highly diverse systems. When combined with traditional biogeochemical techniques, these methods can provide insights into energy and nutrient cycling in and through these systems. Using a comparative approach, the relative role of fermentation in the carbon and hydrogen cycles of these two mat types, and the ecology surrounding these cycles, was examined via a holistic set of molecular, isotopic, and biogeochemical methods. Specifically, we employed pyrotag libraries, functional gene sequencing of [NiFe]-hydrogenase (*hoxH*) and dissimilatory sulfite reductase (*dsrA*), stable isotope probing of labeled ^{13}C-bicarbonate and ^{13}C-acetate, Catalyzed Reporter Deposition Fluorescence In Situ Hybridization (CARD-FISH) probing of *Chloroflexi* and SRB clades, in combination with measurements of hydrogen, hydrogen sulfide, and organic acids to examine microbial mats manipulated with inhibitors that disrupt the sulfur, nitrogen, and carbon cycles. The results from these experiments indicated that in GN-I mats constitutive fermentation served to liberate roughly 80% of the photosynthetically fixed electrons into the bulk pool and thereby formed a basis for close trophic coupling between *Cyanobacteria*, filamentous anoxygenic phototrophs, and SRBs in hypersaline mats of both Guerrero Negro and of Elkhorn Slough.

MATERIALS AND METHODS

FIELD SITE AND SAMPLE COLLECTION

Water and whole 30×30 cm sections of microbial mats were harvested in September 2010 and September 2011 from sites located near Guerrero Negro, B.C.S. Mexico. Two distinct mat types were selected for study. Mats from the seawater concentration area (pond 4 near 5) of the Exportadora de Sal, S.A. (ESSA) Guerrero Negro, B.C.S. saltworks (27° 41′ 20.6″ N, 113° 55′ 1.2′ W) have a well-documented community profile (Spear et al., 2003; Ley et al., 2006; Feazel et al., 2008; and Robertson et al., 2009). These mats have experienced a largely quiescent environment with permanent cover of approximately 0.5–1 m of water at ~80–100‰ salinity, and can grow to approximately 10 cm thick (Des Marais, 1995; Nübel et al., 1999). This mat type, which is constructed

primarily by the cyanobacterium *Microcoleus chthonoplastes* is referred in this study as "GN-S" (Guerrero Negro—Submerged). The second mat type was collected from the intertidal flats bordering Laguna Ojo de Liebre, just outside the ESSA salt works (27° 45′ 30.2″ N, 113° 59′ 42.8″ W), and is referred to as "GN-I" mats (Guerrero Negro—Intertidal). The GN-I mats, constructed primarily by the cyanobacterium *Lyngbya* spp., experience periodic tidal desiccation and breakup along a sandy sloping shore ecological gradient (Rothrock and Garcia-Pichel, 2005). Mats were also collected from Elkhorn Slough, CA, USA in November 2011 in order to facilitate comparison of the present work with previously published studies of microbial mat fermentation and H_2 production and consumption (Burow et al., 2012; Woebken et al., 2012). Samples were returned to a greenhouse facility at NASA Ames and maintained in UV transparent acrylic boxes under ~3 cm of water collected from the site specific to the individual mat types as previously described (Bebout et al., 2002). The mats received natural solar irradiance and regulated temperature environment designed to mimic natural daily fluctuations around the *in situ* average of ~19°C.

DIEL MANIPULATIONS

Diel (24 hour) studies of these mat types were performed with inhibitors of specific metabolic processes and in conjunction with physical disruption by homogenization, all conducted under natural light conditions. All manipulations were performed by placing replicate 11 mm diameter cores of the top 2 mm of the microbial mats into 14 ml glass serum vials with 4 ml site water. Bottles were closed with butyl rubber stoppers, crimp sealed, flushed with nitrogen gas, and incubated under controlled temperature (Burow et al., 2012). Six replicates were prepared for each experimental condition. The four distinct manipulation experiments were:

(1) DCMU [3-(3,4-dichlorophenyl)-1,1-dimethylurea, a specific inhibitor of Photosystem II (Oremland and Capone, 1988; Oettmeier, 1992), and therefore of oxygenic photosynthesis] was used to evaluate the contribution of oxygenic photosynthetic carbon fixation as a driver of the release of fermentation products. DCMU was added to mats at dawn at a final concentration of 20 μM.

(2) Ammonium chloride was added as a source of combined nitrogen in order to suppress dinitrogen-fixing activity (Bebout et al., 1987; Burow et al., 2012) and thereby investigate whether N-fixation activity (and associated energy demand) affected rates of fermentation. Ammonium chloride was added at dusk at a final concentration of 8.8 mM.

(3) Sodium molybdate, a competitive inhibitor of sulfate reduction (Oremland and Capone, 1988), was used to quantify the role of sulfate reducers in consuming fermentation products. Sodium molybdate was added to mat incubations in site water and in sulfate-free artificial ASN5 media (to enhance the competitive effect) at dusk at a final concentration of 30 mM.

(4) Mat microbes were physically disrupted from each other with a Tissue Master (TM125-115, Omni International, Kennesaw, GA, USA) in order to investigate the importance of close spatial associations between microbes in internal cycling

vs. diffusive loss of fermentation products (Burow et al., in press).

BIOGEOCHEMICAL METHODS
Analysis of H₂ and organic acids

To measure H_2 net flux, 25 µL of headspace was withdrawn with a volumetric syringe from the headspace of each replicate microcosm vial at several time points during an incubation period. Samples were either analyzed immediately by direct injection onto a gas chromatograph with an HgO reduction detector (Trace Analytical) (Burow et al., 2012), or were preserved for later analysis as a small (1 mL) gas sample in a serum vial containing saturated NaCl solution that had been sparged with nitrogen gas for 20 min. To analyze organic acids, the entire liquid phase (4 mL) of each of three replicate microcosms for each control or manipulation experiment was sampled (with the associated incubation sacrificed). Liquid was filtered through 0.2 µm syringe filters for storage in ashed glass vials at −20°C. Organic acids (C1–C5) were quantified via high-pressure liquid chromatography (Albert and Martens, 1997).

Analysis of glycogen, Dissolved Inorganic Carbon (DIC), and hydrogen sulfide

Whole 1-cm diameter subcores were collected from incubated microbial mats and immediately frozen in liquid nitrogen. The uppermost 2 mm was sub-sectioned while still frozen and was subsequently freeze dried. The dried samples were ground and sonicated (3 min, setting 6, Fisher 60 Dismembrator) and transferred to a 7 ml glass screwcap vial. The vial was then capped and put in a boiling water bath for 6 min to solubilize glycogen and stop any enzyme activity in the extract. 50 µl of the extract was filter sterilized (0.45 µm) and added to 50 µl of amyloglucosidase solution (Keppler and Decker, 1974) (0.01 g in 2 ml deionized water) in a 1.5 ml screwcap vial. The vial was placed horizontally in a 40°C heating block on an orbital shaker so that the vial could roll slightly to aid mixing. After 1 h, the vial received 100 µl of a derivatization solution consisting of 60 mg anthranilamide (Sigma-Aldrich, St. Louis, MO, USA), 40 mg sodium cyanoborohydride (Sigma-Aldrich), 0.6 ml glacial acetic acid and 1.4 ml dimethyl sulfoxide (Bigge et al., 1995) and then heated to 70°C for 1 h followed by isocratic HPLC separation and fluorescence detection (270 nm excitation, 430 nm emission). The solvent consisted of 20 ml tetrahydrofuran, 6 ml butylamine, 10 ml phosphoric acid, 12 ml tetraethyl ammonium hydroxide and 3950 ml deionized water (Anumula, 1994).

DIC was quantified in separate flux chamber experiments of whole mats based on the methods of Hoehler et al. (2001). 1.5 mL fluid samples were collected in 3 mL plastic syringes from 1.5 L flux chambers placed on mats in the greenhouse, closed by means of a 3-way stopcock, and stored at 4°C until analysis (typically within 2–3 h, but in no case greater than 48 h after collection). Samples were analyzed via flow injection analyzer (FIA) (Hall and Aller, 1992). Duplicate injections were made for each incubation time point but, due to limited volume of incubation fluids, replicate samples were not taken.

Hydrogen sulfide measurements were based on the work of Cline (1969) and consisted of a N,N-Dimethyl-p-phenylenediamine sulfate salt and iron(III) chloride colorimetric method (read at 670 nm) with a sodium sulfide standard.

PHOTOSYNTHESIS/FERMENTATION MASS AND ELECTRON BALANCE

Net flux data collected from several diel experiments were used to estimate the proportion of photosynthetic carbon uptake that was subsequently mobilized in fermentation. We compared DIC uptake and flux of fermentation products on the common basis of "electron equivalents" to enable us to include H_2 in the calculation. The flux amounts were used to determine total daytime and nighttime DIC flux and averaged to obtain the average flux of inorganic carbon taken up and released by mats over a single diel cycle. Glycogen day and night differences were averaged and used to estimate net fixed carbon accumulation and depletion. Net and total production of organic acids and hydrogen were estimated, respectively, from replicated control and physical disruption experiments. All fluxes were normalized to the surface area of incubated mat cores and then converted to "electron equivalents" based on total charge state $[(CH_2O)_n = 4n,$ $CH_3COOH = 8,$ etc.] of each chemical species. Photoautotrophy was assumed to fix four electrons per carbon. Standard deviations were distributions of replicate measurements propagated to final results.

MOLECULAR BIOLOGY METHODS
Nucleic acid isolation and pyrotag sequencing of 16S Small Subunit (SSU) rRNA genes and transcripts

In October 2011 (1 month after field collection), samples for pyrotag libraries of both GN mat types were collected at 1200–2400 h from unaltered mats kept in site water in a greenhouse environment. Nucleic acid extraction [RNA clean extractions for nucleic acids (DNA and RNA) followed by cDNA reverse-transcription] was performed as per Burow et al. (2012). The upper 2 mm phototrophic layer was homogenized and extracted using acid-phenol and DNA/RNA cleanup and separation with the Qiagen RNeasy Mini Kit and the QIAamp DNA Mini Kit as per the manufacturers' protocol (Qiagen, Venlo, The Netherlands). Extractions were done in triplicate and pooled. The V8 hypervariable region of the 16S SSU rRNA gene was amplified from DNA (rRNA gene) or cDNA (rRNA reverse-transcript) templates using the universal primer pair 926f/1392r (Engelbrektson et al., 2010), including the titanium adaptor sequences and a five-base barcode on the reverse primer. Sequencing and bioinformatics processing was completed by Research and Testing Laboratory, LLC (Lubbock, TX, USA) using an in-house denoising, demultiplexing, clustering, and taxonomic assignment pipeline (http://www.researchandtesting.com/). Operational Taxonomic Unit (OTU) tables of denoised and dereplicated sequences with taxonomies identified by BLASTN+ on a custom NCBI database were received from the vendor. Taxonomy and taxonomic levels were based on BLAST percent similarity to best match reference sequences (same species = 97% ID, genus = 95%, etc.). Population statistics were computed using subsampling ($n = 100$) to the smallest library size (11,500) using QIIME (Caporaso et al., 2010) and the average and distribution of Chao1 and ACE metrics determined. Interactive Krona HTML5 (Ondov et al.,

2011) hierarchical pie chart community profiles of all pyrotag libraries have been included in the supplemental information online as Krona_charts_supplemental.zip.

Degenerate bidirectional [NiFe]-hydrogenase (hoxH) and dissimilatory sulfite reductase (dsrA) gene sequencing

Primers HoxH_F37 (5′ -ATHGARGGHCAYGCBAARAT-3′) and HoxH_R518 (5′ -ACNCCICCVGGNAYHGHCCA-3′) developed to amplify reversible nickel-iron (type 3b) hydrogenase transcripts containing the L1 and L2 motifs (Vignais and Billoud, 2007) and used previously (Burow et al., 2012) were used to sequence hydrogenases from both mat types from nighttime manipulations (2 96-well plates each from GN-I and GN-S). Primers dsrA_1R (5′-ACSCACTGGAAGCACG-3′) (Wagner et al., 1998) and dsrA_DGGE_R (5′- CGGTGMAGYTCRTC −3′) (Leloup et al., 2009) designed to detect sulfate reduction capability from SRBs were used to sequence *dsrA* genes from both mat types in nighttime manipulations (Burow et al., in press) (2 cDNA and 1 DNA 96-well plates each from GN-I and GN-S). PCR and cloning (described in Burow et al., 2012) were used to prepare 96-well format clone libraries of both DNA and cDNA for expression ratio profiling. Sequencing was completed using Single Pass T3 primer end sequencing runs at Beckman Coulter Genomics (Danvers, MA, USA) using BigDye Terminator v3.1 sequencing on an ABI PRISM 3730*xl* (Life Technologies, Carlsbad, CA, USA).

Phylogenetic analysis of bidirectional [NiFe]-hydrogenase (hoxH) and dissimilatory sulfite reductase (dsrA) clone sequences

All sequences were quality trimmed (Q30, 10-base average), and filtered for primer and plasmid regions and translated to the amino acid coding frame using Geneious (Biomatters, Auckland, New Zealand). Sequences were exported and then clustered by CD-HIT (Li and Godzik, 2006) at the 97% similarity level and queried against the NCBI non-redundant peptide database using the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990). Sequences were then aligned to custom databases composed of reference *dsrA* or *hoxH* sequences from environmental samples and isolated bacteria using CLUSTALX2 (Larkin et al., 2007) and then manually curated. The final alignment, excluding gapped positions, included 134 residues for 40 OTUs and 84 reference sequences for *dsrA* and 156 residues for 17 OTUs and 46 reference sequences for *hoxH*. For *hoxH* genes, 4 high-variability amino acid positions in the alignment (<50% conserved) were removed from the alignment. For each gene, maximum likelihood phylogenetic trees were constructed using PhyML (Guindon et al., 2010) with the LG+I+G substitution matrix as selected using the AIC criterion using Prottest 2.4 (Abascal et al., 2005; Le and Gascuel, 2008). Neighbor-joining was performed in MEGA5.2 (Tamura et al., 2011) using the JTT+I+G substitution matrix (Jones et al., 1992), which was the best-fit matrix available, as determined by Prottest. ML and NJ analyses underwent 100 and 1000 bootstrap iterations, respectively. Deeply branching or closely related reference sequences were pruned from trees for readability.

DATA ARCHIVING

Representative clone sequences of *hoxH* and *dsrA* OTUs were submitted to GenBank as accession KF582421-KF582479. Pyrotag libraries were submitted as *.sff files to NCBI's Sequence Read Archive (SRA) under BioProject PRJNA219681 (accession SRP030038).

ISOTOPE LABELING METHODS
^{13}C-Bicarbonate and ^{13}C-acetate labeling of GN mats and isotope-ratio mass spectrometry (IRMS)

Small subcores (11-mm diameter, 2-mm depth) of each microbial mat type were cut from whole sections of intact microbial mat and placed in serum vials with 4 ml site water. ^{13}C-bicarbonate was added for 10 daytime hours to examine autotrophic light-driven incorporation. [2-^{13}C]-acetate (0.2 mM) labeled FISH-NanoSIMS was completed according to Burow et al. (2012) and consisted of 10 hour overnight incubations in microcosm vials. Unlabeled mat sections, paraformaldehyde (PFA) fixed mat cores, and continuously dark microcosms served as controls. Mat cores were washed twice in label free media and immediately flash frozen at −80°C prior to FISH and NanoSIMS analyses. Bulk isotope ratios for $^{13}C/^{12}C$ (reported as $\delta^{13}C$ in permil relative to VPDB) were determined by IRMS (ANCA-IRMS; PDZ Europa Limited, Crewe, England) at the University of California, Berkeley, with IAEA and NIST peach leaf standards used for C isotope standard corrections (Woebken et al., 2012). For each mat type, IRMS measurements were analyzed for significant differences using Student's *t*-test between different trials.

FISH and NanoSIMS

As previously described, CARD-FISH was used to hybridize fluorescent oligonucleotides to cells of *Chloroflexi* (CFX1223 and GNSB-941) (Woebken et al., 2012; Burow et al., 2013) and *Desulfosarcina / Desulfobacteraceae* (DSS658) (Burow et al., in press). *M. chthonoplastes* bundles were directly identified by morphology using scanning electron microscopy (SEM) on a FEI Inspect F (Hillsboro, Oregon, USA). Using the FISH, Chlorophyll *a* autofluorescence, and SEM micrographs as a guide to the location of specific cells types, high-resolution secondary ion mass spectrometry (SIMS) was performed at Lawrence Livermore National Laboratory with a Cameca NanoSIMS 50 (Gennevilliers, France). NanoSIMS data were analyzed for normal distribution with the Student's *t*-test and Shapiro-Wilk W test, and in cases where the data did not meet the standard ($p < 0.05$) of normality, the Wilcoxon (Whitney-Mann) test was used to confirm the initial result of significant differences using Student's *t*-test. Statistical tests were computed in R (2.15.1) (Ihaka and Gentleman, 1996). The $^{13}C/^{12}C$ ratio was measured using $^{12}C_2$ and $^{13}C/^{12}C$ corrected for the dimer abundances and converted to permil enrichment (Pett-Ridge and Weber, 2012). Measurements were made in both imaging and spot analysis modes. Uptake of the respective ^{13}C-labeled compound was determined based on ^{13}C enrichment relative to unlabeled, PFA fixed samples as a reference.

RESULTS
PYROTAG ASSAY RESULTS OF GN-I AND GN-S MATS

16S SSU rRNA gene pyrotag libraries provide a way to survey a large fraction of a microbial community to determine community composition and population statistics, as well as to track population changes between samples (Hamady et al.,

2008; Kuczynski et al., 2010). **Table 1** shows the estimated OTU richness and sequencing depth of the samples collected at 1200 and 2400 over the same diel cycle. DNA libraries were much more diverse than cDNA libraries and indicated that a small number of clades were responsible for a large fraction of the ribosomal expression in mats. Overall, coverage based on Chao1 and ACE estimators was approximately half in samples, though the proper usage of alpha diversity estimators in high-throughput sequencing studies (combined with denoising or clustering) is still debated (Reeder and Knight, 2010; Gihring et al., 2012) and therefore should be interpreted as a preliminary measure of coverage. Assuming that dominant organisms received adequate sequencing coverage, only OTU expression ratios (cDNA: DNA) of genera >1% in DNA libraries were examined during downstream analysis and only the top genera in phyla were highlighted. In **Figure 1A**, phylum level results of pyrotag libraries showed *Proteobacteria*, *Cyanobacteria*, and *Chloroflexi* dominate the DNA libraries, with relatively lower levels of *Bacteroidetes* than in previously sequenced GN-S samples (Ley et al., 2006). In GN-S mats, the OTUs from the genus *Microcoleus* had the highest expression ratio, and in GN-I mats, OTUs from the genus *Lyngbya* (followed by *Microcoleus*) had the highest expression ratio. While overall *Alphaproteobacteria* community composition was similar between the two GN mat types (**Figure 1B**), (primarily diazotrophic and purple non-sulfur (PNS) bacteria, Online Information, Krona pie charts), there were generally more *Alphaproteobacteria* sequences in GN-I samples than in GN-S samples. Additionally, *Proteobacteria* represented the largest fraction of the community, with major mat functional groups (Fenchel and Finlay, 1995) represented (i.e., PNS in *Alphaproteobacteria*, SRBs in *Deltaproteobacteria*, sulfur-oxidizing bacteria in *Gammaproteobacteria*) (Online Information). The third-most abundant phyla in the mats, *Chloroflexi* (**Figure 1C**), was divided between the phototrophic (class *Chloroflexi*) and the dark filamentous clades of the *Anaerolineae* and

Caldilineae. The most abundant identifiable genus level OTUs from phylum *Chloroflexi* matched reference sequences for the filamentous anoxygenic phototrophic *Oscillochloris* (and the *Oscillochloridaceae* generally, Online Information).

FIGURE 1 | Relative abundance of pyrotag sequences of intertidal (GN-I) and submerged (GN-S) microbial mats from Guerrero Negro, B.C.S., Mexico. (A) Phylum level overview showing most abundant and most active OTUs [*Microcoleus* (shadowed), *Lyngbya* (black)]; **(B)** abundance within *Proteobacteria*; and **(C)** abundance within *Chloroflexi* with shadowing for *Oscillochloris*.

Table 1 | OTU richness and sampling depth of pyrotag samples from GN-S and GN-I mats.

Library	OTU richness		
	Observed species	Chao1	ACE
		Avg. (*SD*)	Avg. (*SD*)
GN-S 1200 DNA	3042	5252 (49)	5743 (41)
GN-S 1200 cDNA	1944	3818 (179)	4124 (154)
GN-S 2400 DNA	3326	6068 (137)	6766 (123)
GN-S 2400 cDNA	1243	2322 (121)	2511 (97)
GN-I 1200 DNA	3417	6747 (180)	7465 (189)
GN-I 1200 cDNA	1404	2511 (105)	2643 (88)
GN-I 2400 DNA	3559	6679 (124)	7285 (120)
GN-I 2400 cDNA	1545	2840 (128)	3016 (107)

Average population statistics are shown with standard deviations of subsample distributions (n = 100, 11,500 sequences) in parentheses. Observed Species statistics were computed at the full sampling depth.
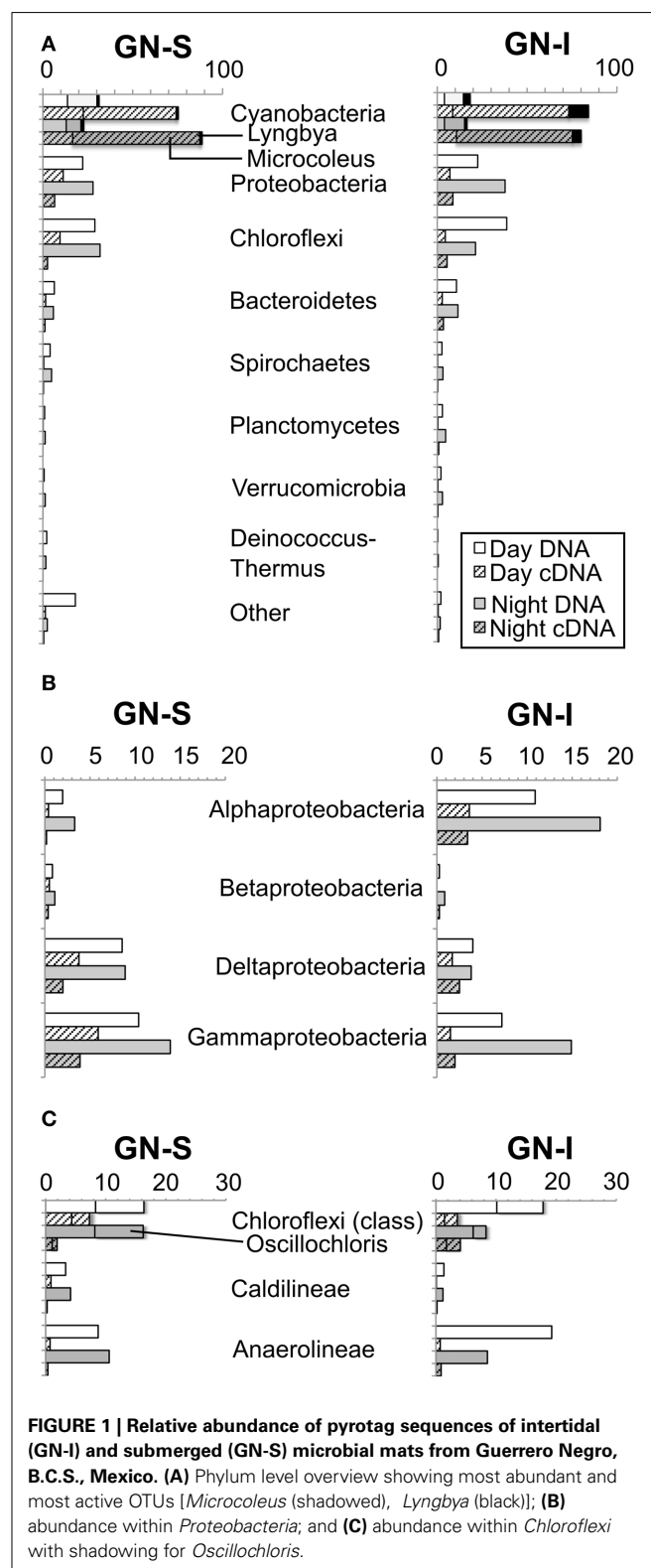
MICROCOSM STUDIES OF THE IMPACT OF NITROGEN FIXATION AND OXYGENIC PHOTOSYNTHESIS ON FERMENTATION

To determine whether similar processes and ecology account for the chemical cycling observed in Guerrero Negro mats as in Elkhorn Slough mats, both GN-S and GN-I mats were subjected to microcosm-based studies. Experiments were conducted on cores of harvested mat cut down to the top 2 mm and placed in microcosm vials with site water and sealed. Reduced gases and fermentation products were collected at periodic intervals over the diel. Hydrogen gas and organic acids were produced only during the dark portion of the diel cycle with more of each produced in GN-I mats (**Figure 2**), with the relative abundance of organic acids consistently corresponding to: acetate > formate > propionate (**Figure 2B**). Additionally, two separate manipulation experiments were conducted to differentiate the influence of nitrogen fixation and oxygenic photosynthesis on nighttime fermentation. Ammonium chloride manipulations were intended to suppress the effect of nitrogenase and nitrogen fixation to examine the impact on fermentation rates. No significant difference was observed on net hydrogen and organic acid production between control and nitrogen replete treatments (**Figures 2A,B**), with the exception of a reduction in net formate production in

GN-I mats. This is consistent with previous work conducted on Elkhorn Slough mats that showed that hydrogen production at night was largely uninfluenced by nitrogen fixation and constitutive fermentation was the main contribution to hydrogen production (Burow et al., 2012).

In the second experiment, DCMU (which inhibits photosystem II found in *Cyanobacteria*) was added in an effort to help clarify the role of oxygenic photosynthesis on fermentation activity in these mats. Inhibition was verified by oxygen microelectrode profiles showing anoxic conditions within 200 μm of the mat water interface (data not shown). Net hydrogen production was significantly lower in DCMU treated mats (**Figure 2C**), while net acetic acid production was higher (**Figure 2D**). This was the only manipulation experiment in which hydrogen and net organic acid production did not change in concert in a microcosm. **Figure 3** shows isotopic enrichment following ^{13}C-bicarbonate labeling of GN mats exposed to DCMU and light. In both mat types, the ^{13}C enrichment of DCMU-exposed mats resembled that of control mats incubated in the dark rather than light, suggesting that mats were not fixing carbon during the day. These observations are consistent with past studies that used DCMU to switch mats into an anoxic, nitrogen-fixing, sulfate-reducing
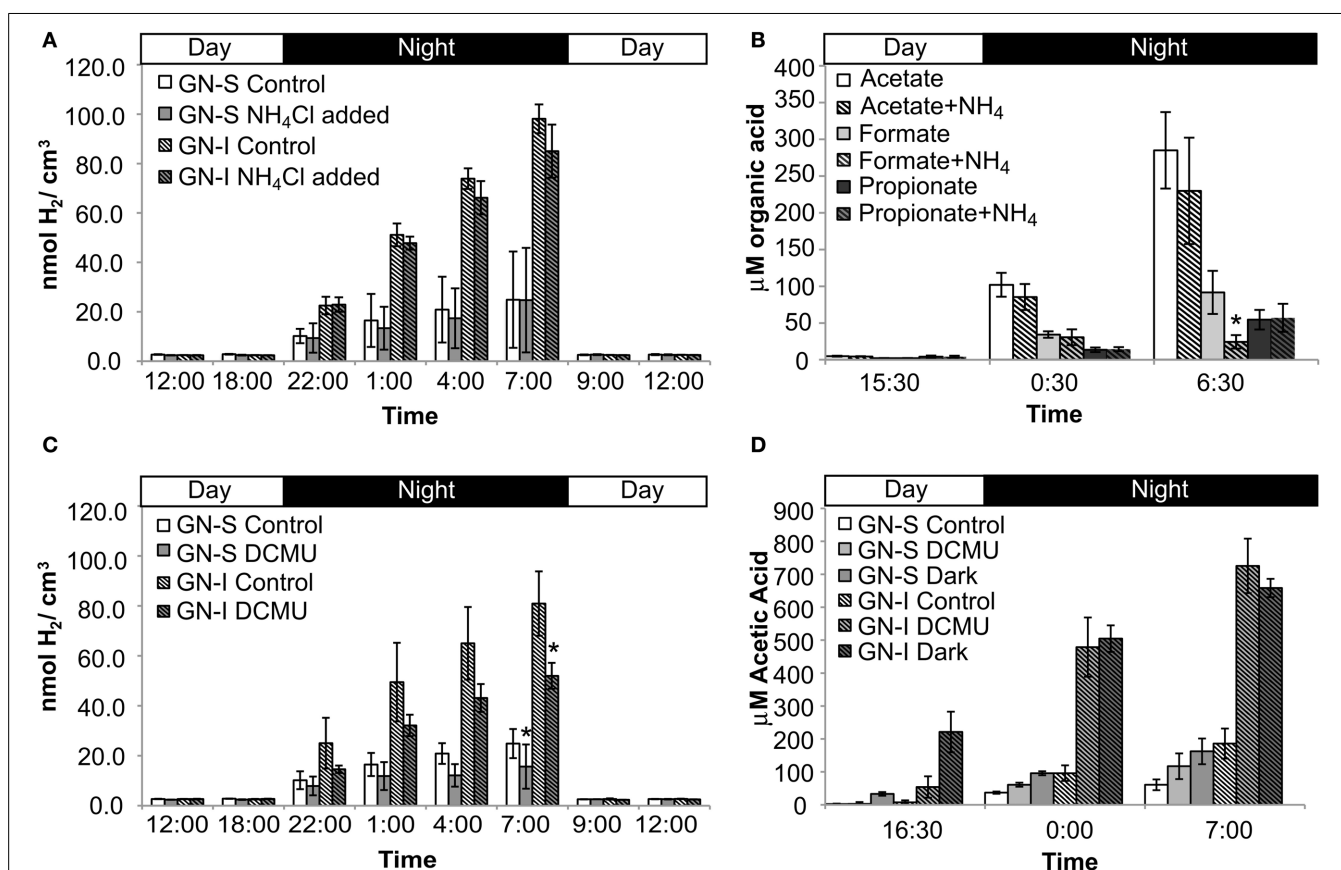
FIGURE 2 | Hydrogen and organic acid production in diel manipulation experiments with two microbial mats types from Guerrero Negro, B.C.S., Mexico. NH₄Cl manipulation (inhibition of nitrogen fixation): **(A)** H₂ production **(B)** net organic acid production (GN-I only). DCMU manipulation (photosynthesis inhibition via photosystem II disruption): **(C)** H₂ production and **(D)** net acetic acid production. Error bars indicate standard deviation of 6 replicate trials. Asterisk indicates $p < 0.05$ to control case by Student's *t*-test.

mode during daylight (Bebout et al., 1987, 1993; Steppe and Paerl, 2002). We further verified our findings by measuring daytime ^{13}C-bicarbonate uptake in individual *Microcoleus* morphotype filaments by NanoSIMS isotopic imaging (**Figure 4**). Control filaments incorporated significantly more labeled bicarbonate relative to DCMU-exposed filaments during daylight.

IDENTIFYING MAJOR HYDROGEN PRODUCERS AND CONSUMERS IN MATS BY HYDROGENASE (HoxH) PHYLOGENY

To identify the organisms responsible for the production and consumption of hydrogen in Guerrero Negro mats, 117 transcripts

FIGURE 3 | Isotopic enrichment (measured by IRMS) of GN-S and GN-I mats following ^{13}CHO$_3$- labeling and DCMU inhibition of photosystem II and photoautotrophy. Samples were collected after either 10 h of lighted label exposure, 10 h of light plus 12 h of darkness, or 12 h of darkness. Error bars are standard deviation of 4 replicate trials.
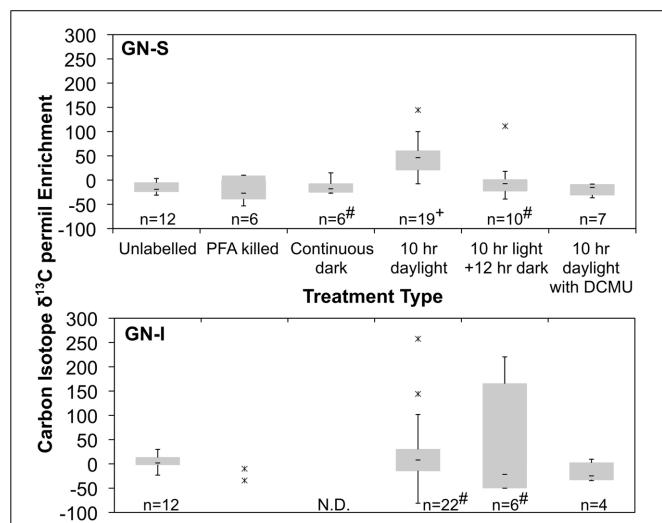
FIGURE 4 | Quartile box plots of NanoSIMS isotope ratio measurements of *Microcoleus* morphotype filaments labeled with ^{13}C-bicarbonate for trials of unlabeled, PFA killed, continuous dark, 10 h daylight, 10 h daylight followed by 12 h night, and 10 h daylight with DCMU added. Data was collected via multiple spot measurements of multiple filaments for each mat type (GN-S, at top; GN-I, at bottom). Asterisks are outlier datapoints identified as 1.5*IQR (Inter Quartile Range) outside quartiles. Outliers were retained in analysis. "#" denotes distributions for which the Shapiro–Wilk test $p < 0.05$. "+" denotes significantly different to all other GN-S trials (Wilcoxon test, $p < 0.05$). "N.D." denotes no data taken.

(in 17 OTUs) of [NiFe]-hydrogenases of type 3b were sequenced and passed quality control. **Figure 5** shows the phylogenetic relationship of OTUs detected from GN-I and GN-S mats. In both mat types, the majority of the observed hydrogenases were from *Cyanobacteria* of both filamentous and unicellular types. In GN-I mats, transcript OTUs related to *Lyngbya* sp. PCC 8106 were seen to be most abundant, and additional OTUs of mat-associated filamentous *Cyanobacteria* such as *Microcoleus chthonoplastes* PCC 7420 were also seen. In GN-S mats, the most abundant OTUs came from an unknown cyanobacterium in *Oscillatoriales* or *Pleurocapsales* whose closest BLAST match (∼94% ID by BLASTp) was *Pleurocapsa* sp. PCC 7319. Transcripts related to *Microcoleus chthonoplastes* PCC 7420 were detected as well. The remaining hydrogenases observed in these mats were most related to clusters of hydrogenases previously reported in Elkhorn Slough mats (Burow et al., 2012). Therefore, the distribution of abundant hydrogenases in hypersaline microbial mats appears to fit a pattern of cyanobacterial clusters, with additional diverse unidentified groupings of lower abundances. Notably, no sequences for hydrogenases from the anoxygenic phototrophs of the *Chloroflexi* were detected.

ACCUMULATION OF FIXED CARBON AND RELEASE OF FERMENTATION BYPRODUCTS IN MICROCOSM STUDIES

The presence of fermentation products in Guerrero Negro mat types, though detectable, may not be in meaningful quantity to influence the overall cycling of carbon in the ecosystem. Therefore, several experiments were completed to measure the relative magnitude of fixed carbon in relation to fermentation products released in GN-I mats. Changes in concentration of DIC in the water overlying incubated mats were taken to reflect total flux of carbon due to fixation or respiration (**Figure 6**). The observed flux was into the mats (negative) primarily during the daylight period and into the water column (positive) primarily during the nighttime period, consistent with previous work (Des Marais, 1995). Glycogen measurements from mat cores at different times were used to estimate the amount of stored fixed carbon in mats across the diel. Both mat types did exhibit a daytime accumulation (net gain) and nighttime loss (net loss) of glycogen, though the absolute magnitude of glycogen was 2.5-fold higher in GN-S mats than in GN-I mats while the net diel change differed by only 25% (higher in the GN-S mat) (**Figure 7**).

To relate these different fluxes and pools to each other, and to the fermentation results from the microcosm studies, an electron mass balance was constructed (**Table 2**) (See Section Photosynthesis/Fermentation Mass and Electron Balance for methods and conditions used). DIC and glycogen fluxes over the day and night period were averaged to determine a single flux over the diel. For organic acids and hydrogen, both "net" and "total" values are given. "Net" values were derived from fluxes observed under control conditions, and were assumed to incorporate the effects of H$_2$ consumption by spatially-associated accessory populations as well as H$_2$ production in whole intact mats. "Total" values were derived from fluxes observed when mats were homogenized to disrupt physical associations and thereby decrease or eliminate consumption by accessory organisms. Under these conditions, the flux of hydrogen, acetate, and propionate roughly

FIGURE 5 | Phylogenetic relationship of type 3b [NiFe]-hydrogenases detected from GN-I and GN-S mats. Maximum likelihood (ML) phylogram inferred from partial *hoxH* transcript sequence data for selected taxa, with an emphasis on the *Cyanobacteria* and *Chloroflexi*. Symbols denoting bootstrap support values are for both ML and neighbor joining analyses. Representative sequences for OTUs identified from this study are in bold. Numbers in parentheses after study sequences denote number of sequences for each OTU from GN-S mats (left), and GN-I mats (right).

FIGURE 6 | Flux chamber measurements of dissolved inorganic carbon (DIC) flux from GN-I mats over a diel period. Values reflect net flux of DIC between the indicated times. Error bars are standard deviations of 4 replicate trials.

FIGURE 7 | Glycogen accumulation over the diel cycle in GN-I and GN-S mats over the diel. Error bars represent standard deviation of 9 replicate cores per time point. Asterisk indicates $p < 0.05$ by unpaired 2-tailed Student's t-test to both 0700 sample sets.

doubled over controls. Because it is not certain that disruption completely eliminates consumption of fermentation products, this "total" value should be taken as a lower bound on the maximum possible production of fermentation products. Approximately 81% of the photosynthetically fixed electrons are released as total fermentation products during the night. Based on the differing "net" and "total" fluxes of fermentation products, it appears that about 41% of net photosynthetic carbon fixation was ultimately consumed by accessory populations within the mat. Quantitatively, hydrogen was a minor contributor to the overall cycling of photosynthetically-fixed electrons in these mats. The majority was accounted for as organic acids, specifically as acetate.

MICROCOSM STUDIES OF HYDROGEN CONSUMPTION IN MICROBIAL MATS

Similar to the microcosms used to test nitrogen fixation and oxygenic photosynthesis in hydrogen production, a combination of experimental microcosms were used to characterize and quantify microbial hydrogen consumption. Sulfate deprivation and molybdate addition were intended to reduce the activity of sulfate reduction, and were conducted both individually and together in microcosms. Disruption was intended to physically separate mat community members and was conducted in a separate microcosm experiment. Both sulfate deprivation alone, and sulfate deprivation with molybdate addition resulted in significantly diminished accumulation of hydrogen sulfide (the end product of sulfate reduction) within microcosm vials, but with the second case having a much larger effect (**Figure 8A**). In GN-I mats, net hydrogen production was enhanced by both disruption and molybdate addition with sulfate deprivation (**Figure 8B**), while hydrogen production in GN-S mats was enhanced only by sulfate deprivation with molybdate inhibition (**Figure 8D**). Because our method for quantifying organic acids is not compatible with molybdate, we could only examine the effect of physical disruption on organic acid production, which showed that production of acetate and propionate roughly double in GN-I mats due to physical disruption (**Figure 8C**).

ORGANIC ACID UPTAKE IN MICROBIAL MATS

Organic acids produced by nighttime fermentation have the potential to serve as substrates for both sulfate reduction and photoheterotrophy. The former was suggested by the significant enhancement of nighttime organic acid production upon physical disruption (**Figure 8C**). Photoheterotrophy was suggested in Elkhorn Slough mats by Burow et al. (2013) to be related to the activity of *Chloroflexi*. To further characterize and quantify organic acid uptake, mats from both GN and Elkhorn Slough were incubated overnight with ^{13}C-labeled acetate, and analyzed for isotopic enrichment by NanoSIMS. The final amended ^{13}C-acetate concentration (0.2 mM) was not greater than maximum values of accumulation in mats and represented similar concentrations in natural unamended mats. Probes specifically for phylum *Chloroflexi* (CFX1223 and GNSB-941) and for *Desulfobacteraceae* (DSS658) were used to determine if hybridized cells took up acetate. **Figure 9** shows NanoSIMS measurements of hybridized filaments for three mat types compared to microbial filaments with no hybridization. **Figures 9A–D** show paired Chlorophyll *a* autofluorescence (A), *Chloroflexi* probe CARD-FISH (B), and NanoSIMS ^{13}C enrichment (C), and NanoSIMS secondary electron images (D). **Figures 9E–H** shows a similar set of paired images for the *Desulfobacteraceae* probe. **Figure 9I** shows a quartile box plot summarizing NanoSIMS isotopic spot measurements over a number of different mat types and across replicate samples and the significance of the distribution when compared to filaments that did not hybridize with probes. All three mat types showed enrichment in CFX1223/GNSB-941 hybridized cells by NanoSIMS imagery and spot measurements, but varying patterns of enrichment between filaments within samples was observed. DSS658 hybridized filaments also showed enrichment, but generally much less than for CFX1223/GNSB-941 filaments. In GN-I mats, no DSS658 hybridization was visible in microscopy. In all cases, significant enrichment was observed over unhybridized filaments, but enrichment in *Chloroflexi*

Table 2 | Fermentation product concentration and equivalent electron balance in GN-I mat microcosms showing the amount of light captured electrons in fixed carbon analytes (during the day) and in fermentation by-products (produced at night).

	Daytime flux/storage	Nighttime flux/storage	Avg. diel flux/storage	e⁻ equivalents	e⁻ flux/ storage
	mmol/m² Avg. (*SD*)	mmol/m² Avg. (*SD*)	mmol/m² Avg. (*SD*)		mmol e⁻/m² Avg. (*SD*)
Fixed inorganic-C (CH_2O)	−121.4(10.3)	77.5(6.9)	99.5(8.6)	4	397.9(34.4)
Glycogen	21.2(25.2)	−31.6(30.4)	26.4(27.8)	24	633.8(667.2)
NET FERMENTATION					
Acetate	–	14.9(0.9)	14.9(0.9)	8	119.2(7.5)
Formate	–	8.3(1.2)	8.3(1.2)	1	8.3(1.2)
Propionate	–	2.4(0.5)	2.4(0.5)	13	31.7(6.1)
H_2	–	0.83(0.05)	0.83(0.05)	2	1.66(0.09)
TOTAL FERMENTATION					
Acetate	–	28.0(1.3)	28.0(1.3)	8	224.2(10.1)
Formate	–	10.0(1.0)	10.0(1.0)	1	10.0(1.0)
Propionate	–	6.6(1.2)	6.6(1.2)	13	85.3(15.1)
H_2	–	1.77(0.18)	1.77(0.18)	2	3.55(0.36)
% Fixed e⁻ fermented					81.2(14.6)
% Fixed e⁻ taken up by accessory microbes					40.8(14.0)

Values were normalized to microcosm mat diel flux or storage rate and core cross-sectional area, mmol/m². Net fermentation by-product measurements were from control microcosms, and total fermentation byproducts were from homogenized microcosms.

and *Desulfobacteraceae* was notably higher in Elkhorn Slough samples.

SURVEYING THE DIVERSITY OF SULFATE REDUCING BACTERIA THROUGH DsrA PHYLOGENY IN GUERRERO NEGRO MATS

DsrA phylogeny of sequences derived from both mat types are shown in **Figure 10**. Nodes of the family *Desulfobacteraceae* did not generally bootstrap well, but genera of the family *Desulfobacteraceae* formed a monophyletic clade, as did genera within the family *Desulfobulbaceae*. The constructed phylogeny was consistent with the work of Leloup et al. (2009) which showed *Desulfovibrionales* affiliating with *Desulfobulbaceae* within *Desulfobacterales* and also identified both the orthologous and xenologous *Desulfotomaculum*. In total, 347 cDNA and 105 DNA reads were obtained and binned into 40 OTUs for *dsrA* genes from the GN-I and GN-S mats. The maximum likelihood analysis for the translated *dsrA* data in **Figure 10** revealed that OTUs were distributed throughout the phylogeny. However, the majority of sequences were clustered into five main groups: two within the *Desulfobacteraceae* and three within the deeply branching regions of the tree. Though DSS658 labeled cells could not be found in GN-I mats by CARD-FISH, *dsrA* genes and transcripts belonging to *Desulfobacteraceae* were detected in both mat types. One main cluster, primarily from the GN-I mats, formed a well-supported clade with a DsrA sequence identified from the intertidal mats found at Elkhorn Slough, CA, USA (accession JX502749, Burow et al., in press), this may represent a novel "intertidal" lineage of *Desulfobacteraceae*. A second cluster did not affiliate closely with any described SRB taxa, and was not supported by bootstrap analysis, but was placed within the *Desulfobacteraceae*. This grouping was particularly high in

transcript content (197 cDNA reads vs. only 6 DNA reads) and distributed in both GN-I and GN-S environments. Finally, several OTUs grouped deep in the tree together in what Harrison et al. (2009) refers to as the "deep-branching *dsrA*." These OTUs tended to cluster with other cloned sequences from studies of sediments and marshes (Castro et al., 2002; Bahr et al., 2005) and have an as yet unknown role in these mats. No naming convention for clades was apparent in the numerous published DsrA phylogenies, but these deep OTUs affiliated with sequences belonging to what have been previously termed clade IV (Dhillon et al., 2003; Bahr et al., 2005; Zhang et al., 2008), clade V (Kaneko et al., 2007; Zhang et al., 2008; Harrison et al., 2009), and/or clade DSR-2 (Castro et al., 2002) and at least one clade may contain members derived from horizontal gene transfer (Mussmann et al., 2005). Of the deeply-branching clades, one deeply branching group contained an abundant OTU from GN-S, one contained an abundant OTU from GN-I, and the last contained rare OTUs from GN-I mats.

DISCUSSION

Previous work (Skyring et al., 1989; Burow et al., 2012) suggested that nighttime production of reduced gases results from photoautotrophy and storage of reduced carbon by *Cyanobacteria* with subsequent fermentation of stored photosynthate following the onset of dark, anoxic conditions (Hoehler et al., 2001; Des Marais, 2003). In Elkhorn Slough mats, *Cyanobacteria* were indicated as the dominant fermenter (Burow et al., 2012) and sulfate reducers as a key consumer of H_2 (Burow et al., in press). This study builds on previous reports of H_2 cycling in microbial mats in three important regards. First, bulk chemical cycling and the underlying ecology are shown to be common
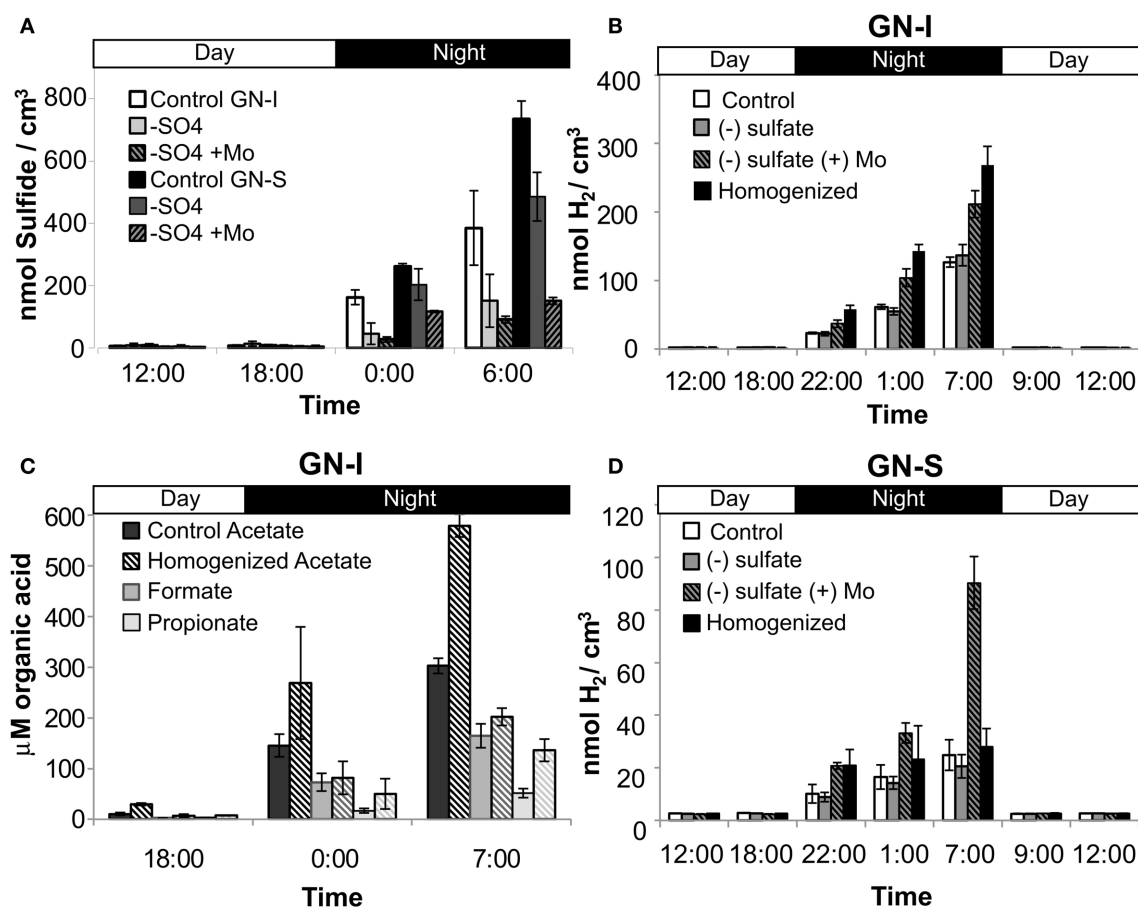
FIGURE 8 | Hydrogen consumption and SRB activity in two microbial mat types from Guerrero Negro, B.C.S., Mexico. Net hydrogen sulfide production (both mats) **(A)**, net hydrogen production in GN-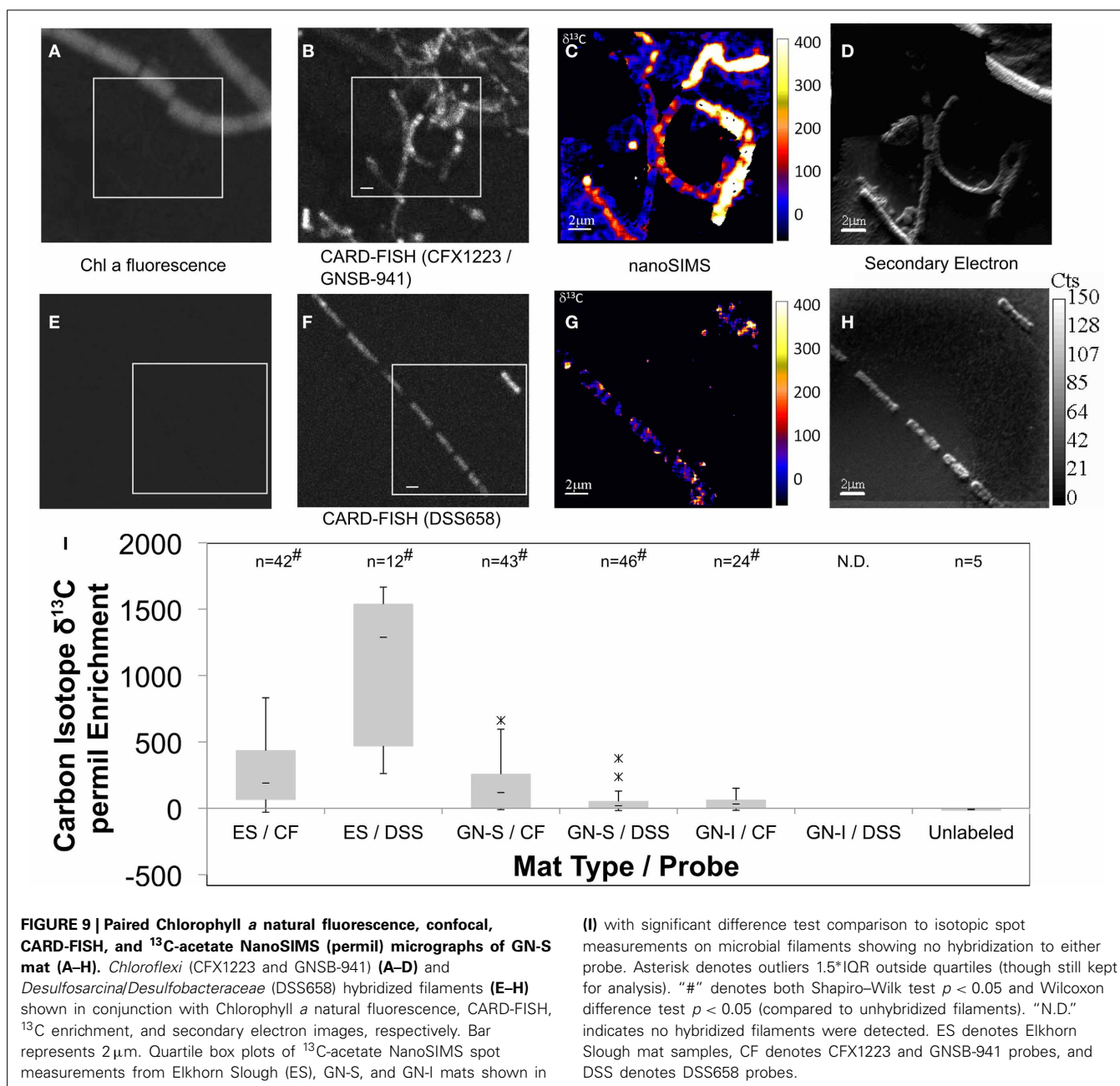I mats **(B)**, net organic acid production in GN-I mats **(C)**, and net hydrogen production in GN-S mats **(D)**, showing the effect of physical disruption, sulfate deprivation, and molybdate addition on microcosms. Error bars indicate standard deviation of 6 replicate trials.

features of geographically diverse mats, as well as mats constructed by distinctly different *Cyanobacteria*. Second, organic acid cycling is characterized and quantified, and shown to represent a significant component of overall carbon and electron flow in the studied mats. Last, multiple methods are utilized to demonstrate that exchange of fermentation products serves to directly link *Cyanobacteria* with sulfate reducers and anoxygenic phototrophs.

FERMENTATION IS QUANTITATIVELY IMPORTANT IN MATS

If daytime oxygenic photosynthetic fixation of carbon drives the subsequent nighttime fermentation to hydrogen and organic acids, a series of changes in metabolites should be observable as: (1) accumulation during the day and release at night of inorganic carbon based on measurements of the flux of dissolved inorganic carbon (DIC) across the mat-water interface, (2) accumulation during the day and depletion at night of small stored carbon polymers within the mat (e.g., glycogen), and (3) a rise in organic acid flux within the mat at night. These were all observed in the studies presented here. As shown in **Table 2** and

Figure 6, inorganic carbon was incorporated into mats during daytime. **Figure 7** indicated glycogen as the primary fixed carbon storage molecule in hypersaline microbial mats, with GN-S mats accumulating much more glycogen. Approximately ~81% of all carbon fixed during the day in GN-I mats was subsequently fermented at night, with most of the net accumulation of fermentation products occurring as organic acids rather than as hydrogen. The small amount of hydrogen released by fermentation activity relative to organic acids was unsuspected given that the hydrogen concentrations in the mat go through a four order of magnitude change in concentration throughout a diel cycle and particularly given that net hydrogen fluxes in GN-I mat are 10 times greater than in the GN-S mat. A stoichiometric fermentation of glucose to acetic acid, carbon dioxide, and hydrogen would produce nearly 2:1 electron ratios of acetate and hydrogen, but the ratios measured in this study were typically closer to 100:1 (**Table 2**) and suggests that organic acids provide the most quantitatively observable flux of reductant and energy available to the broader community of microbes within the mat under dark/anoxic conditions.

FIGURE 9 | Paired Chlorophyll *a* natural fluorescence, confocal, CARD-FISH, and ^{13}C-acetate NanoSIMS (permil) micrographs of GN-S mat (A–H). *Chloroflexi* (CFX1223 and GNSB-941) **(A–D)** and *Desulfosarcina/Desulfobacteraceae* (DSS658) hybridized filaments **(E–H)** shown in conjunction with Chlorophyll *a* natural fluorescence, CARD-FISH, ^{13}C enrichment, and secondary electron images, respectively. Bar represents 2 μm. Quartile box plots of ^{13}C-acetate NanoSIMS spot measurements from Elkhorn Slough (ES), GN-S, and GN-I mats shown in

(I) with significant difference test comparison to isotopic spot measurements on microbial filaments showing no hybridization to either probe. Asterisk denotes outliers 1.5*IQR outside quartiles (though still kept for analysis). "#" denotes both Shapiro–Wilk test *p* < 0.05 and Wilcoxon difference test *p* < 0.05 (compared to unhybridized filaments). "N.D." indicates no hybridized filaments were detected. ES denotes Elkhorn Slough mat samples, CF denotes CFX1223 and GNSB-941 probes, and DSS denotes DSS658 probes.

CONSTITUTIVE FERMENTATION BY *CYANOBACTERIA* WAS RESPONSIBLE FOR A MAJORITY OF HYDROGEN AND ORGANIC ACID PRODUCTION

The results of this study indicated similar diel patterns in organic acid and hydrogen production in hypersaline microbial mats from different locations and of different types. Because fermentation of photosynthate represents a loss of reducing power and delivers the lowest energy yield among potential catabolic processes, it could be viewed as a process to be minimized (that is, to be employed only if required by the demands of nighttime metabolism). One potentially large demand for such fermentation would be the energy required to fuel dinitrogen fixation

and, for this reason, we examined whether inhibition of dinitrogen fixation (by addition of ammonium as a source of combined nitrogen) diminished the yield of fermentation products. As shown in **Figure 2A**, and consistent with previous observations in Elkhorn Slough, ammonium addition does not appear to affect fermentation. This suggests that fermentation may be a constitutive aspect of metabolism in these mats, rather than being regulated in response to energetic or other demands of metabolism. The glycogen data and relative ratios of organic acid flux indicate that fermentation of glycogen to acetate was the dominant nighttime pathway, with a smaller component of mixed acid fermentation.
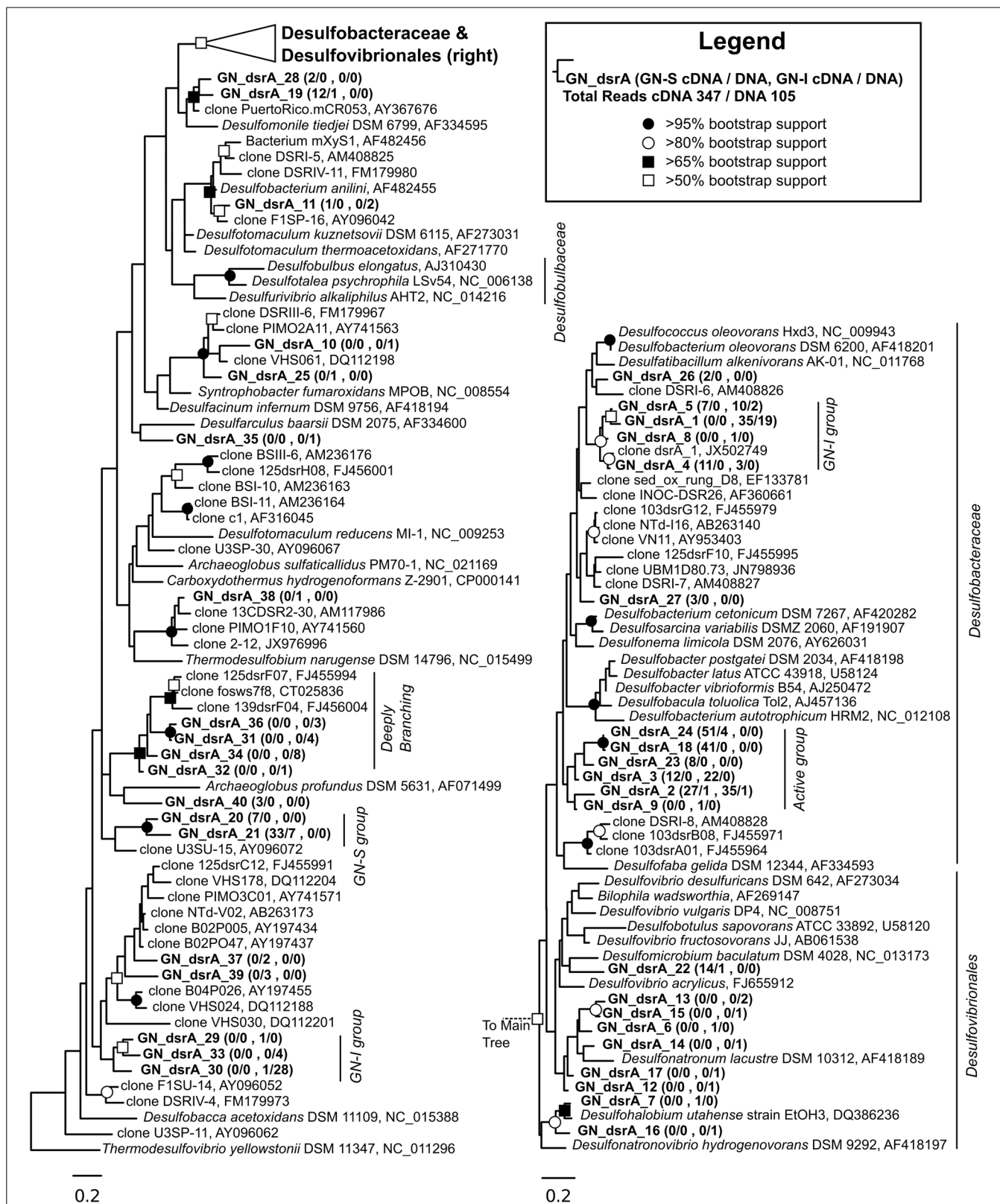
FIGURE 10 | Phylogenetic relationship of translated *dsrA* genes and transcripts detected from GN-I and GN-S mats. Maximum likelihood phylogram inferred from partial DsrA sequence data for selected taxa. Symbols denoting bootstrap support values are for both ML and neighbor joining analyses. Representative sequences for OTUs identified from this study are in bold. Numbers in parentheses after study sequences denote number of sequences for each OTU from GN-S mats (first), and GN-I mats (last), with ratios for cDNA/DNA reads shown, respectively.

Photosynthetically-fixed carbon was indicated to be the feed-stock for fermentation. In GN-I mats, enhanced levels of bicarbonate incorporation (**Figure 3**) were reflected in significantly higher net efflux of fermentation products than in GN-S mats (**Figure 2**). Indeed, when DCMU was used to limit oxygenic photosynthesis, net fermentation dropped more than in GN-S mats, indicating oxygenic photosynthesis in *Cyanobacteria* was the most likely cause behind net fermentation productivity of these mats. At the same time, NanoSIMS measurements of bicarbonate incorporation (**Figure 4**) show increased accumulation of label in GN-I *Microcoleus* filaments over GN-S filaments. These findings point to the accumulation of glycogen being more common in GN-S mats, but increased fixation and increased catabolic metabolism being more common in GN-I mats and fits the observation that GN-I mats are adapted to exist in a dynamic turbulent intertidal zone and GN-S mats are adapted to quiescence (Bebout et al., 1994).

Pyrotag assays (**Figure 1**) corroborated previous reports of the abundance (approximately one fourth of DNA sequences from pyrotag libraries) of *Chloroflexi* in GN-S mats (Ley et al., 2006) as well as a similar level of *Chloroflexi* pyrotags seen in GN-I mats. Moreover, previous studies have shown that anoxygenic phototrophy was found to account for 10–40% of carbon fixation in GN-S mats (Finke et al., 2013), assumed to be by phototrophic sulfide oxidation. It has also been demonstrated, in the case of hot springs microbial mats that filamentous photoautotrophic *Chloroflexi* can have a role in fermentation, including hydrogenase transcript expression at night (Klatt et al., 2013), or a role indirectly driving fermentation (Otaki et al., 2012). Yet, in this study anoxygenic phototrophy appears to play only a minor role in nighttime fermentation in Guerrero Negro mats. Expression ratio (cDNA:DNA pyrotags) data demonstrated that *Cyanobacteria* (specifically genus *Microcoleus* in GN-S mat and both *Microcoleus* and genus *Lyngbya* in GN-I mat) maintain a consistent level of ribosomal expression between day and night and at a level much higher than any other phylogenetic group detected. The only hydrogenases attributable to phototrophs that were expressed at night (**Figure 5**) were associated with *Cyanobacteria*; no type 3b [NiFe]-hydrogenases from any anoxygenic phototrophic *Chloroflexi* group were recovered. However, given that marine filamentous anoxygenic phototrophs are diverse and mostly uncharacterized in Guerrero Negro mats (Nübel et al., 2001; Ley et al., 2006) the phylogeny of novel *Chloroflexi* hydrogenases present in these systems is an avenue of future study. But overall, given the dominance of several types of *Cyanobacteria* in pyrotags and hydrogenase transcripts, *Cyanobacteria* were likely the metabolically dominant phototrophic fermenters in mats. Interestingly, the HoxH tree (**Figure 5**) does suggest that a *Cyanobacteria* other than *Microcoleus chthonoplastes* PCC 7420 was the main hydrogen producer in GN-S mats and that different species of *Cyanobacteria* may differ in their capacity for hydrogen production.

Under natural conditions, both the GN mats and Elkhorn Slough mats (Burow et al., 2012) were characterized by net fluxes of hydrogen and organic acids out of the mats at night due to fermentation activity. However, in mats incubated with DCMU for the previous photoperiod, net hydrogen fluxes were reduced relative to the unamended treatments, whereas the flux of organic acids out of the mats was not significantly different (**Figures 2C,D**). This differential effect of DCMU on hydrogen vs. organic acid flux was not previously observed in Elkhorn Slough samples (Burow et al., 2012). Overall, though the DCMU in mats has been shown to inhibit photosystem II in *Cyanobacteria* and the establishment of anoxic conditions, the response of whole mat communities to this photosystem shutdown in daylight was variable across different mat types and is still poorly understood, especially with respect to the daytime sulfide cycling. Mechanisms in both *Cyanobacteria* and in *Cyanobacteria*-associated members, such as phototrophic sulfide oxidation, may be acting in concert to alter both the cycling of hydrogen as well as the original production of hydrogen.

HYDROGEN AND ORGANIC ACID AVAILABILITY LEADS TO UPTAKE BY *CYANOBACTERIA*-ASSOCIATED MICROBES IN HYPERSALINE MICROBIAL MATS

Release of cyanobacterial fermentation products within the closely packed matrix of the mat offers a flux of potential substrate to a range of terminal metabolizers. We hypothesized that SRB were the primary consumers of hydrogen and organic acids under dark, anoxic conditions, due to the abundance of sulfate. Like in hot spring microbial mats (Otaki et al., 2012) they were suspected to require close physical proximity for hydrogen uptake. In GN-I mats, inhibition with molybdate significantly increased accumulation of hydrogen at quantitatively similar levels to physical disruption (**Figure 8B**). This was consistent with previous findings in the hypersaline microbial mats of Elkhorn Slough (Burow et al., 2012), and suggests that a physical association between *Cyanobacteria* and SRB underlies most of the observable consumption of fermentation products within these mats (Burow et al., in press). Molybdate also enhanced accumulation of hydrogen in the GN-S mats, but physical disruption in those mats did not result in significantly greater net hydrogen flux relative to controls (**Figure 8D**). Thus, while SRB appear to be the dominant sink for fermentation products under dark conditions in GN-S mats, physical associations appear to be less important than in the GN-I and Elkhorn Slough mats, though there is evidence that physical proximity could still be necessary (Fike et al., 2008). However, the failure of physical disruption techniques to separate mat members apart in GN-S mats cannot be discounted, nor can the possibility of unique motile SRBs in GN-S mats be discounted.

We show here that disruption of the GN-I microbial mat, with a presumed separation of diverse members of the mat community from *Cyanobacteria*, led to a proportionally greater increase in the flux of organic acids, relative to hydrogen. This was consistent with the idea that organic acid consumption was also dependent (and possibly even more dependent than hydrogen) on tight physical association between producing and consuming organisms. Preliminary efforts were made to identify organisms that may be consuming acetate, via NanoSIMS analysis of samples incubated with ^{13}C-labeled acetate under dark/anoxic conditions. NanoSIMS analysis (**Figure 9**) confirmed that filamentous members of *Chloroflexi* and *Desulfobacteraceae* were significant consumers of acetate at night and may be

important members of these close spatial associations, though no known filamentous SRB could be identified in DsrA phylogenetic analysis. Pyrotags of *Proteobacteria* show that purple nonsulfur bacteria were also quantitatively important, particularly in the GN-I mats (**Figure 1B**). The present study did not specifically investigate acetate uptake by these organisms. However, the diverse and robust nature of the metabolism within this group (with various members being able to conduct autotrophic, heterotrophic, photosynthetic, chemotrophic, aerobic, and anaerobic metabolisms) suggests that they should also be examined for significant nighttime consumption of fermentation products in GN-I mats.

CONCLUSION

In this study, a suite of methods identified that a variety of *Cyanobacteria* were the dominant fermentive organism responsible for hydrogen production during nighttime constitutive metabolism. Furthermore, hydrogen production was driven by daytime carbon storage, and total hydrogen produced was a fraction of the total fermentation potential, with the majority of fermentation products being organic acids (especially acetate). This work also identified uptake of acetate during nighttime by both sulfate reducing bacteria and filamentous *Chloroflexi* provided an important linkage to *Cyanobacteria*. Taken together, these results indicate the nighttime fermentation of stored light energy can explain the close association of the filamentous *Chloroflexi* and of the SRB with cyanobacterial filaments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/Journal/10.3389/fmicb.2014.00061/abstract

REFERENCES

Abascal, F., Zardoya, R., and Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21, 2104–2105. doi: 10.1093/bioinformatics/bti263

Albert, D. B., and Martens, C. S. (1997). Determination of low-molecular-weight organic acid concentrations in seawater and pore-water samples via HPLC. *Mar. Chem.* 56, 27–37. doi: 10.1016/S0304-4203(96)00083-7

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Anumula, K. R. (1994). Quantitative determination of monosaccharides in glycoproteins by high-performance liquid chromatography with highly sensitive fluorescence detection. *Anal. Biochem.* 220, 275–283. doi: 10.1006/abio.1994.1338

Bahr, M., Crump, B. C., Klepac-Ceraj, V., Teske, A., Sogin, M. L., and Hobbie, J. E. (2005). Molecular characterization of sulfate-reducing bacteria in a New England salt marsh. *Environ. Microbiol.* 7, 1175–1185. doi: 10.1111/j.1462-2920.2005.00796.x

Bebout, B. M., Carpenter, S. P., Des Marais, D. J., Discipulo, M., Embaye, T., Garcia-Pichel, F., et al. (2002). Long-term manipulations of intact microbial mat communities in a greenhouse collaboratory: simulating earth's present and past field environments. *Astrobiology* 2, 383–402. doi: 10.1089/153110702762470491

Bebout, B. M., Fitzpatrick, M. W., and Paerl, H. W. (1993). Identification of the sources of energy for nitrogen fixation and physiological characterization of nitrogen-fixing members of a marine microbial mat community. *Appl. Environ. Microbiol.* 59, 1495–1503.

Bebout, B. M., Hoehler, T. M., Thamdrup, B., Albert, D., Carpenter, S. P., Hogan, M., et al. (2004). Methane production by microbial mats under low sulphate concentrations. *Geobiology* 2, 87–96. doi: 10.1111/j.1472-4677.2004.00024.x

Bebout, B. M., Paerl, H. W., Bauer, J. E., Canfield, D. E., and Marais, D. J. D. (1994). "Nitrogen cycling in microbial mat communities: the quantitative importance of N-fixation and other sources of N for primary productivity," in *Microbial Mats NATO ASI Series*, eds L. J. Stal and P. Caumette (Berlin; Heidelberg: Springer), 265–271.

Bebout, B. M., Paerl, H. W., Crocker, K. M., and Prufert, L. E. (1987). Diel interactions of oxygenic photosynthesis and N2 fixation (acetylene reduction) in a marine microbial mat community. *Appl. Environ. Microbiol.* 53, 2353–2362.

Bigge, J. C., Patel, T. P., Bruce, J. A., Goulding, P. N., Charles, S. M., and Parekh, R. B. (1995). Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. *Anal. Biochem.* 230, 229–238. doi: 10.1006/abio.1995.1468

Burow, L. C., Woebken, D., Bebout, B. M., McMurdie, P. J., Singer, S. W., Pett-Ridge, J., et al. (2012). Hydrogen production in photosynthetic microbial mats in the Elkhorn Slough estuary, Monterey Bay. *ISME J.* 6, 863–874. doi: 10.1038/ismej.2011.142

Burow, L. C., Woebken, D., Marshall, I. L., Singer, S. W., Pett-Ridge, J., Prufert-Bebout, L., et al. (in press). Identification of *Desulfobacterales* as primary hydrogenotrophs in a complex microbial mat community. *Geobiology*.

Burow, L. C., Woebken, D., Marshall, I. P., Lindquist, E. A., Bebout, B. M., Prufert-Bebout, L., et al. (2013). Anoxic carbon flux in photosynthetic microbial mats as revealed by metatranscriptomics. *ISME J.* 7, 817–829. doi: 10.1038/ismej.2012.150

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Meth.* 7, 335–336. doi: 10.1038/nmeth.f.303

Castro, H., Reddy, K. R., and Ogram, A. (2002). Composition and function of sulfate-reducing prokaryotes in eutrophic and pristine areas of the florida everglades. *Appl. Environ. Microbiol.* 68, 6129–6137. doi: 10.1128/AEM.68.12.6129-6137.2002

Cline, J. D. (1969). Spectrophotometric determination of hydrogen sulfide in natural waters. *Limnol. Oceanogr.* 14, 454–458. doi: 10.4319/lo.1969.14.3.0454

Des Marais, D. J. (1995). The biogeochemistry of hypersaline microbial mats. *Adv. Microb. Ecol.* 14, 251–274. doi: 10.1007/978-1-4684-7724-5_6

Des Marais, D. J. (2003). Biogeochemistry of hypersaline microbial mats illustrates the dynamics of modern microbial ecosystems and the early evolution of the biosphere. *Biol. Bull.* 204, 160–167. doi: 10.2307/1543552

Dhillon, A., Teske, A., Dillon, J., Stahl, D. A., and Sogin, M. L. (2003). Molecular characterization of sulfate-reducing bacteria in the guaymas basin. *Appl. Environ. Microbiol.* 69, 2765–2772. doi: 10.1128/AEM.69.5.2765-2772.2003

Engelbrektson, A., Kunin, V., Wrighton, K. C., Zvenigorodsky, N., Chen, F., Ochman, H., et al. (2010). Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J.* 4, 642–647. doi: 10.1038/ismej.2009.153

Feazel, L. M., Spear, J. R., Berger, A. B., Harris, J. K., Frank, D. N., Ley, R. E., et al. (2008). Eucaryotic diversity in a hypersaline microbial mat. *Appl. Environ. Microbiol.* 74, 329–332. doi: 10.1128/AEM.01448-07

Fenchel, T., and Finlay, B. J. (1995). *Ecology and Evolution in Anoxic Worlds*. Oxford; New York, NY: Oxford University Press.

Fike, D. A., Gammon, C. L., Ziebis, W., and Orphan, V. J. (2008). Micronscale mapping of sulfur cycling across the oxycline of a cyanobacterial mat: a paired nanoSIMS and CARD-FISH approach. *ISME J.* 2, 749–759. doi: 10.1038/ismej.2008.39

Finke, N., Hoehler, T. M., Polerecky, L., Buehring, B., and Thamdrup, B. (2013). Competition for inorganic carbon between oxygenic and anoxygenic phototrophs in a hypersaline microbial mat, Guerrero Negro, Mexico. *Environ. Microbiol.* 15, 1532–1550. doi: 10.1111/1462-2920.12032

Gihring, T. M., Green, S. J., and Schadt, C. W. (2012). Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. *Environ. Microbiol.* 14, 285–290. doi: 10.1111/j.1462-2920.2011.02550.x

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Hall, P. O. J., and Aller, R. C. (1992). Rapid, small-volume, flow injection analysis for SCO2, and NH4+ in marine and freshwaters. *Limnol. Oceanogr.* 37, 1113–1119. doi: 10.4319/lo.1992.37.5.1113

Hamady, M., Walker, J. J., Harris, J. K., Gold, N. J., and Knight, R. (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat. Meth.* 5, 235–237. doi: 10.1038/nmeth.1184

Harrison, B. K., Zhang, H., Berelson, W., and Orphan, V. J. (2009). Variations in archaeal and bacterial diversity associated with the sulfate-methane transition zone in continental margin sediments (Santa Barbara Basin, California). *Appl. Environ. Microbiol.* 75, 1487–1499. doi: 10.1128/AEM.01812-08

Hoehler, T. (2005). "Biogeochemistry of dihydrogen (H2)," in *Metal Ions in Biological Systems, Volume 43 - Biogeochemical Cycles of Elements*, eds A. Sigel, H. Sigel, and R. Sigel (Boca Raton: CRC Press), 9–48.

Hoehler, T. M., Bebout, B. M., and Des Marais, D. J. (2001). The role of microbial mats in the production of reduced gases on the early Earth. *Nature* 412, 324–327. doi: 10.1038/35085554

Ihaka, R., and Gentleman, R. (1996). R: a language for data analysis and graphics. *J. Comput. Grap. Stat.* 5, 299–314. doi: 10.1080/10618600.1996.10474713

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275

Kaneko, R., Hayashi, T., Tanahashi, M., and Naganuma, T. (2007). Phylogenetic diversity and distribution of dissimilatory sulfite reductase genes from deep-sea sediment cores. *Mar. Biotechnol.* 9, 429–436. doi: 10.1007/s10126-007-9003-7

Keppler, D., and Decker, K. (1974). "Glycogen determination with amyloglucosidase," in *Methods of Enzymatic Analysis*, ed H. U. Bergmeyer (Weinheim: Verlag Chemie), 1127–1131.

Klatt, C. G., Liu, Z., Ludwig, M., Kühl, M., Jensen, S. I., Bryant, D. A., et al. (2013). Temporal metatranscriptomic patterning in phototrophic Chloroflexi inhabiting a microbial mat in a geothermal spring. *ISME J.* 7, 1775–1789. doi: 10.1038/ismej.2013.52

Kuczynski, J., Liu, Z., Lozupone, C., McDonald, D., Fierer, N., and Knight, R. (2010). Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. *Nat. Meth.* 7, 813–819. doi: 10.1038/nmeth.1499

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320. doi: 10.1093/molbev/msn067

Leloup, J., Fossing, H., Kohls, K., Holmkvist, L., Borowski, C., and Jørgensen, B. B. (2009). Sulfate-reducing bacteria in marine sediment (Aarhus Bay, Denmark): abundance and diversity related to geochemical zonation. *Environ. Microbiol.* 11, 1278–1291. doi: 10.1111/j.1462-2920.2008.01855.x

Ley, R. E., Harris, J. K., Wilcox, J., Spear, J. R., Miller, S. R., Bebout, B. M., et al. (2006). Unexpected diversity and complexity of the guerrero negro hypersaline microbial mat. *Appl. Environ. Microbiol.* 72, 3685–3695. doi: 10.1128/AEM.72.5.3685-3695.2006

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158

Mussmann, M., Richter, M., Lombardot, T., Meyerdierks, A., Kuever, J., Kube, M., et al. (2005). Clustered genes related to sulfate respiration in uncultured prokaryotes support the theory of their concomitant horizontal transfer. *J. Bacteriol.* 187, 7126–7137. doi: 10.1128/JB.187.20.7126-7137.2005

Nübel, U., Bateson, M. M., Madigan, M. T., Kühl, M., and Ward, D. M. (2001). Diversity and distribution in hypersaline microbial mats of bacteria related to chloroflexusspp. *Appl. Environ. Microbiol.* 67, 4365–4371. doi: 10.1128/AEM.67.9.4365-4371.2001

Nübel, U., Garcia-Pichel, F., Kühl, M., and Muyzer, G. (1999). Quantifying microbial diversity: morphotypes, 16S rRNA genes, and carotenoids of oxygenic phototrophs in microbial mats. *Appl. Environ. Microbiol.* 65, 422–430.

Oettmeier, W. (1992). "Herbicides of photosystem II," in *Topics in Photosynthesis*, ed J. Barber (Amsterdam: Elsevier Science Publishers), 295–348.

Omoregie, E. O., Crumbliss, L. L., Bebout, B. M., and Zehr, J. P. (2004a). Comparison of diazotroph community structure in Lyngbya sp. and Microcoleus chthonoplastes dominated microbial mats from Guerrero Negro, Baja, Mexico. *FEMS Microbiol. Ecol.* 47, 305–308. doi: 10.1016/S0168-6496(03)00301-5

Omoregie, E. O., Crumbliss, L. L., Bebout, B. M., and Zehr, J. P. (2004b). Determination of nitrogen-fixing phylotypes in lyngbya sp. and microcoleus chthonoplastes cyanobacterial mats from Guerrero Negro, Baja California, Mexico. *Appl. Environ. Microbiol.* 70, 2119–2128. doi: 10.1128/AEM.70.4.2119-2128.2004

Ondov, B. D., Bergman, N. H., and Phillippy, A. M. (2011). Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* 12:385. doi: 10.1186/1471-2105-12-385

Oremland, R., and Capone, D. (1988). Use of specific inhibitors in biogeochemistry and microbial ecology. *Adv. Microb. Ecol.* 10, 285–383. doi: 10.1007/978-1-4684-5409-3_8

Otaki, H., Everroad, R. C., Matsuura, K., and Haruta, S. (2012). Production and consumption of hydrogen in hot spring microbial mats dominated by a filamentous anoxygenic photosynthetic bacterium. *Microbes Environ.* 27, 293–299. doi: 10.1264/jsme2.ME11348

Pett-Ridge, J., and Weber, P. K. (2012). "NanoSIP: NanoSIMS applications for microbial biology," in *Microbial Systems Biology Methods in Molecular Biology*, ed A. Navid (New York, NY: Humana Press), 375–408.

Reeder, J., and Knight, R. (2010). Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nat. Meth.* 7, 668–669. doi: 10.1038/nmeth0910-668b

Robertson, C. E., Spear, J. R., Harris, J. K., and Pace, N. R. (2009). Diversity and stratification of archaea in a hypersaline microbial mat. *Appl. Environ. Microbiol.* 75, 1801–1810. doi: 10.1128/AEM.01811-08

Rothrock, M. J., and Garcia-Pichel, F. (2005). Microbial diversity of benthic mats along a tidal desiccation gradient. *Environ. Microbiol.* 7, 593–601. doi: 10.1111/j.1462-2920.2005.00728.x

Skyring, G. W., Lynch, R. M., and Smith, G. D. (1989). "Quantitative relationships between carbon, hydrogen, and sulfur metabolism in cyanobacterial mats," in *Microbial Mats: Physiological Ecology of Benthic Microbial Communities*, eds Y. Cohen and E. Rosenberg (Washington, DC: American Society for Microbiology), 170–179.

Spear, J. R., Ley, R. E., Berger, A. B., and Pace, N. R. (2003). Complexity in natural microbial ecosystems: the guerrero negro experience. *Biol. Bull.* 204, 168–173. doi: 10.2307/1543553

Steppe, T. F., and Paerl, H. W. (2002). Potential N2 fixation by sulfate-reducing bacteria in a marine intertidal microbial mat. *Aquat. Microb. Ecol.* 28, 1–12. doi: 10.3354/ame028001

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Vignais, P. M., and Billoud, B. (2007). Occurrence, classification, and biological function of hydrogenases: an overview. *Chem. Rev.* 107, 4206–4272. doi: 10.1021/cr050196r

Wagner, M., Roger, A. J., Flax, J. L., Brusseau, G. A., and Stahl, D. A. (1998). Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration. *J. Bacteriol.* 180, 2975–2982.

Woebken, D., Burow, L. C., Prufert-Bebout, L., Bebout, B. M., Hoehler, T. M., Pett-Ridge, J., et al. (2012). Identification of a novel cyanobacterial group as active diazotrophs in a coastal microbial mat using NanoSIMS analysis. *ISME J.* 6, 1427–1439. doi: 10.1038/ismej.2011.200

Zhang, W., Song, L., Ki, J.-S., Lau, C.-K., Li, X.-D., and Qian, P.-Y. (2008). Microbial diversity in polluted harbor sediments II: sulfate-reducing bacterial community assessment using terminal restriction fragment length polymorphism and clone library of dsrAB gene. *Estuar. Coast. Shelf Sci.* 76, 682–691. doi: 10.1016/j.ecss.2007.07.039

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Spatial patterns and links between microbial community composition and function in cyanobacterial mats

Mohammad A. A. Al-Najjar[1,2]*, Alban Ramette[3], Michael Kühl[4,5], Waleed Hamza[6], Judith M. Klatt[1] and Lubos Polerecky[1,7]*

[1] Microsensor Group, Max-Planck Institute for Marine Microbiology, Bremen, Germany
[2] Marine Microbial Ecology Group, Red Sea Research Center, KAUST, Thuwal, Saudi Arabia
[3] Microbial Habitat Group, Max-Planck Institute for Marine Microbiology, Bremen, Germany
[4] Marine Biological Section, Department of Biology, University of Copenhagen, Helsingør, Denmark
[5] Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, Sydney, NSW, Australia
[6] Biology Department, UAE University, Al-Ain, UAE
[7] Department of Earth Sciences – Geochemistry, Utrecht University, Utrecht, Netherlands

We imaged reflectance and variable fluorescence in 25 cyanobacterial mats from four distant sites around the globe to assess, at different scales of resolution, spatial variabilities in the physiological parameters characterizing their photosynthetic capacity, including the absorptivity by chlorophyll a (A_{chl}), maximum quantum yield of photosynthesis (Y_{max}), and light acclimation irradiance (I_k). Generally, these parameters significantly varied within individual mats on a sub-millimeter scale, with about 2-fold higher variability in the vertical than in the horizontal direction. The average vertical profiles of Y_{max} and I_k decreased with depth in the mat, while A_{chl} exhibited a sub-surface maximum. The within-mat variability was comparable to, but often larger than, the between-sites variability, whereas the within-site variabilities (i.e., between samples from the same site) were generally lowest. When compared based on averaged values of their photosynthetic parameters, mats clustered according to their site of origin. Similar clustering was found when the community composition of the mats' cyanobacterial layers were compared by automated ribosomal intergenic spacer analysis (ARISA), indicating a significant link between the microbial community composition and function. Although this link is likely the result of community adaptation to the prevailing site-specific environmental conditions, our present data is insufficient to identify the main factors determining these patterns. Nevertheless, this study demonstrates that the spatial variability in the photosynthetic capacity and light acclimation of benthic phototrophic microbial communities is at least as large on a sub-millimeter scale as it is on a global scale, and suggests that this pattern of variability scaling is similar for the microbial community composition.

Keywords: spatial link between structure and function, photosynthetic microbial mats, imaging PAM, biogeography, hyperspectral imaging, microbial community structure

INTRODUCTION

A major challenge in microbial ecology is to understand how the structure, composition, and function of microbial communities are linked, how microbial communities are influenced by environmental conditions, and how they contribute to the local and global cycling of elements. Examples of microbial communities that have been extensively studied from all of these perspectives are cyanobacterial mats, which are highly compacted microbial ecosystems consisting of diverse phototrophic and heterotrophic populations (van Gemerden, 1993; Stal, 2000). The interest in their study stems from the generally accepted assumption that they represent a modern analog of the earliest complex ecosystems on Earth (Seckbach and Oren, 2010). Additionally, their vertically stratified and compact structure offers good possibilities for studying microbial and biogeochemical interactions in well-controlled laboratory mesocosms.

Environments that harbor cyanobacterial mats vary greatly with respect to parameters such as nutrient concentrations, temperature, salinity, or input of mineral particles. This variability leads to differences in the structure of the mats with respect to the density and distribution of microbial cells, content of inorganic particles, and characteristics of exopolymers that bind the mat matrix together. In spite of these differences, one feature that these environments have in common is their extremity with respect to at least one of the environmental parameters, typically salinity, temperature, or pH (Caumette et al., 1994; Garcia-Pichel et al., 1999; Des Marais, 2003; Abed et al., 2006; Seckbach and Oren, 2010). This environmental constraint is essential for excluding, or at least minimizing, the influence of grazing and bioturbation, which would otherwise disturb the perennial growth and laminated structure of the mats.

A typical feature of cyanobacterial mats are steep vertical gradients of physical and chemical parameters such as light, O_2, pH, and H_2S (van Gemerden, 1993; Stal, 2000). While the gradients in light intensity and spectral composition are the consequence of strong and wavelength-dependent absorption and scattering by photosynthetically active (photopigments) and inactive (mineral particles, organic detritus) components in the mat (Kühl and Jørgensen, 1992, 1994; Kühl et al., 1994), the steep chemical gradients form due to mass transfer limitation in a volume that is densely packed with active microbial cells (Kühl et al., 1996; Wieland and Kühl, 2000a; Jonkers et al., 2003; Garcia de Lomas et al., 2005).

A critical process in cyanobacterial mats is the photosynthetic activity of the cyanobacterial population, which occurs in the uppermost layer of the mat (so-called euphotic zone) and supports diverse heterotrophic populations in the mat ecosystem through the production of organic substrates and O_2 (Nübel et al., 1999; Roeselers et al., 2007). Cyanobacterial photosynthesis depends on a number of environmental parameters, including temperature, salinity, concentration of nutrients, intensity and spectral quality of light, and exposure to H_2S (Kühl, 1993, 2005; Kühl and Fenchel, 2000; Wieland et al., 2003; Pinckney et al., 2011). Since these parameters vary between different sites that harbor cyanobacterial mats as well as within the mats themselves (see above), it is expected that the rate and efficiency of photosynthesis in the mats will exhibit strong geographical as well as micrometer-scale variability. Although previous measurements demonstrated that photosynthetic rates and efficiency are strongly variable within mats (Kühl et al., 1994; Al-Najjar et al., 2010), presently it is not known how this within-mat variability compares to the variability between mats from different environments. Additionally, it is not known whether there is a link between the photosynthetic capacity in the euphotic zone of the mats and the composition of the corresponding microbial community, and to which extent are these properties of the mats determined by the parameters characterizing their environment.

To address these issues, we compared 25 samples of cyanobacterial mats collected from four distant geographical locations (United Arab Emirates, Australia, Brazil, and Spain) with respect to their photosynthetic capacity and microbial community composition. Our focus was on the cyanobacterial layer at the top of the mats, which was assumed to be a good proxy for the mats' photosynthetically active zone. We used absorptivity by chlorophyll a (A_{chl}), maximum quantum yield of photosystem II (Y_{max}), and light acclimation irradiance (I_k) as parameters characterizing the photosynthetic capacity and adaptation of the cyanobacterial populations. These parameters were measured with a sub-millimeter spatial resolution across vertical sections of the cyanobacterial layers using hyper-spectral and variable chlorophyll fluorescence imaging. Differences in the microbial community composition of the cyanobacterial layers were quantified by automated ribosomal intergenic spacer analysis (ARISA). Possible links between the parameters characterizing the photosynthetic capacity, microbial community composition, and environmental parameters were identified using multivariate statistical methods. We hypothesized that the microbial community composition and photosynthetic capacity in the cyanobacterial layers are linked, and that the photosynthetic potential and light acclimation of the cyanobacterial populations are more strongly influenced by the steep vertical gradients within the mat ecosystem than by the environmental parameters characterizing their habitat.

MATERIALS AND METHODS

SAMPLES

The studied cyanobacterial mats originated from four sites: an intertidal flat near Abu-Dhabi, UAE (AD mats), an intertidal flat in the Exmouth Gulf in Australia (AU mats; Lovelock et al., 2010; Adame et al., 2012), the hypersaline lake Lagoa Vermelha in Brazil (BR mats ; Vasconcelos et al., 2006), and the hypersaline lake La Salada de Chiprana in Spain (SP mats; Jonkers et al., 2003). The mat samples were collected between 2003 and 2008 in at least two replicates from each site and incubated under artificial illumination (10 h light/14 h dark cycles, wavelength range 400–700 nm) and at approximately constant temperature until the measurements, which were conducted in 2009. More details about the collection sites and incubation conditions are given in **Table 1**. During the incubation, the appearance of the AD, AU, and BR mats did not change, whereas the SP mats gradually changed from thinly laminated structures (see Jonkers et al., 2003) to thicker structures featuring a loosen upper layer composed of a mixture of exopolymers and suspended particles.

MEASUREMENT PROTOCOL

First, the mat samples were pre-incubated for about 12 h at room temperature and incident irradiance of 100 μmol photons m^{-2} s^{-1}. Subsequently, they were vertically sectioned and immediately afterwards variable chlorophyll fluorescence and spectral reflectance were measured to characterize, respectively, the photosynthetic potential and pigments in the mats. This was done with a high spatial resolution ($\sim$20 μm) across the vertical sections of the mats using imaging cameras (see below). Immediately after imaging, cyanobacteria-dominated layers close to the mats surface were cut-off with a sterile scalpel and prepared for the ARISA (see below). Identification of these layers was based on their characteristic dark-green appearance, microscopic observations and the results of the imaging analyses, which were rapidly obtained by the image processing routines developed during this study (see below).

PAM IMAGING OF THE VARIABLE CHLOROPHYLL FLUORESCENCE

Pulse amplitude modulated (PAM) imaging of the variable chlorophyll fluorescence was done with the Imaging-PAM system (Walz GmbH, Germany), using red light-emitting diodes for the excitation of the chlorophyll a fluorescence from cyanobacteria. For each mat, a vertical section of the mat sample was placed on its side in a Petri dish and covered with a few millimeters of seawater (salinity of 32, temperature 15°C). After 15 min of dark adaptation, images of the minimum (F_o) and maximum (F_m) fluorescence yields in the dark-adapted state were recorded. Subsequently, rapid light curves (RLC) (Schreiber et al., 1996, 1997) were measured by increasing the actinic irradiance from 0 to 1700 μmol photons m^{-2} s^{-1} in time intervals of 3 min for each irradiance and acquiring the fluorescence yield images under

Table 1 | Characteristics of the sampling sites, collection details, and incubation conditions for the studied microbial mats.

Site abbreviation	AD	AU	SP	BR
ENVIRONMENTAL PARAMETERS[a]				
Type of site	Intertidal flat	Intertidal flat	Hypersaline lake	Lagoon (mixed with seawater)
Salinity range (‰)	35–200	35–50	78–90	40–150
Temperature (°C)	25–55	24–28	15–28	25–29
Inundation period	Few hours everyday	For ∼6 h per month	Continuous	Continuous
Water depth when inundated (m)	1	2	1.5	1–2
Type of substrate	Very fine grained sediment	Very fine grained sediment	Very fine grained sediment	Organic material
N (mg/l)	2	840	1.2	NA[b]
P (mg/l)	0.03	96	0.04	0.23
SAMPLE COLLECTION				
Year	2007	2008	2003	2008
Geographic location	24°31′20″N, 54°26′50″E	22°43′7″S, 114°34′6″E	41°14′30″N, 0°10′50″W	22°53′15.2″S, 42°6′38.14″W
Replicates	13	8	2	2
Within-site replicates separation (m)	∼50	∼10	∼1[c]	∼10
INCUBATION CONDITIONS				
Salinity range (‰)	35–200	35–50	80–90	35–200
T (°C):	26–28	26–28	20	26–28
Downwelling irradiance[d] (µmol photons m^{-2} s^{-1}):	380[e]	120–180[f]	240[g]	380[e]

[a] Information extracted from literature: Al-Najjar et al. (2010) for AD mats; Lovelock et al. (2010) for AU mats; Jonkers et al. (2003) for SP mats; Vasconcelos et al. (2006) for BR mats.

[b] Value not measured.

[c] Taken from an aquarium where the mats were grown for ∼6 years.

[d] Quantified by a calibrated PAR quantum irradiance sensor (LI-190 Quantum) connected to a light meter (LI-250, both from LI-COR Biosciences).

[e] Light source: AQUALINE 10000, MH 400W, Germany.

[f] Light source: Envirolite, UK.

[g] Light source: cool white fluorescent tubes T8 (32W), Philips, Germany.

actinic illumination (F') and during the saturating pulse (F_m') at the end of each interval. The intensity and duration of the saturating pulse was ∼2400 µmol photons m^{-2} s^{-1} and 0.8 s, respectively, and the irradiance levels of the actinic light were calibrated using a PAR (400–700 nm) quantum irradiance sensor (LI-190 Quantum) connected to a light meter (LI-250, both from LI-COR Biosciences) positioned in the same place as the mat sample. As the sample was relatively small and it was lying on its side, irradiance was evenly distributed across the vertical section of the mat.

IMAGING OF CHLOROPHYLL A ABSORPTIVITY

In addition to variable fluorescence imaging, the Imaging-PAM system was used to image reflectance of the mats in the red (R_r) and near-infrared (R_{nir}) region. These images were used to calculate the chlorophyll a absorptivity as $A_{chl} = (R_{nir} - R_r)/R_{nir}$, which was taken as a proxy for chlorophyll a concentration in the mats. Because the spectral resolution of these measurements was insufficient, chlorophyll a absorptivity was additionally quantified by hyper-spectral imaging (Kühl and Polerecky, 2008; Polerecky et al., 2009). Specifically, each mat sample in the Petri dish was placed on a motorized stage, illuminated with a halogen bulb (Philips, type 6423) emitting in the visible to near-infrared

range (400–900 nm), and scanned with a hyper-spectral imaging system (VNIR-100, Themis, Themis Vision, USA). Spectral normalization was achieved by scanning a gray reference standard with 40% reflectance (SRS-40-020, Labsphere Inc., USA). After verification that the reflectance spectrum had a characteristic chl a edge in the wavelength range of 700–720 nm and a flat plateau above 720 nm (see, e.g., Polerecky et al., 2009), the image of chl a absorptivity was calculated as $A_{chl} = (R_{750} - R_{675})/R_{750}$, where R_{750} and R_{675} are the reflectance images measured at 750 nm (no absorption by chl a) and 675 nm (maximal chl a absorption), respectively. No difference was found between the chl a absorptivity values obtained by hyper-spectral imaging and by the Imaging-PAM system (data not shown). Therefore, the latter were used in the subsequent analysis, as they allowed perfect alignment with the variable fluorescence images.

DNA EXTRACTION AND ARISA

DNA from the cyanobacterial layers in the studied mats was extracted and purified using the UltraClean soil DNA isolation kit (MO BIO Laboratories, Inc., Carlsbad, CA, USA) according to the manufacturer's instructions. For each mat sample, PCR (50 µl) were conducted in triplicates and contained 1× PCR

buffer (Promega, Madison, WI, USA), 2.5 mM MgCl$_2$ (Promega), 0.25 mM of 40 mM dNTP mix (Promega), bovine serum albumin (3 µg/µl, final concentration), 25 ng extracted DNA, 400 nM each of universal primer ITSF (5′-GTCGTAACAAGGTAGCCGTA-3′) and eubacterial ITSReub (5′-GCCAAGGCATCCACC-3′; Cardinale et al., 2004) labeled with the phosphoramidite dye HEX, and 0.05 units GoTaq polymerase (Promega). All subsequent steps, including the PCR protocol, purity of the PCR products, labeling of the products, discrimination of the PCR-amplified fragments via capillary electrophoresis (ABI PRISM 3130*xl* Genetic Analyzer, Applied Biosystems) and the subsequent statistical analysis of ARISA profiles (quality control, binning, merging) were done as previously described (Boer et al., 2009; Ramette, 2009).

PROCESSING AND ANALYSIS OF THE FLUORESCENCE YIELD IMAGES

Using the fluorescence yield images, the quantum yield of PSII in the dark-adapted state was calculated as $Y_0 = (F_m − F_o)/F_m$ and the effective quantum yield of PSII at a given actinic irradiance, $I > 0$, as $Y = (F_m′ − F′)/F_m′$. Subsequently, the values of Y were plotted as a function of I to determine the maximum effective quantum yield of PSII (denoted as Y_{max}) and the actinic irradiance at which Y reached Y_{max} (denoted as I_{max}). Based on the theoretical background of the saturation pulse method (Baker, 2008), Y_{max} represents the maximal photosynthetic potential of the cyanobacterial population in the mat. In plants, Y decreases monotonously with I and this maximal potential is reached in the dark-adapted state, i.e., $Y_{max} = Y_0$ and $I_{max} = 0$ (White and Critchley, 1999). However, in our measurements with cyanobacterial mats the Y-I relationship was not monotonous and these parameters were typically related as $Y_{max} > Y_0$ and $I_{max} > 0$ (**Figure 1**), which is why we additionally determined also I_{max}.

In photosynthesis research, light acclimation is determined from the relationship between the rate of photosynthesis, P, and irradiance, I. This so-called P-I curve is close to linear at low irradiance levels and approaches saturation at high irradiance levels. The actual value of the light acclimation irradiance (denoted here as I_k) depends on the mathematical model that describes it, and

is typically obtained by fitting the P-I data (see, e.g., Platt and Jassby, 1976). To obtain I_k from our data, we assumed that the parameter Y measured by the saturation pulse method is given by $Y = \alpha P/I$, where α is a proportionality constant whose value is not important in this study (but see Campbell et al., 1998), and considered the following model to describe the P-I relationship:

$$P(I) = P_{max}[1 − \exp(− I/I_k)] − \delta Y/\alpha\, I \exp(− I/I_m). \quad (1)$$

The first term in this model is equivalent to that proposed by Webb et al. (1974), whereas the second term was necessary to describe the non-monotonous relationship between Y and I observed in this study (**Figure 1**). Using Equation (1) and assuming that $I_m \ll I_k$, which was the case in our measurements, Y can be written approximately as

$$Y \approx Y_m(I_k/I)[1 − \exp(− I/I_k)] − \delta Y \exp(− I/I_{max}), \quad (2)$$

where $Y_m = Y_{max}/(1 − I_{max}/I_k)$, $\delta Y = Y_m − Y_0$ and $I_{max} \approx I_m$ if $I_m \ll I_k$ (see **Figure 1B**). Thus, by determining the values of Y_0, Y_{max}, and I_{max} and fitting the rest of the measured Y-I values with the model in Equation (2), it was possible to determine the remaining fitting parameter I_k. Data processing required for this analysis, including quantification of the images of Y_{max}, I_{max}, and I_k, their variabilities across the vertical and horizontal directions as well as average vertical and horizontal profiles, was done in Matlab (The MathWorks Inc., Natick, MA) using the program Look@PAM developed during this study. This program is available on the internet (http://www.microsen-wiki.net/pamimaging:lookatpam).

STATISTICAL ANALYSES

The significance of univariate response data as a function of categorical factors was tested using one-way analysis of variance (ANOVA), after verifying the normality (Shapiro–Wilk normality test) of the response variable at $p = 0.05$. Community differences were visualized by non-metric multidimensional scaling (MDS) ordination based on the Bray–Curtis dissimilarity
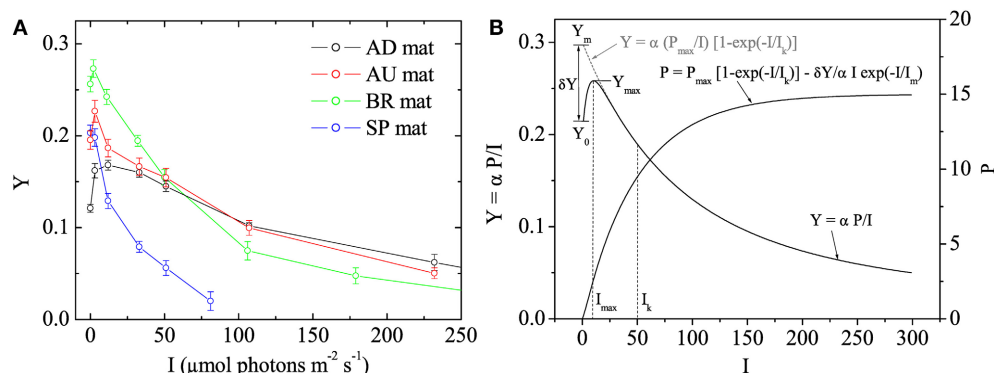
FIGURE 1 | (A) Examples of quantum yields of photosystem II, *Y*, in the studied cyanobacterial mats, as measured by pulse amplitude modulated (PAM) imaging at different irradiances, *I*. Symbols and error bars represent, respectively, the mean and SD calculated from 5 × 5

pixels in the image. **(B)** Mathematical model describing the *Y* vs. *I* relationship. The main features of the relationship are annotated. The corresponding relationship between photosynthesis, *P*, and *I* is also shown.

matrix between samples, and significance of community differences between groups of samples was determined by Analysis of Similarity (ANOSIM) tests. Prior to analyzing ARISA profiles conjointly with functional or environmental variables, a consensus community profile was obtained for each sample by merging the triplicate ARISA PCR and by considering an OTU present if it appeared at least twice among the triplicates (Ramette, 2009). The merged table was Hellinger-transformed to minimize the effects of the strongly right-skewed distribution curve (Ramette, 2007). To assess the link between the microbial community composition and function, the Procrustes superimposition approach was used to estimate the concordance of scores originating from two independent ordinations after rotating, translating, and dilating one of them, while keeping the other ordination coordinates constant (Gower, 1975). Significance of the rotation statistic was assessed by Monte-Carlo permutations (Peres-Neto and Jackson, 2001). All statistical tests were carried out with the statistical platform *R* (http://cran.r-project.org/) and multivariate community analyses with the *vegan* package (Oksanen et al., 2012).

RESULTS

All studied mats had a clear laminated structure, with a characteristic dark-green layer at or close to the mat surface in each of them (see examples of true color images in **Figure 2**). The AD and AU mats had a brown layer and a red layer underneath the dark-green layer, while the deepest layer was black. Additionally, the AD mats were covered by a thin orange gelatinous layer. The BR mats showed a different layering pattern with an upper dark-green layer followed by a thicker dark-pink gelatinous layer. The

SP mats had a similar structure as the AD mats, except the distinct layers were thicker and the surface gelatinous layer had whitish to light-brown appearance.

Hyperspectral imaging revealed that the dark-green layer in the mats had a pronounced absorption at wavelengths corresponding to the maximal absorption by chlorophyll *a* (675 nm) and phycocyanin (625 nm; Figure S1). Since these pigments are characteristic for cyanobacteria, we could be confident that the dark-green layer contained abundant cyanobacterial populations and could therefore be referred to as the cyanobacterial layer. This conclusion was supported by microscopic observations (data not shown).

VARIABILITY OF PHYSIOLOGICAL PARAMETERS ON DIFFERENT SPATIAL SCALES

Imaging of A_{chl}, Y_{max}, I_{max}, and I_k in mats collected from different sites made it possible to investigate how the variability in these physiological parameters changes depending on the scale at which it is determined, including the micrometer-scale (within-mat variability), meter-scale (within-site variability), and global-scale (between-site variability). The within-mat variability was quantified in three ways: as standard deviation (SD) of the values from the entire image of the cyanobacterial layer and as SD of the average vertical and horizontal profiles. The within-site variability was calculated as SD of the average values over the cyanobacterial layer in each mat from the respective site, while the global variability as SD of the average values for each site.

As demonstrated by the images, all studied physiological parameters exhibited pronounced micrometer-scale variability

FIGURE 2 | Example images of the AD, AU, BR, and SP mats obtained by hyperspectral imaging (column "true color"), reflectance imaging (A_{chl}), and PAM imaging (Y_{max}, I_k, I_{max}). Scale bar is 1 mm.
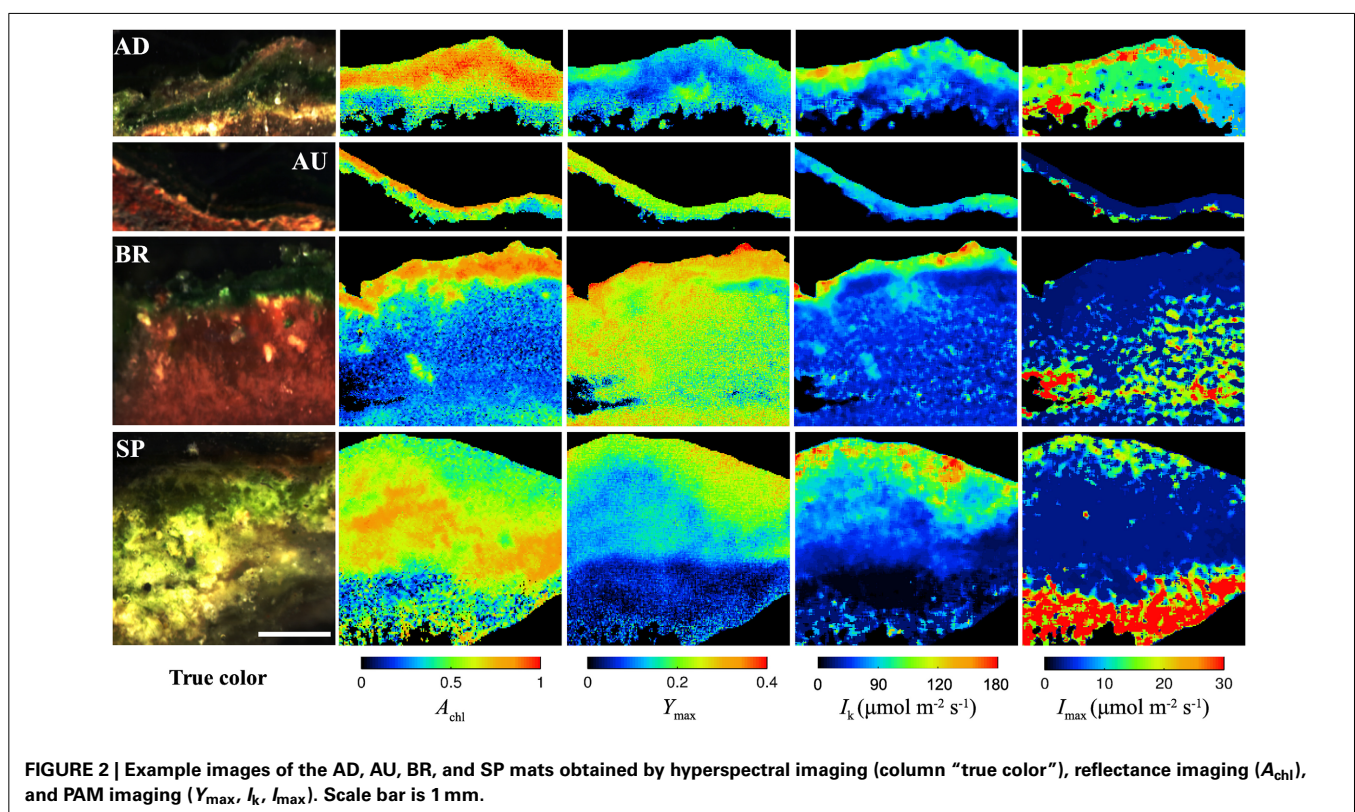
within the cyanobacterial layer of the studied mats (**Figure 2**). Two-factorial analysis of variance performed on individual images of A_{chl}, Y_{max}, and I_k, using vertical and horizontal position in the mat as factors, revealed that the percentage of the total within-mat variance explained by the vertical position (25–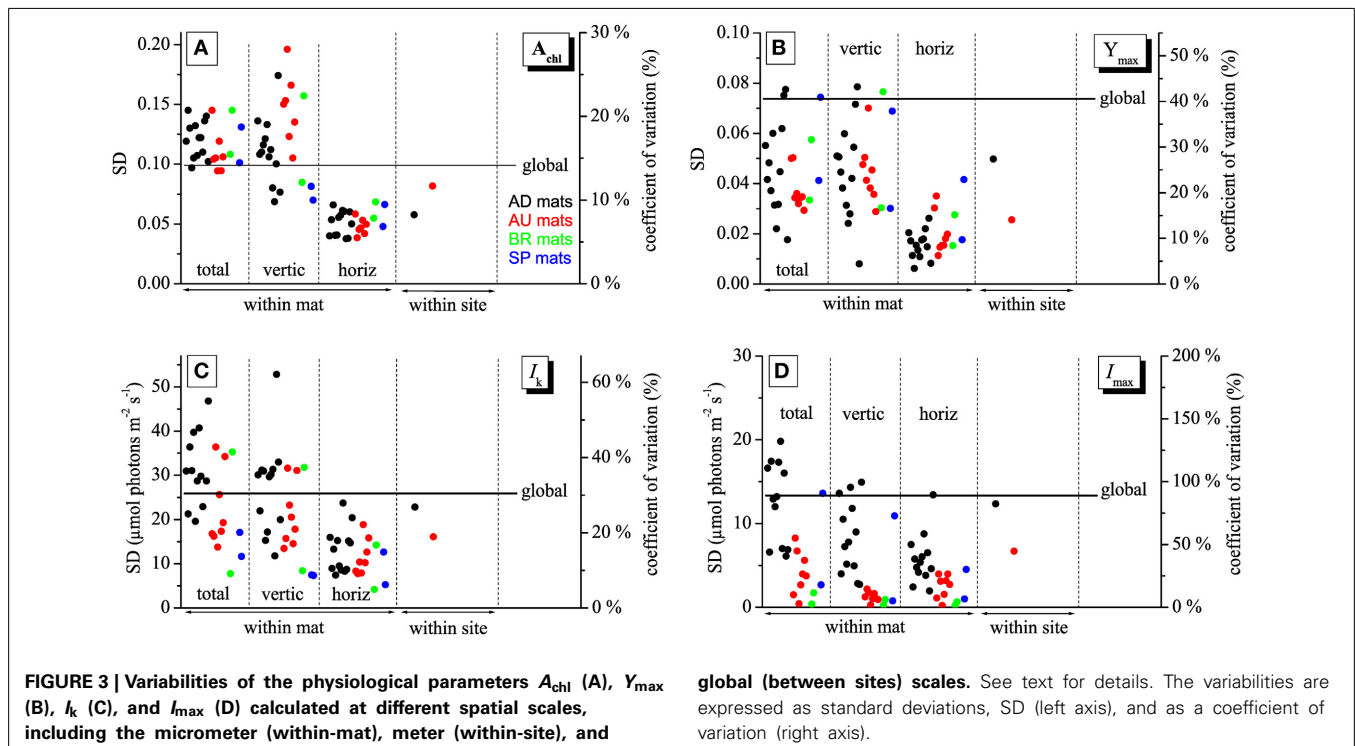30% for A_{chl}, 45–50% for Y_{max}, and 30–35% for I_k) was about 1.5 to 2-fold larger than the percentage of variance explained by the horizontal position (15–20% for A_{chl}, 15–30% for Y_{max}, 15–25% for I_k). This was consistent with the comparison of the SD for the average vertical and horizontal profiles, which showed that the former were about 2-fold larger than the latter (**Figures 3A–C**). Thus, on average, the vertical variability in A_{chl}, Y_{max}, and I_k was about twice as high as the horizontal one. A significant portion of the within-mat variance was explained by the interaction between the vertical and horizontal position (55–60% for A_{chl}, 30–35% for Y_{max}, and 45–50% for I_k), consistent with the clearly visible variation of the vertical profiles of these parameters along the horizontal direction for each individual mat sample (**Figure 2**). With respect to I_{max}, the vertical and horizontal variabilities were similar (**Figure 3D**), each explaining about 10% of the total within-mat variability, while the remaining 80% was explained by their interaction. The within-mat variability of I_{max} in the AD mats was significantly larger than in the other mats (**Figure 3D**), which was mainly because the I_{max} values were generally larger in the AD mats.

Average vertical profiles of A_{chl} for AD, BR, and SP mats were characterized by a subsurface maximum, whereas a monotonous decrease with depth was observed for AU mats (Figure S2A). Average Y_{max} and I_k generally decreased with depth (Figures S2B,C) and were significantly correlated ($p < 0.001$) for each mat

sample. Spatial trends in I_{max} were not so clear, having patchy distributions for some mats while on average increasing with depth for others (**Figure 2** and Figure S2D). Depending on the mats, vertical variation across the cyanobacterial layer, expressed as a coefficient of variation (SD/mean) of the average vertical profile, reached 10–30% for A_{chl}, 5–45% for Y_{max}, 10–60% for I_k, and 5–100% for I_{max} (**Figure 3**).

The within-site variability of A_{chl} calculated for the mats from the AD and AU sites was lower than the within-mat variabilities calculated for the individual mats from these sites (**Figure 3A**). This was often the case also for Y_{max}, I_k, and I_{max}, but sometimes the within-mat variability was lower than the within-site one (**Figures 3B–D**). Such comparison could not be done reliably for the BR and SP mats because of the limited number of replicates ($N = 2$) for these two sites.

The global-scale variability was larger than the within-site variability for all measured physiological parameters (**Figure 3**). In contrast, in many cases it was lower than the within-mat variability. This relationship was most pronounced for A_{chl} and I_k, where it was observed in 21 (for A_{chl}) and 12 (for I_k) out of 24 mat samples (**Figures 3A,C**). Clearly, this was primarily due to the pronounced variability in the vertical direction, whereas the horizontal within-mat variability was almost always lower than the global-scale variability. On the other hand, the global-scale variability in Y_{max} and I_{max} was mostly larger than the within-mat variability, except for a few AD mats (**Figures 3B,D**). The global-scale variability, expressed as a coefficient of variation, was about 15% for A_{chl}, 40% for Y_{max}, 30% for I_k, and 80% for I_{max} (**Figure 3**).

FIGURE 3 | Variabilities of the physiological parameters A_{chl} (A), Y_{max} (B), I_k (C), and I_{max} (D) calculated at different spatial scales, including the micrometer (within-mat), meter (within-site), and **global (between sites) scales.** See text for details. The variabilities are expressed as standard deviations, SD (left axis), and as a coefficient of variation (right axis).

CLUSTERING OF MATS BASED ON AVERAGE PHYSIOLOGICAL PARAMETERS

Physiological parameters A_{chl}, Y_{max}, I_{max}, and I_k, when averaged over the cyanobacterial layers, varied significantly between the sampled sites (**Figure 4**, **Table 2**). For example, AU and BR mats had, on average, about 2-fold larger Y_{max} than AD and SP mats, SP mats had the lowest I_k, whereas AD mats had the largest values of I_{max}. When a distance matrix was calculated from the average values of A_{chl}, Y_{max}, I_k, and I_{max} using Euclidean metric, its visualization in an MDS plot revealed clear clustering of the mats according to the site of their origin (**Figure 4C**).

CLUSTERING OF MATS BASED ON THE MICROBIAL COMMUNITY COMPOSITION

ARISA fingerprinting revealed that the microbial communities in the mats from a given site were more similar to each other than to those from other sites (**Figure 5A**). This marked endemism was further supported by significant ANOSIM test ($P = 0.0001$). Overall, the sampled bacterial communities shared between 18 and 34% OTUs, with 84 out of 398 OTUs (21.1%) found everywhere (i.e., at least in one mat from a given site) and 314 OTUs being mat-specific.

LINK BETWEEN COMMUNITY COMPOSITION AND FUNCTION

To assess the degree of concordance between community composition and their potential function (photosynthetic capacity), we compared the bacterial community composition as depicted by the NMDS plot in **Figure 5A** with that of the configuration of A_{chl}, Y_{max}, I_k, and I_{max} (**Figure 4C**) by Procrustes analyses. A significant concordance between ordinations was found ($r = 0.605$, $P = 0.002$, based on 1000 permutations; **Figure 5B**), suggesting that distinct communities were associated with distinct (potential) functions.

ENVIRONMENTAL PARAMETERS

Most parameters characterizing the habitat of the studied mats were obtained from literature and are summarized in **Table 1**. The

sites strikingly differed with respect to nutrient concentrations in the overlying water, which were extremely high for the AU mats. This was very likely because the sediments at this site contained large biomass of insect larvae (alive and dead), which could be a source of nutrients (Behie et al., 2012). Additionally, the site was inundated in brief and infrequent intervals (during a spring tide). Another notable difference between the sites was due to the salinity fluctuations, which were very large at the AD and BR sites.

DISCUSSION

SCALING OF VARIABILITY IN PHYSIOLOGICAL PARAMETERS

The imaging approach used in this study enabled us to assess how the spatial variability in parameters characterizing the photosynthetic potential in cyanobacterial mats scales depending on the sampling resolution (from sub-millimeter to thousands of km). Our main result is that this scaling pattern differed depending on the studied parameter, with chlorophyll *a* absorptivity (A_{chl}) and acclimation intensity (I_k) having on average the largest variability on the micrometer-scale (mostly in the vertical direction) while the maximum quantum yield of PSII (Y_{max}) and the irradiance at which this maximum yield is reached (I_{max}) were most variable on the global scale (between sites). This contrast suggests that the different physiological parameters are controlled by different environmental factors, as discussed below.

The pronounced micrometer-scale vertical variability in A_{chl} was most likely due to the combined effects of light and nutrients. The growth of photosynthetic microbial mats is typically limited by nutrients, with most nutrients assimilated in the cyanobacterial layer originating from organic matter remineralization underneath the layer (Jonkers et al., 2003). On the other hand, light quantity strongly attenuates with depth due to intense absorption by photopigments and abiotic components of the mat matrix (Kühl and Jørgensen, 1992, 1994; Kühl et al., 1994). Thus, because of the opposing gradients in light and nutrients, cyanobacterial growth at the top and bottom of the cyanobacterial layer is likely limited by nutrients and light, respectively, whereas a location with an optimal supply of light and nutrients is somewhere in the

FIGURE 4 | (A,B) Relationships between physiological parameters A_{chl}, Y_{max}, I_k, and I_{max} in the studied cyanobacterial mats. Symbols represent averages over the cyanobacterial layers, thin and thick error-bars depict the total and vertical variability (expressed as standard deviation, SD) within the layers, respectively. **(C)** Multidimensional scaling plot of the distance matrix calculated based on the average values of the parameters shown in **(A,B)** using Euclidean metric.
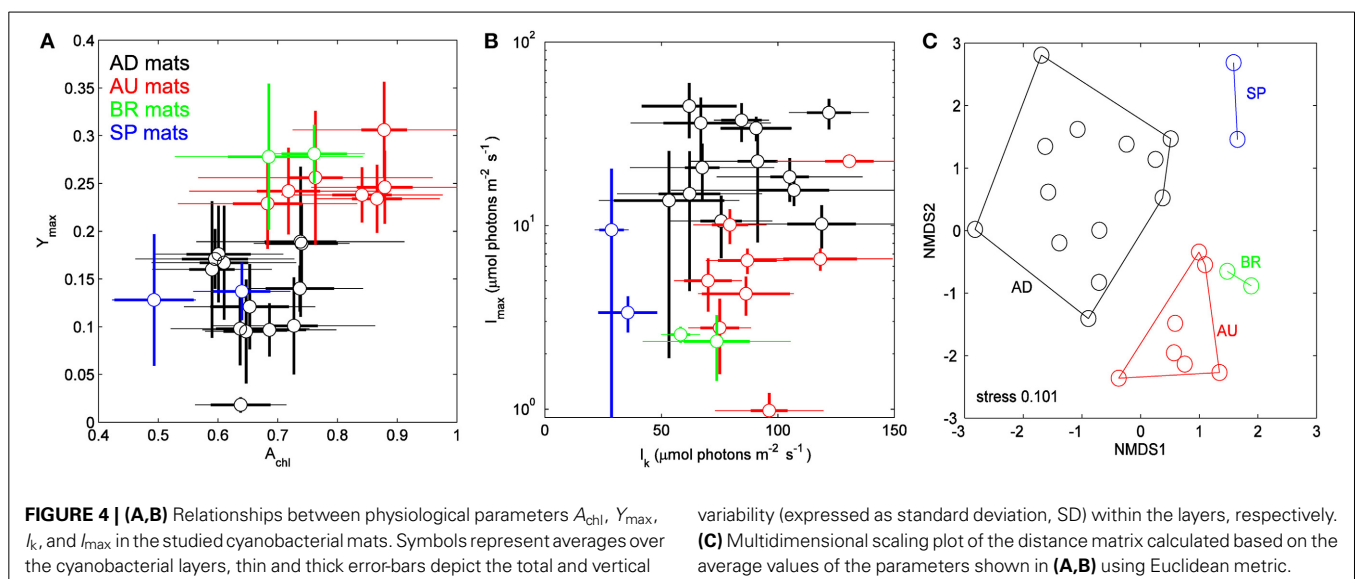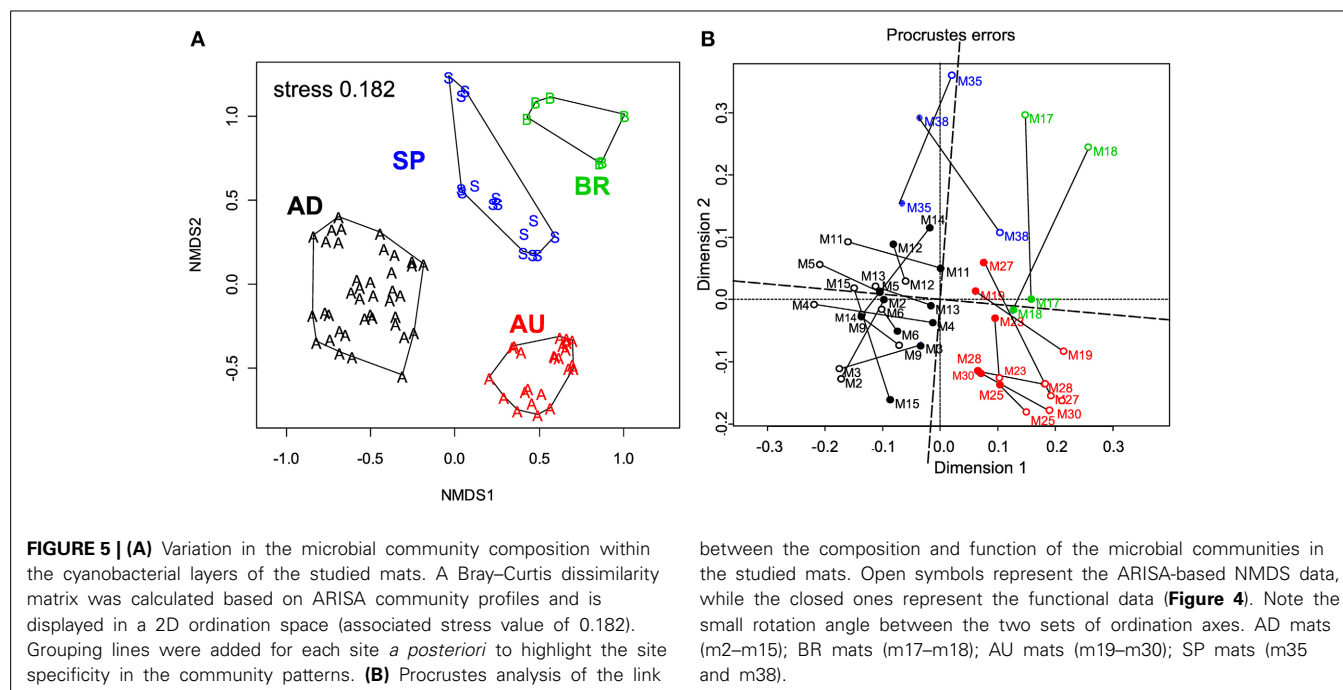
Table 2 | Average values and standard deviations of the physiological parameters characterizing the mats from the different sites.

	AD	AU	BR	SP	p^a
A_{chl}	0.66 ± 0.06 [2][b]	0.80 ± 0.08 [1]	0.72 ± 0.05 [1–2]	0.57 ± 0.10 [2]	0.0004
Y_{max}	0.13 ± 0.05 [2]	0.25 ± 0.03 [1]	0.280 ± 0.002 [1]	0.13 ± 0.01 [2]	8×10^{-6}
$I_k (\mu mol\ m^{-2}\ s^{-1})$	85 ± 23 [1]	88 ± 16 [1]	66 ± 11 [1–2]	32 ± 5 [2]	0.01
$I_{max} (\mu mol\ m^{-2}\ s^{-1})$	24 ± 12 [1]	5 ± 3 [2]	2.4 ± 0.2 [2]	6 ± 4 [2]	1×10^{-5}

[a] Probability that the means between the different sites are equal, as determined by ANOVA. The I_{max} values were log-transformed before ANOVA to ensure variance homogeneity.

[b] Means ranked as [1] are significantly larger than means ranked as [2], means ranked as [1–2] are not significantly different from those ranked as [1] and [2].

FIGURE 5 | (A) Variation in the microbial community composition within the cyanobacterial layers of the studied mats. A Bray–Curtis dissimilarity matrix was calculated based on ARISA community profiles and is displayed in a 2D ordination space (associated stress value of 0.182). Grouping lines were added for each site *a posteriori* to highlight the site specificity in the community patterns. **(B)** Procrustes analysis of the link between the composition and function of the microbial communities in the studied mats. Open symbols represent the ARISA-based NMDS data, while the closed ones represent the functional data (**Figure 4**). Note the small rotation angle between the two sets of ordination axes. AD mats (m2–m15); BR mats (m17–m18); AU mats (m19–m30); SP mats (m35 and m38).

middle. This suggests that the cyanobacterial biomass should have a maximum somewhere around this optimal location. Assuming that the measured chlorophyll *a* absorptivity, A_{chl}, is a proxy for cyanobacterial biomass, our results are consistent with this interpretation: while A_{chl} had a clear subsurface maximum in mats from sites where the N and P concentrations in the overlying water were low (AD, SP, BR), it decreased sharply from the surface in the mats from the AU site, where the nutrient concentrations in the overlying water were very high (**Table 1**, Figure S2). This latter characteristic of the AU site was most likely also the main factor responsible for on average the highest values of A_{chl} (**Figure 4A**, **Table 2**) and its vertical variability (**Figure 3A**) in the AU mats as compared to the mats from the other sites. Additional explanation for the observed within-mat variability in A_{chl} is the tendency of phototrophic cells to have a larger cellular pigment content when grown at lower irradiances (Falkowski, 1980; Kirk, 2011). Thus, at least in the mats from AD, SP, and BR, the observed subsurface maximum in A_{chl} could additionally be due to this adaptation of cyanobacterial cells to light limitation that progressively increases with depth in the mat. Together, this suggests that the average cyanobacterial biomass in the cyanobacterial layer of

the studied mats was mostly determined by the overlying water nutrient content, whereas the micrometer-scale distribution of the biomass, and possibly also of the average pigment content in the cyanobacterial cells, was additionally shaped by light.

With respect to the light acclimation intensity, I_k, light availability was likely the most important factor that determined both its micrometer-scale and global-scale variation. Assuming that cyanobacteria adapt to local light conditions, the sharp decrease in light intensity with depth in mats should result in cyanobacterial populations in deeper parts of the euphotic zone being acclimated to lower light intensities. This is what we generally observed for mats from all sites (Figure S2). Although we do not have reliable data on site-specific downwelling irradiances, the observation that the global-scale variability of I_k was comparable to, and in many occasions lower than, the micrometer-scale variability suggested that light variation within the photosynthetically active cyanobacterial layer was similar or larger than the variation in the average light dose received by the mats from the different sites.

Although the micrometer-scale variability of the maximum effective quantum yield, Y_{max}, was strongly positively correlated

with the light acclimation intensity for each studied mat (see e.g., Figure S2), light availability is not a likely factor that controls its spatial distribution. This is primarily because Y_{max} is, by definition, a proxy for the maximum photosynthetic *potential*, which is reached at light intensities that are considerably lower than the acclimation intensity (in the dark for plants, at $I_{max} \ll I_k$ for the cyanobacterial populations in this study; **Figure 1**). Since the quantum yield of PSII determined by the PAM measurement of variable chlorophyll fluorescence relates to the redox state of the plastoquinone (PQ) pool in the photosynthetic electron transport chain (Allen, 2003), parameters that affect this redox state are likely affecting also Y_{max}. In the context of cyanobacterial mats, H_2S and O_2 are likely candidates.

The response of cyanobacterial photosynthesis to H_2S is well-documented (Cohen et al., 1986; Jørgensen et al., 1986; Miller and Bebout, 2004), with several species being able to perform anoxygenic photosynthesis using H_2S as the electron donor. In such cyanobacteria, H_2S oxidation is facilitated by the activity of sulfide-quinone-reductase (SQR) (Bronstein et al., 2000; Griesbeck et al., 2000), which reduces the PQ molecule using the electrons transferred from H_2S and can therefore decrease the apparent quantum yield of PSII. Our preliminary experiments with an axenic cyanobacterial culture embedded in agarose showed that exposure to H_2S in the μM range lead to a rapid decrease in the quantum yield Y, both in the dark and light, and the yield recovery in the light occurred only after H_2S decreased below a threshold in the μM range (Figure S3). A similar negative effect of H_2S on Y was observed in other cyanobacterial systems (e.g., living stromatolites; Kromkamp et al., 2007). Therefore, the pronounced decrease in Y_{max} with depth in the cyanobacterial layer may be due to the steep increase in H_2S in this layer at low-light conditions, which is a well-documented phenomenon in mats (e.g., Wieland and Kühl, 2000b; Jonkers et al., 2003; Garcia de Lomas et al., 2005; Al-Thani et al., 2014).

The effect of O_2 on Y is likely because photosynthetic and respiratory electron transport chains in a cyanobacterial cell share three electron carriers in the thylakoid membrane (PQ, cytochrome b_6f and the terminal oxidases; Vermaas, 2001), which makes O_2 an important electron acceptor for electrons from the reduced PQ (chlororespiration) and also for those coming from PSI via ferredoxin (Mehler reaction) (Schreiber et al., 2002). Presence of O_2 can therefore contribute to a partial oxidation of the PQ pool, which could manifest itself as an increase in the apparent Y. The steep decrease in O_2 with depth, which is typical for cyanobacterial mats under low light conditions (Wieland and Kühl, 2000a, 2006; Al-Najjar et al., 2012), could therefore additionally be responsible for the observed decline of Y_{max} with depth.

In addition to environmental parameters mentioned above, part of the observed spatial variability of Y_{max}, I_k and A_{chl} could be due to artifacts linked to autofluorescence (for Y_{max} and I_k) and absorption (for A_{chl}) by photosynthetically inactive components in the cyanobacterial layer, such as dead/inactive cyanobacterial cells, pigment degradation products or mineral particles. Such components are abundant in the cyanobacterial layer of microbial mats (e.g., Kühl et al., 1994; Al-Najjar et al., 2012), and because their fluorescence is not variable but

possibly comparable to that from the PSII of photosynthetically active cells, the detected quantum yield of PSII can appear lower. Similarly, their absorption in the same wavelength range as chlorophyll *a* could increase the apparent A_{chl}. These effects may have contributed to the differences in the "baseline" of these parameters and thus also their apparent global variability. The importance of this contribution could, however, not be estimated based on our present data.

In contrast to the monotonous decrease in the effective quantum yield, Y, with the actinic irradiance, I, which is characteristic for plants (White and Critchley, 1999) and eukaryotic algae (Flameling and Kromkamp, 1998), our measurements in cyanobacterial mats showed that the relationship between Y and I was not monotonous (**Figure 1**). Instead, Y reached a maximum at actinic intensities I_{max} that varied in the range from one to several tens of μmol photons m^{-2} s^{-1} (**Figure 4B**). The spatial patterns for I_{max} within a given mat were not very clear (see, e.g., **Figure 2**), primarily because of the noise in the fluorescence yield images. Although we were able to find a suitable mathematical expression for this relationship (Equation 2), further research is required to understand why Y reached a maximum at intensities $I_{max} > 0$ and what controls this behavior.

IMPLICATIONS FOR MICROBIAL ECOLOGY

Our imaging data demonstrate that the variability of a specific function (here photosynthesis) in a microbial community (here the cyanobacterial layer) can be at least as large on the micrometer-scale within the community as it is on the global scale between communities from different locations. As discussed above, this is primarily due to the strong within-mat variability of the physico-chemical parameters that control the function. Furthermore, our comparison of the cyanobacterial layers based on their average functional (photosynthetic) potential and microbial community composition revealed that the mats clustered according to their site of origin (**Figures 4C, 5A**) and that the functional and compositional data were significantly linked (**Figure 5B**). Based on these results we expect that if physico-chemical parameters exhibit pronounced variability within a microbial community, which is common in transport-limited systems such microbial mats and sediments, the community composition will also vary at least as much on the micrometer-scale (within community) as it does on the global scale (between communities). This implies that analyses of the microbial community composition that ignore this micrometer-scale variability (e.g., by sampling bulk volumes of sediments rather than layer-by-layer) may have a limited value in identifying correlations between the microbial composition, function and the corresponding environmental settings. In other words, when sampling microbial communities in order to identify these correlations, the prefix "micro" should refer not only to the size of the inhabitants but also to the required spatial resolution of sampling.

Interestingly, despite the large micrometer-scale variability of the functional potential within each cyanobacterial layer community, distinct communities from a given site were on average characterized by a distinct, albeit highly variable, functional potential (**Figure 4**). This suggests that out of the environmental parameters discussed in the previous section, there are possibly

one or more factors that have the most dominant influence on the composition and function of the studied cyanobacterial mats. Unfortunately, these dominant factors could not be identified from the limited dataset presented in this study.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fmicb.2014.00406/abstract

REFERENCES

Abed, R. M. M., Polerecky, L., Al Najjar, M., and De Beer, D. (2006). Effect of temperature on photosynthesis, oxygen consumption and sulfide production in an extremely hypersaline cyanobacterial mat. *Aquat. Microb. Ecol.* 44, 21–30. doi: 10.3354/ame044021

Adame, M. F., Reef, R., Grinham, A., Holmes, G., and Lovelock, C. E. (2012). Nutrient exchange of extensive cyanobacterial mats in an arid subtropical wetland. *Marine Freshw. Res.* 63, 457–467. doi: 10.1071/MF11133

Allen, J. F. (2003). State transitions - a question of balance. *Science* 299, 1530–1532. doi: 10.1126/science.1082833

Al-Najjar, M. A. A., De Beer, D., Jørgensen, B. B., Kühl, M., and Polerecky, L. (2010). Conversion and conservation of light energy in a photosynthetic microbial mat ecosystem. *ISME J.* 4, 440–449. doi: 10.1038/ismej.2009.121

Al-Najjar, M. A. A., De Beer, D., Kühl, M., and Polerecky, L. (2012). Light utilization efficiency in photosynthetic microbial mats. *Environ. Microbiol.* 4, 982–992. doi: 10.1111/j.1462-2920.2011.02676.x

Al-Thani, R., Al-Najjar, M. A., Al-Raei, A. M., Ferdelman, T., Thang, N. M., Al Shaikh, I., et al. (2014). Community structure and activity of a highly dynamic and nutrient-limited hypersaline microbial mat in Um Alhool Sabkha, Qatar. *PLoS ONE* 9:e92405. doi: 10.1371/journal.pone.0092405

Baker, N. R. (2008). Chlorophyll fluorescence: a probe of photosynthesis *in vivo. Annu. Rev. Plant Biol.* 59, 89–113. doi: 10.1146/annurev.arplant.59.032607.092759

Behie, S. W., Zelisko, P. M., and Bidochka, M. J. (2012). Endophytic insect-parasitic fungi translocate nitrogen directly from insects to plants. *Science* 336, 1576–1577. doi: 10.1126/science.1222289

Boer, S. I., Hedtkamp, S. I. C., Van Beusekom, J. E. E., Fuhrman, J. A., Boetius, A., and Ramette, A. (2009). Time- and sediment depth-related variations in bacterial diversity and community structure in subtidal sands. *ISME J.* 3, 780–791. doi: 10.1038/ismej.2009.29

Bronstein, M., Schutz, M., Hauska, G., Padan, E., and Shahak, Y. (2000). Cyanobacterial sulfide-quinone reductase: cloning and heterologous expression. *J. Bacteriol.* 182, 3336–3344. doi: 10.1128/JB.182.12.3336-3344.2000

Campbell, D., Hurry, V., Clarke, A. K., Gustafsson, P., and Oquist, G. (1998). Chlorophyll fluorescence analysis of cyanobacterial photosynthesis and acclimation. *Microbiol. Mol. Biol. Rev.* 62, 667–683.

Cardinale, M., Brusetti, L., Quatrini, P., Borin, S., Puglia, A. M., Rizzi, A., et al. (2004). Comparison of different primer sets for use in automated ribosomal intergenic spacer analysis of complex bacterial communities. *Appl. Environ. Microbiol.* 70, 6147–6156. doi: 10.1128/AEM.70.10.6147-6156.2004

Caumette, P., Matheron, R., Raymond, N., and J.C., R. (1994). Microbial mats in the hypersaline ponds of Mediterranean salterns (Salins-de-Girard, France). *FEMS Microbiol. Ecol.* 13, 273–286. doi: 10.1111/j.1574-6941.1994.tb00074.x

Cohen, Y., Jørgensen, B. B., Revsbech, N. P., and Poplawski, R. (1986). Adaptation to hydrogen-sulfide of oxygenic and anoxygenic photosynthesis among Cyanobacteria. *Appl. Environ. Microbiol.* 51, 398–407.

Des Marais, D. J. (2003). Biogeochemistry of hypersaline microbial mats illustrates the dynamics of modern microbial ecosystems and the early evolution of the biosphere. *Biol. Bull.* 204, 160–167. doi: 10.2307/1543552

Falkowski, P. G. (1980). "Light-shade adaptation in marine phytoplankton," in *Primary Productivity in the Sea*, ed P. G. Falkowski (New York, NY: Plenum Press), 99–119.

Flameling, I. A., and Kromkamp, J. (1998). Light dependence of quantum yields for PSII charge separation and oxygen evolution in eucaryotic algae. *Limnol. Oceanogr.* 43, 284–297. doi: 10.4319/lo.1998.43.2.0284

Garcia de Lomas, J., Corzo, A., Garcia, C. M., and Van Bergeijk, S.,A (2005). Microbenthos in a hypersaline tidal lagoon: factors affecting microhabitat, community structure and mass exchange at the sediemnt-water interface. *Aquat. Microb. Ecol. gy* 38, 53–69. doi: 10.3354/ame038053

Garcia-Pichel, F., Kühl, M., Nübel, U., and Muyzer, G. (1999). Salinity-dependent limitation of photosynthesis and oxygen exchange in microbial mats. *J. Phycol.* 35, 227–238. doi: 10.1046/j.1529-8817.1999.3520227.x

Gower, J. C. (1975). Generalized procrustes analysis. *Phychometrika* 40, 33–51. doi: 10.1007/BF02291478

Griesbeck, C., Hauska, G., and Schütz, M. (2000). "Biological sulfide-oxidation: sulfide-quinone reductase (sqr), the primary reaction," in *Recent Research Developments in Microbiology*, Vol. 4, ed S. G. Pandalai (Trivadrum: Research Signpost), 179–203.

Jonkers, H. M., Ludwig, R., De Wit, R., Pringault, O., Muyzer, G., Niemann, H., et al. (2003). Structural and functional analysis of a microbial mat ecosystem from a unique permanent hypersaline inland lake: 'La Salada de Chiprana' (NE Spain). *FEMS Microbiol. Ecol.* 44, 175–189. doi: 10.1016/S0168-6496(02)00464-6

Jørgensen, B. B., Cohen, Y., and Revsbech, N. P. (1986). Transition from anoxygenic to oxygenic photosynthesis in a Microcoleus-chthonoplastes cyanobacterial mat. *Appl. Environ. Microbiol.* 51, 408–417.

Kirk, J. T. O. (2011). *Light and Photosynthesis in Aquatic Ecosystems.* Cambridge: Cambridge University Press.

Kromkamp, J. C., Perkins, R., Dijkman, N., Consalvey, M., Andres, M., and Reid, R. P. (2007). Resistance to burial of cyanobacteria in stromatolites. *Aquat. Microb. Ecol.* 48, 123–130. doi: 10.3354/ame048123

Kühl, M. (1993). "Photosynthesis, O_2 respiration and sulfur cycling in a cyanobacterial biofilm," in *Trends in Microbial Ecology*, eds R. Guerrero and C. Pedros-Alio (Spanish Society for Microbiology), 163–167.

Kühl, M. (2005). Optical microsensors for analysis of microbial communities. *Methods Enzymol.* 397, 166–199. doi: 10.1016/S0076-6879(05)97010-9

Kühl, M., and Fenchel, T. (2000). Bio-optical characteristics and the vertical distribution of photosynthetic pigments and photosynthesis in an artificial cyanobacterial mat. *Microb. Ecol.* 40, 94–103. doi: 10.1007/s002480000061

Kühl, M., Glud, R. N., Ploug, H., and Ramsing, N. B. (1996). Microenvironmental control of photosynthesis and photosynthesis-coupled respiration in an epilithic cyanobacterial biofilm. *J. Phycol.* 32, 799–812. doi: 10.1111/j.0022-3646.1996.00799.x

Kühl, M., and Jørgensen, B. B. (1992). Spectral light measurements in microbenthic phototrophic communities with a fiber-optic microprobe coupled to a sensitive diode array detector. *Limnol. Oceanogr.* 37, 1813–1823. doi: 10.4319/lo.1992.37.8.1813

Kühl, M., and Jørgensen, B. B. (1994). The light-field of microbenthic communities - radiance distribution and microscale optics of sandy coastal sediments. *Limnol. Oceanogr.* 39, 1368–1398. doi: 10.4319/lo.1994.39.6.1368

Kühl, M., Lassen, C., and Jørgensen, B. B. (1994). "Optical properties of microbial mats: light measurements with fiber-optic microprobes," in *Microbial Mats: Structure, Development and Environmental Significance,* eds J. Stal and P. Caumette (Berlin: Springer), 149–167.

Kühl, M., and Polerecky, L. (2008). Functional and structural imaging of phototrophic microbial communities and symbioses. *Aquat. Microb. Ecol.* 53, 99–118. doi: 10.3354/ame01224

Lovelock, C., Grinham, A., Adame, M. F., and Penrose, H. (2010). Elemental composition and productivity of cyanobacterial mats in an arid zone estuary in north Western Australia. *Wetlands Ecol. Manag.* 18, 37–47. doi: 10.1007/s11273-009-9146-6

Miller, S. R., and Bebout, B. M. (2004). Variation in sulfide tolerance of photosystem II in phylogenetically diverse cyanobacteria from sulfidic habitats. *Appl. Environ. Microbiol.* 70, 736–744. doi: 10.1128/AEM.70.2.736-744.2004

Nübel, U., Garcia-Pichel, F., Kühl, M., and Muyzer, G. (1999). Spatial scale and the diversity of benthic cyanobacteria and diatoms in a salina. *Hydrobiologia* 401, 199–206. doi: 10.1023/A:1003702814150

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Peter, R., Minchin O'Hara, R. B., et al. (2012). *Vegan: Community Ecology Package.* R package Version 2.0–4.

Peres-Neto, P. R., and Jackson, D. A. (2001). How well do multivariant data sets match? The advantages of procrustean superimposition approach over the Mantel test. *Oecologia* 129, 169–178. doi: 10.1007/s004420100720

Pinckney, J. L., Long, R. A., and Paerl, H. W. (2011). Structural and functional responses of microbial mats to reductions in nutrient and salinity stressors in a Bahamian hypersaline lagoon. *Aquat. Microb. Ecol.* 62, 289–298. doi: 10.3354/ame01475

Platt, T., and Jassby, A. D. (1976). The relationship between photosynthesis and light for natural assemblages of coastal marine phytoplankton. *J. Phycol.* 12, 421–430.

Polerecky, L., Bissett, A., Al-Najjar, M., Faerber, P., Osmers, H., Suci, P. A., et al. (2009). Modular spectral imaging system for discrimination of pigments in cells and microbial communities. *Appl. Environ. Microbiol.* 75, 758–771. doi: 10.1128/AEM.00819-08

Ramette, A. (2007). Multivariate analyses in microbial ecology. *FEMS Microbiol. Ecol.* 62, 142–160. doi: 10.1111/j.1574-6941.2007.00375.x

Ramette, A. (2009). Quantitative community fingerprinting methods for estimating the abundance of operational taxonomic units in natural microbial communities. *Appl. Environ. Microbiol.* 75, 2495–2505. doi: 10.1128/AEM.02409-08

Roeselers, G., Norris, T. B., Castenholz, R. W., Rysgaard, S., Glud, R. N., Kühl, M., et al. (2007). Diversity of phototrophic bacteria in microbial mats from Arctic hot springs (Greenland). *Environ. Microbiol.* 9, 26–38. doi: 10.1111/j.1462-2920.2006.01103.x

Schreiber, U., Gademann, R., Bird, P., Ralph, P. J., Larkum, A. W. D., and Kühl, M. (2002). Apparent light requirement for activation of photosynthesis upon rehydration of disiccated beachrock microbial mat. *J. Phycol.* 38, 125–134. doi: 10.1046/j.1529-8817.2002.01103.x

Schreiber, U., Gademann, R., Ralph, P. J., and Larkum, A. W. D. (1997). Assessment of photosynthetic performance of Prochloron in Lissoclinum patella in hospite by chlorophyll fluorescence measurements. *Plant Cell Physiol.* 38, 945–951. doi: 10.1093/oxfordjournals.pcp.a029256

Schreiber, U., Kühl, M., Klimant, I., and Reising, H. (1996). Measurement of chlorophyll fluorescence within leaves using a modified PAM Fluorometer with a fiber-optic microprobe. *Photosyn. Res.* 47, 103–109. doi: 10.1007/BF00017758

Seckbach, J., and Oren, A. (2010). *Microbial Mats: Modern and Ancient Microorganisms in Stratified Systems.* Heidelberg: Springer.

Stal, L. (2000). "Cyanobacterial mats and stromatolites," in *The Ecology of Cyanobacteria,* eds A. B. Whitton and M. Potts (Dordrecht: Kluwer Academic Publishers), 61–120.

van Gemerden, H. (1993). Microbial mats: a joint venture. *Marine Geol.* 113, 3–25.

Vasconcelos, C., Warthmann, R., Mckenzie, J. A., Visscher, P. T., Bittermann, A. G., and Van Lith, Y. (2006). Lithifying microbial mats in Lagoa Vermelha, Brazil: modern precambrian relics? *Sediment. Geol.* 185, 175–183. doi: 10.1016/j.sedgeo.2005.12.022

Vermaas, W. F. J. (2001). "Photosynthesis and respiration in cyanobacteria," in *Encyclopedia of Life Sciences* (London: Nature Publishing Group), 245–251.

Webb, W. L., Newton, M., and Starr, D. (1974). Carbon-dioxide exchange of Alnus-Rubra - mathematical-model. *Oecologia* 17, 281–291. doi: 10.1007/BF00345747

White, A. J., and Critchley, C. (1999). Rapid light curves: a new fluorescence method to assess the state of the photosynthetic apparatus. *Photosyn. Res.* 59, 63–72. doi: 10.1023/A:1006188004189

Wieland, A., and Kühl, M. (2000a). Irradiance and temperature regulation of oxygenic photosynthesis and O-2 consumption in a hypersaline cyanobacterial mat (Solar Lake, Egypt). *Mar. Biol.* 137, 71–85. doi: 10.1007/s002270000331

Wieland, A., and Kühl, M. (2000b). Short-term temperature effects on oxygen and sulfide cycling in a hypersaline cyanobacterial mat (Solar Lake, Egypt). *Marine Ecol. Progress Series* 196, 87–102. doi: 10.3354/meps196087

Wieland, A., and Kühl, M. (2006). Regulation of photosynthesis and oxygen consumption in a hypersaline cyanobacterial mat (Camargue, France) by irradiance, temperature and salinity. *FEMS Microbiol. Ecol.* 55, 195–210. doi: 10.1111/j.1574-6941.2005.00031.x

Wieland, A., Kühl, M., Mcgowan, L., Fourcans, A., Duran, R., Caumette, P., et al. (2003). Microbial mats on the Orkney Islands revisited: microenvironment and microbial community composition. *Microb. Ecol.* 46, 371–390. doi: 10.1007/s00248-002-0108-2

Nutrient requirements and growth physiology of the photoheterotrophic Acidobacterium, *Chloracidobacterium thermophilum*

Marcus Tank[1] and Donald A. Bryant[1,2]*

[1] *Department of Biochemistry and Molecular Biology, Eberly College of Science, The Pennsylvania State University, PA, USA,*
[2] *Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA*

A novel thermophilic, microaerophilic, anoxygenic, and chlorophototrophic member of the phylum *Acidobacteria*, *Chloracidobacterium thermophilum* strain B[T], was isolated from a cyanobacterial enrichment culture derived from microbial mats associated with Octopus Spring, Yellowstone National Park, Wyoming. *C. thermophilum* is strictly dependent on light and oxygen and grows optimally as a photoheterotroph at irradiance values between 20 and 50 μmol photons m^{-2} s^{-1}. *C. thermophilum* is unable to synthesize branched-chain amino acids (AAs), L-lysine, and vitamin B_{12}, which are required for growth. Although the organism lacks genes for autotrophic carbon fixation, bicarbonate is also required. Mixtures of other AAs and 2-oxoglutarate stimulate growth. As suggested from genomic sequence data, *C. thermophilum* requires a reduced sulfur source such as thioglycolate, cysteine, methionine, or thiosulfate. The organism can be grown in a defined medium at 51°C (T_{opt}; range 44–58°C) in the pH range 5.5–9.5 (pH_{opt} = ~7.0). Using the defined growth medium and optimal conditions, it was possible to isolate new *C. thermophilum* strains directly from samples of hot spring mats in Yellowstone National Park, Wyoming. The new isolates differ from the type strain with respect to pigment composition, morphology in liquid culture, and temperature adaptation.

Keywords: bacteriochlorophyll, anoxygenic photosynthesis, photoheterotroph, thermophile, Acidobacteria

Introduction

"*Candidatus* Chloracidobacterium thermophilum" was first detected by bioinformatics analyses of metagenomic sequence data for phototrophic microbial mats from alkaline (~pH 8.2) siliceous hot springs (50–65°C) in Yellowstone National Park, Wyoming, USA (Bryant et al., 2007). Inferences from these analyses indicated that this previously uncharacterized bacterium had a photosynthetic apparatus that was highly similar to that of obligately anaerobic, anoxygenic, and chlorophotrophic members of the *Chlorobiales* (i.e., green sulfur bacteria). It was inferred that "*Ca*. C. thermophilum" synthesized BChl *a* and *c*, and that it had a photosynthetic apparatus comprising chlorosomes as light-harvesting antenna complexes, the Fenna-Matthews-Olson, BChl *a*-binding protein (FmoA),

---

Abbreviations: BChl, bacteriochlorophyll; Chl, chlorophyll; HPLC, high performance liquid chromatography.

and homodimeric type-1 photochemical reaction centers. However, in contrast to green sulfur bacteria, BChl biosynthesis and some other cellular processes (e.g., tyrosine biosynthesis) apparently are oxygen-dependent (Bryant et al., 2007; Garcia Costas et al., 2012a).

Phylogenetic analyses of the sequences of 16S rRNA and RecA assigned the new bacterium to subdivision 4 of the highly diverse phylum, *Acidobacteria* (Barns et al., 1999; Bryant et al., 2007; Tank and Bryant, 2015). *C. thermophilum* is currently the only cultivated, phototrophic member of this phylum, and together with *Pyrinomonas methylaliphatogenes, Blastocatella fastidiosa,* and *Aridibacter famidurans* and *A. kavangonen*sis, *C. thermophilum* (Tank and Bryant, 2015) is one of the few axenic strains in this subdivision (Foesel et al., 2013; Crowe et al., 2014; Huber et al., 2014). At the time of its initial discovery, "*Ca.* C. thermophilum" extended the number of bacterial phyla known to have members capable of Chl -dependent phototrophic growth (i.e., chlorophototrophy) from five to six. These include the phyla *Cyanobacteria, Chloroflexi, Chlorobi, Proteobacteria, Firmicutes,* and now *Acidobacteria* (Bryant et al., 2007). This number has very recently expanded from six to seven because of the discovery of a BChl *a*-producing, anoxygenic photoheterotroph from the phylum *Gemmatimonadetes* (Zeng et al., 2014).

As opposed to the "*in silico*" evidence describing this organism, the existence of *C. thermophilum* was confirmed by identifying living cells that exhibited similar DNA signatures to those of the organism described by metagenomic analysis. Those cells were detected in a cyanobacterial enrichment generated from Octopus Spring and were used to initiate the characterization of *C. thermophilum*. The cyanobacterium in the enrichment, a *Synechococcus* sp., was rather easily eliminated by providing a mixture of carbon sources to the enrichment culture and by adding atrazine to inhibit the growth of the cyanobacterium (Bryant et al., 2007). However, in spite of considerable effort, attempts to eliminate two heterotrophic contaminants, *Anoxybacillus* sp. and *Meiothermus* sp., were unsuccessful over a period of years. This suggested that these bacteria were providing essential nutrients to *C. thermophilum*, removing growth inhibitory substances, and/or otherwise changing the culture conditions in a way that was essential for the growth of *C. thermophilum*.

Using physical methods (e.g., low-speed centrifugation) to obtain enriched populations of cells, it was possible to obtain highly enriched DNA preparations for *C. thermophilum*. This allowed the complete genome of the organism to be determined (Garcia Costas et al., 2012a). The genome sequence verified many of the inferences from the metagenomic data concerning this bacterium, and these data further revealed important clues about the physiology and metabolism of *C. thermophilum*. Consistent with inferences from the metagenomic data, the genome lacked genes for known CO_2 fixation pathways, and it additionally lacked genes for oxidoreductases that could provide electrons for CO_2 reduction. Moreover, the genome lacked all genes except those encoding aminotransferases for the synthesis of the branched-chain AAs, L-leucine, L-isoleucine, and L-valine. Surprisingly, genes encoding enzymes for the complete degradation of branched-chain AAs were present, however,

(Garcia Costas et al., 2012a). The genome additionally lacked genes for nitrate reductase, nitrite reductase, nitrogenase, and assimilatory sulfate reduction. Based upon these findings, Garcia Costas et al. (2012a) concluded that *C. thermophilum* was an aerobic anoxygenic photoheterotroph. A metatranscriptomic study of the microbial mats from which the organism was enriched over a complete diel cycle suggested that *C. thermophilum* might require alternating oxic and anoxic conditions for optimal growth or might prefer constantly microoxic conditions (Liu et al., 2011, 2012).

Armed with all of this background information, we describe here how the insights gained from metagenomic, genomic, and metatranscriptomic studies were evaluated and used, in combination with biochemical and classical microbiological methods, to isolate an axenic culture of *C. thermophilum*. By establishing the conditions for axenic growth, we elucidated preferred carbon, nitrogen, and sulfur sources, optimum pH, temperature and light intensities as well as the specific oxygen relationship of *C. thermophilum*. Our findings have allowed us to isolate new axenic cultures of *C. thermophilum* strains directly from microbial mats in Yellowstone National Park.

Materials and Methods

Source Material

The original source material of the organism described here was cyanobacterial enrichment culture B′-NACy10o, which was derived from a sample collected by Allewalt et al. (2006) on July 10, 2002 from Octopus Spring at a site temperature of 51–61°C. The genome sequence of the cyanobacterium in this enrichment culture, denoted as *Synechococcus* sp. JA-2-3B′ a(2-13), was reported by Bhaya et al. (2007). The enrichment culture was first simplified by elimination of the *Synechococcus* sp., which produced a stable co-culture of *C. thermophilum, Anoxybacillus* sp. and *Meiothermus* sp. as previously described (Bryant et al., 2007; Garcia Costas et al., 2012a). Procedures used to obtain the axenic type strain (strain B^T; Tank and Bryant, 2015) are described here.

Medium and Medium Preparation

In order to isolate an axenic culture of *C. thermophilum*, it was necessary to modify the medium several times in an iterative fashion. The C. thermophilum Midnight Medium (CTM medium) description that follows was the result of this process and was the basis for all further growth experiments. As described in the text, this medium was modified in some experiments to test various parameters. Several stock solutions were used during the preparation of CTM medium, pH 8.5, which was used as the basal medium for growth of *C. thermophilum*. One liter of 50 × stock solution I contained 3.75 g magnesium sulfate ($MgSO_4 \cdot 7H_2O$), 1.8 g calcium chloride ($CaCl_2 \cdot 2H_2O$), 0.45 g sodium citrate, 0.5 ml trisodium-EDTA (stock: 89.6 g L^{-1}, pH 8.0,), and 50 ml of the trace elements solution. Solution II contained 15.3 g potassium hydrogen phosphate per liter. Solution III contained 12 g of ferric ammonium citrate per liter. Solution IV contained 168.1 g of 2-oxoglutarate (sodium salt) per liter. Solutions I to IV were autoclaved prior to use. One liter of the trace elements stock

solution contained: 2.86 g boric acid (H_3BO_3), 1.81 g manganese chloride ($MnCl_2 \cdot 4H_2O$), 0.222 g zinc sulfate ($ZnSO_4 \cdot 7H_2O$), 0.39 g sodium molybdate ($Na_2MoO_4 \cdot 2H_2O$), 0.079 g cupric sulfate ($CuSO_4 \cdot 5H_2O$), and 0.0494 g cobaltous nitrate hexahydrate ($Co(NO_3)_2 \cdot 6H_2O$). A stock solution of vitamin B_{12} was prepared by adding 100 mg of cobalamin to 100 ml of double-distilled H_2O, and the resulting solution was titrated with 1 M hydrogen chloride until the cobalamin dissolved (final pH 2.7). A 100-ml stock solution of 10 mM potassium phosphate buffer containing a mixture of 13 vitamins was prepared with 10 mg each of riboflavin and biotin and 100 mg each of thiamine hydrochloride, ascorbic acid, D-calcium-panthothenate, folic acid, nicotinamide, nicotinic acid, 4-aminobenzoic acid, pyridoxine hydrochloride, lipoic acid, nicotinamide adenine dinucleotide (NAD^+), and thiamine pyrophosphate. The solution was titrated with 1 M sodium hydroxide until all compounds were dissolved (final pH 9.5). The two vitamin solutions were filter-sterilized through a 0.22-μm cellulose acetate filter and were stored in the dark at 4°C until required.

One liter of growth medium was produced by mixing 20 ml of solution I, 3 ml of solution II, 2 ml of solution III, and 2.5 ml of solution IV, together with 2.4 g of HEPES and double-distilled H_2O to a final volume of ∼970 ml. The medium was adjusted to pH 8.5 with 2 M potassium hydroxide and autoclaved at 121°C for 40 min. The medium bottle was sealed immediately after autoclaving and cooled to ∼60°C. The medium was finalized after cooling by adding 30 ml of a freshly prepared and filter-sterilized solution containing 0.125 g sodium thioglycolate, 0.625 g sodium bicarbonate, 1 ml of a Bacto™ Peptone (BD Biosciences, Sparks, MD, USA) solution (stock: 100 mg ml^{-1}), and 500 μl each of the vitamin B_{12} and 13-vitamin mixture solution was added. To produce a completely defined growth medium, Bacto™ Peptone was replaced by a mixture of all 20 common L-AAs at a concentration of 5 mg L^{-1} each. Solidified medium additionally contained 1% (w/v) Bacto™ Agar (BD Biosciences, Sparks, MD, USA) that had been washed three times with double-distilled H_2O. Agar plates were filled with ∼50 ml of microoxic medium. Plates were incubated in the light in translucent sealed plastic jars (Becton Dickinson, Franklin Lakes, NJ, USA) that had been flushed with a 10% hydrogen/10% carbon dioxide/80% nitrogen (v/v/v) gas mixture to produce reduced oxygen conditions. Vessels used for experiments requiring liquid media were typically filled to ∼75% of the volume and were not shaken or mixed during growth experiments except for sampling. If not stated otherwise, the incubation temperature was 52.5°C, the pH was 8.5, and the irradiance was 20–50 μmol photons m^{-2} s^{-1} provided by a tungsten bulb.

High Performance Liquid Chromatography

Changes in concentrations of major nutrients (carbon, organic nitrogen and sulfur substrates) over time were followed by HPLC (UFLC module system, Shimadzu Scientific Instruments, Columbia, MD, USA). Carbon and organic sulfur substrates were analyzed using a SUPELCOGEL™ C-610H column 59320-U (30 cm × 7.8 cm ID) and a SUPELGUARD C610H 5319 (5 cm × 4.6 mm ID) guard column (Supelco, Bellefonte, PA, USA). An isocratic elution protocol with 4 mM sulfuric acid

as solvent and a total run time of 60 min and a flow rate of 0.5 ml min^{-1} were employed. The column oven had a temperature of 30°C. Prior to injection of 20-μl aliquots of medium into the column, cells and debris were removed by centrifugation for 2 min at ∼12,800 × *g*. An aliquot (1.0 ml) of the supernatant was then additionally filtered through a 0.22-μm polytetrafluoroethylene filter prior to injection. Carbon compounds of interest were detected with a refractive index detector at 210 nm or with a UV/VIS detector. The identification of each analyzed substrate was confirmed using standard solutions of the corresponding substrates tested, and which were treated in the same way as the medium before analysis.

Changes in concentrations of AAs were also analyzed using the Shimadzu HPLC system. These analyses were performed with a Kinetex 5-μm C18 100Å column (15 cm × 4.6 mm ID) protected by a SecurityGuard ULTRA cartridge UHPLC C18 for 4.6-mm ID columns (Phenomenex, Torrance, CA, USA). AAs were derivatized with phenylisothiocyanate (PITC, Edman's reagent) prior to detection at 254 nm. The derivatization procedure was based on the Thermo Scientific (Waltham, MA, USA) Example protocol for AA standard H, with some minor modifications. An aliquot (1.0 ml) of cell culture was centrifuged to remove cells and debris, and the resulting supernatant was evaporated at 55°C under a stream of nitrogen gas. The dried sample was dissolved in 100 μl of coupling solution containing 5 μl of PITC, and the solution was incubated for 10 min at room temperature in the dark. The sample was again evaporated to dryness, and the residue was dissolved in 500 μl of solvent A and evaporated again. Finally, the dried sample was dissolved in 250 μl of solvent A and filtered through a 0.22-μm polytetrafluoroethylene filter prior to injection and analysis of a 20-μl aliquot of the solution. The HPLC analysis method consisted of a 2-solvent gradient developed over a 40-min period with a flow rate of 0.5 ml min^{-1}. The initial condition was 100% solvent A, which decreased over 10 min to 82.5%, from 10–22 min to 80%, from 22–34.5 min to 30%, and from 35–40 min to 0%. Solvent A was 0.14 M sodium acetate, pH 6.2 containing 0.5 mM triethanolamine. Solvent B was a 40:60 (v/v) mixture of HPLC-grade water and acetonitrile. AAs were identified using AA standard mixture H (Thermo Scientific, Waltham, MA, USA). L-asparagine, L-glutamine, and L-tryptophan are not included in this standard mixture, and their elution times were established by analyzing each compound individually.

Pigment analyses of *C. thermophilum* strains were analyzed by a previously described method. Pigments were extracted from cells in acetone/methanol (7:2, v/v) and analyzed as described (Garcia Costas et al., 2012b).

Bacterial Growth

Growth experiments were performed with liquid and/or solidified media. The cell inoculum was equivalent to 2% v/v of the fresh medium. Depending on the particular compound being tested, nutritional tests were made in duplicate or triplicate. Tests for essentiality of vitamins and AAs included up to four serial transfers with medium lacking the compound(s) being tested. The effects of different substrates on growth were tested for both the presence and the absence of the corresponding substrate.

Growth of *C. thermophilum* in liquid cultures was monitored by the Q_y band absorption of the BChl *c* at 667 nm in a Genesys™ 10S UV/Vis scanning spectrophotometer (Thermo Scientific Rochester, NY, USA). An aliquot of cell culture (1.0 ml) was centrifuged at $\sim$15,000 $\times$ *g* for 4 min. The supernatant was removed and the cell pellet was resuspended in HPLC-grade methanol (1.0 ml) to extract the BChls. After 5-min incubation in the dark, the suspension was centrifuged for 2 min at $\sim$12,800 $\times$ *g* and immediately analyzed.

Oxygen Relationship/Requirement

The growth response of *C. thermophilum* to oxygen was tested with agar slants and plates at different oxygen concentrations [20% (atmospheric), 10%, 5%, and $\sim$0%, v/v] in the headspace of the medium, which was flushed for $\sim$1 min with the corresponding mixture of sterile nitrogen and oxygen. This was repeated every second day to ensure stable maintenance of the oxygen concentration. To produce an oxygen gradient in agar deeps, culture tubes were filled with CTM medium containing 1% (w/v) agar and were sealed with a headspace of air.

Temperature and pH Range

To determine the optimal growth temperature and the temperature range over which growth could occur, cells were grown in CTM-Medium at pH 8.5 over the temperature range of $\sim$37–70°C ($\pm$1°C). To determine the optimal pH and the pH range over which growth could occur, cells were grown in CTM-medium from pH range 4–11 at 52.5°C. The pH values of the medium were adjusted prior to autoclaving the medium and were measured again at the end of the experiment to ensure that the pH had not changed during the experiment.

16S rRNA Analyses

Genomic DNA was extracted according to the JGI-Standard protocol for DNA extraction of Gram-negative bacteria. Amplification was performed with a *C. thermophilum* specific forward primer (Cab f) and a universal reverse primer (1390r) which produced a $\sim$1300 bp 16S rRNA fragment using a standard PCR protocol as written elsewhere (Garcia Costas, 2010). Purified PCR products were sequenced with the Sanger DNA sequencing method using the same primers. Sequences were assembled, manually refined, and curated in SeqMan Pro Version 11 which is included in the Lasergene software package (DNASTAR, Madison, WI, USA). The 16S rRNA sequences were aligned with ClustalW prior determining the sequence similarities with DNAdist. Both are implemented in the BioEdit software program (Hall, 1999) and were used with default settings. The 16S rRNA sequences were deposited in GenBank under the accession numbers: KP300942- KP300947.

Results and Discussion

Historical Context for Growth and Isolation of *C. thermophilum*

Prior to the cultivation studies reported here that led to an axenic culture, *C. thermophilum* had been studied in several ways to gain information about this unusual bacterium. For example, the complete genome had been determined, and the photosynthetic apparatus including chlorosomes, BChl *a*-binding FMO protein, and type-1 photochemical reaction centers, as well as pigments, lipids and hopanoids, of *C. thermophilum* had been characterized in considerable detail (Bryant et al., 2007; Tsukatani et al., 2010, 2012; Wen et al., 2011; Garcia Costas et al., 2012a,b). These studies classified *C. thermophilum* as an anoxygenic, chlorophototrophic Acidobacterium, and the data strongly indicated that this organism relies on organic carbon source(s) (i.e., branched chain AAs), reduced sulfur source(s) and oxygen for BChl, carotenoid, and tyrosine biosynthesis. In spite of this considerable knowledge, the enrichment cultures of *C. thermophilum,* although relatively stable, only produced low and variable cell yields. The two heterotrophic contaminants severely limited further progress on biochemical, metabolic, and physiological studies.

Discoveries and descriptions of new bacteria often occur in a simple and familiar manner. A newly acquired sample is inoculated into a previously known growth medium, and those organisms that grow are screened for representatives of new species. Potentially interesting representatives are rendered axenic and are initially characterized using the same growth medium. The situation was completely different for *C. thermophilum*. *C. thermophilum* could be grown under non-optimal conditions, but almost nothing was known about the specific nutrients and growth conditions that would be required to grow this bacterium axenically. Although bacteria have considerable metabolic versatility, they typically require only a limited number of macronutrients, to provide major elements (carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur) as well as trace minerals, vitamins, or other growth factors. The challenge in this case was to establish how to provide the right substrates at the right concentrations under the right physico-chemical conditions to allow the growth *C. thermophilum*. As a starting point in obtaining an axenic culture of *C. thermophilum,* we established that phosphate, trace elements and vitamins (other than vitamin B$_{12}$; see below) did not negatively affect the growth of *C. thermophilum*, and thus the concentrations of these nutrients were not varied during the testing that ultimately led to the axenic culture. With the exception of vitamins (see below), changes in the concentrations of these substrates had little or no effects on the growth of axenic culture of *C. thermophilum*. Thus, the main focus was to identify substrates that could serve as sulfur, carbon, and nitrogen sources, respectively. Secondarily, the role of oxygen and the optimal pH and temperature conditions had to be determined.

Because *C. thermophilum* is the first anoxygenic chlorophototrophic member of a poorly characterized and highly diverse phylum, *Acidobacteria*, there were no comparable organisms that could provide guidance for its purification and cultivation. Genomic data were screened for the presence and absence of key metabolic pathways, intermediates, and transporters/permeases using the Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg/). Promising candidate substrates (as described below) were tested in cultivation experiments, and the outcomes of these cultivation experiments were refined and

used for an amended characterization of *C. thermophilum*, which ultimately led to the establishment of a defined medium.
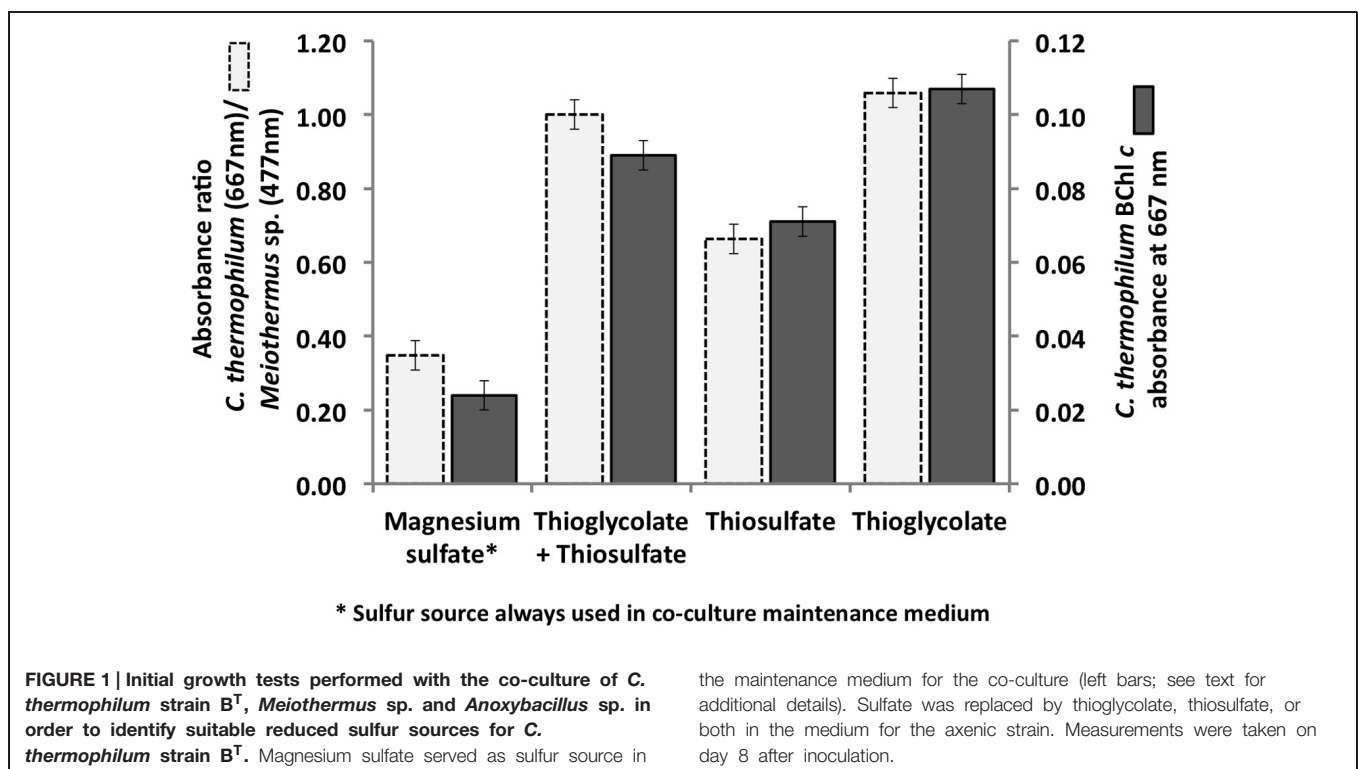
Sulfur Sources

As predicted from the genomic sequence data (Garcia Costas et al., 2012a), *C. thermophilum* was unable to use sulfate as a sulfur source because it does not have the genes for enzymes of assimilatory sulfate reduction. *C. thermophilum* instead uses reduced sulfur sources. Adding reduced sulfur sources to co-cultures of *C. thermophilum* and the two heterotrophs considerably improved the growth of *C. thermophilum* (**Figure 1**). Thioglycolate produced the greatest growth enhancement, and therefore it was used as the preferred sulfur source in a revised medium. Subsequent growth tests with axenic cultures of *C. thermophilum* reconfirmed the need for reduced sulfur sources for growth. Doubling the thioglycolate concentration clearly retarded or sometimes even inhibited the growth of *C. theromphilum*; this may have occurred because of depletion of dissolved oxygen by the reaction of the thiologlycolate with oxygen. Axenic cultures of *C. thermophilum* also grow well with L-methionine and L-cysteine/cystine as reduced sulfur sources (**Figure 2**). Thiosulfate and elemental sulfur can also serve as the sulfur source, but lower cell yields were obtained with these compounds. The use of sulfide by *C. thermophilum* is enigmatic. Cultures containing sodium sulfide did not show sustained growth, but microscopic analyses showed that sulfur globules were produced. Similar to green sulfur bacteria, these globules remained associated with the outer surfaces of cells, and suggested that sulfide oxidation occurred (data not shown). The genome lacks any known enzymes for

the oxidation of sulfide, so how sulfide oxidation occurs is not clear.

Carbon Sources

Analyses of the *C. thermophilum* genome clearly indicated that this bacterium should be a photoheterotroph, because key enzymes belonging to all known CO_2 fixation pathways were missing. Therefore, we focused on the search for suitable organic carbon sources, and we began by testing each of the compounds used in the co-culture medium individually. Acetate, butyrate, citrate, glycolate, pyruvate, lactate, and succinate were all added to the co-culture medium, because they are compounds that have been detected, or are expected to be present, in the microbial mats from which the *C. thermophilum* was enriched (Anderson et al., 1987; Bateson and Ward, 1988; also see Kim et al., 2015). The best growth yields of *C. thermophilum* were achieved with butyrate (10 mM), followed by acetate and succinate as single carbon sources in these initial tests. Although this had seemed to be a logical and reasonable place to start cultivation studies— and indeed a stable co-culture was maintained for several years on this mixture of compounds—*C. thermophilum* actually does not use any of these seven carbon sources. **Figure 3** shows data for butyrate, and these results were eventually verified by retesting each of these compounds with the axenic culture. When the consumption of these potential carbon sources was monitored by HPLC, their disappearance was actually correlated with growth of *Anoxybacillus* sp. and *Meiothermus* sp. (data not shown) but not with *C. thermophilum* (see **Figure 3**). Because the kinetics of the appearance of BChl *c* in the cultures did not match the kinetics of disappearance of any of the seven initial carbon substrates (i.e.,

FIGURE 1 | Initial growth tests performed with the co-culture of *C. thermophilum* strain B^T, *Meiothermus* sp. and *Anoxybacillus* sp. in order to identify suitable reduced sulfur sources for *C. thermophilum* strain B^T. Magnesium sulfate served as sulfur source in the maintenance medium for the co-culture (left bars; see text for additional details). Sulfate was replaced by thioglycolate, thiosulfate, or both in the medium for the axenic strain. Measurements were taken on day 8 after inoculation.
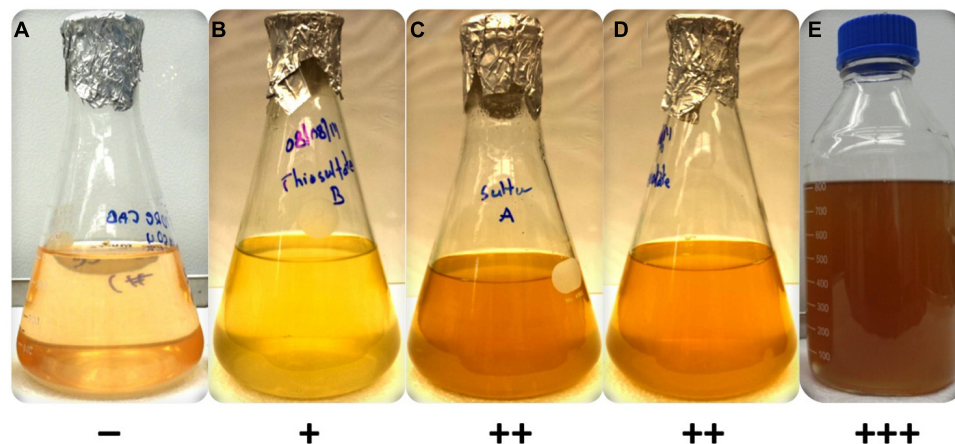
FIGURE 2 | Growth of axenic culture of *C. thermophilum* strain B^T on different sulfur sources in CTM-medium. Growth occurred only in the presence of a reduced sulfur source (**B–E**). *C. thermophilum* strain B^T did not grow with magnesium sulfate (**A**) as it relies on reduced sulfur compounds. The reduced sulfur substrates tested were (**B**) sodium thiosulfate, (**C**) sulfur (**D**) sodium thioglycolate and (**E**) L-methionine/L-cysteine. The symbols at the bottom reflect a qualitative assessment of growth of *C. thermophilum* as assessed by BChl *c* synthesis.
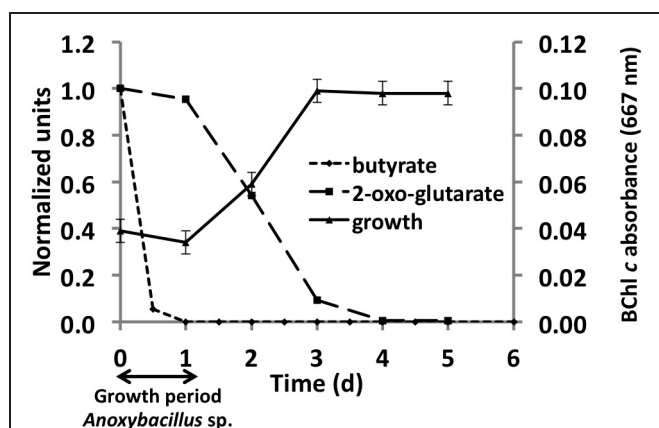

FIGURE 3 | Correlation between growth of *C. thermophilum* strain B^T and the consumption of 2-oxoglutarate and butyrate in a co-culture with *Anoxybacillus* sp. At the end of the initial growth period (∼24 h) when butyrate was completely consumed, only spores of *Anoxybacillus* sp. were observed. After *Anoxybacillus* sp. had sporulated, consumption of 2-oxoglutarate commenced and concomitantly, the BChl *c* concentration in the medium, indicative of the growth of *C. thermophilum,* increased until the 2-oxoglutarate was consumed.

acetate, butyrate, citrate, glycolate, pyruvate, lactate, and succinate), we tested 2-oxoglutarate as an alternative growth substrate. When 2-oxoglutarate was added together with one or more of the other carbon substrates, consumption of the 2-oxoglutarate began about 24 h after fresh growth medium was inoculated with the mixture of cells (**Figure 3**), and the accumulation of BChl *c* (i.e., *C. thermophilum* cells) ceased when the 2-oxoglutarate was depleted from the growth medium. Because it improved the growth of *C. thermophilum* in the co-cultures, 2-oxoglutarate was an important substrate on the way to obtaining an axenic culture of *C. thermophilum*. However, it is not essential for growth of *C. thermophilum*, because growth still occurred without added 2-oxoglutarate in subsequent experiments with axenic cultures (data not shown).

Additional growth experiments showed that bicarbonate and the AAs L-isoleucine, L-leucine, L-lysine, and L-valine are essential for growth of *C. thermophilum* (also see the section on nitrogen sources below). *C. thermophilum* did not grow when bicarbonate was omitted from the growth medium. We hypothesize that bicarbonate may be used in anaplerotic reactions to replenish the pool of certain required organic molecules (Tang et al., 2011). However, photoautotrophic growth with bicarbonate as the sole carbon source was never observed. HPLC analyses of culture medium showed that *C. thermophilum* can take up and utilize all AAs except aspartate and glutamate. Their disappearance from the medium could be correlated with the growth of *C. thermophilum* in experiments with the axenic culture (see **Figure 4**). However, it still needs to be determined whether the AAs are used primarily as carbon sources, nitrogen sources, or both.

Monosaccharides, disaccharides and polysaccharides, including mannose, glucose, fructose, rhamnose, maltose, glycogen, starch, cellulose, and chitin, were tested as additional carbon sources. Only mannose, glucose and maltose showed any effect on the growth of *C. thermophilum*. Compared to cells grown in their absence, when cells were grown in a medium containing one of these three sugars, the cells showed an obvious increase in size (data not shown). However, distinct growth rate differences in sugar-supplemented medium compared to sugar-free medium were not observed based on BChl measurements (**Figure 5**). Because the sugars were only depleted by 20–35% from the growth medium, their effects on the growth of *C. thermophilum* are still unclear. Other organic carbon sources, including pyruvate, fumarate, malate, oxaloacetate, formate, ethanol, and methanol, did not produce any clear effects on the growth of *C. thermophilum*. The BChl contents of cultures containing these
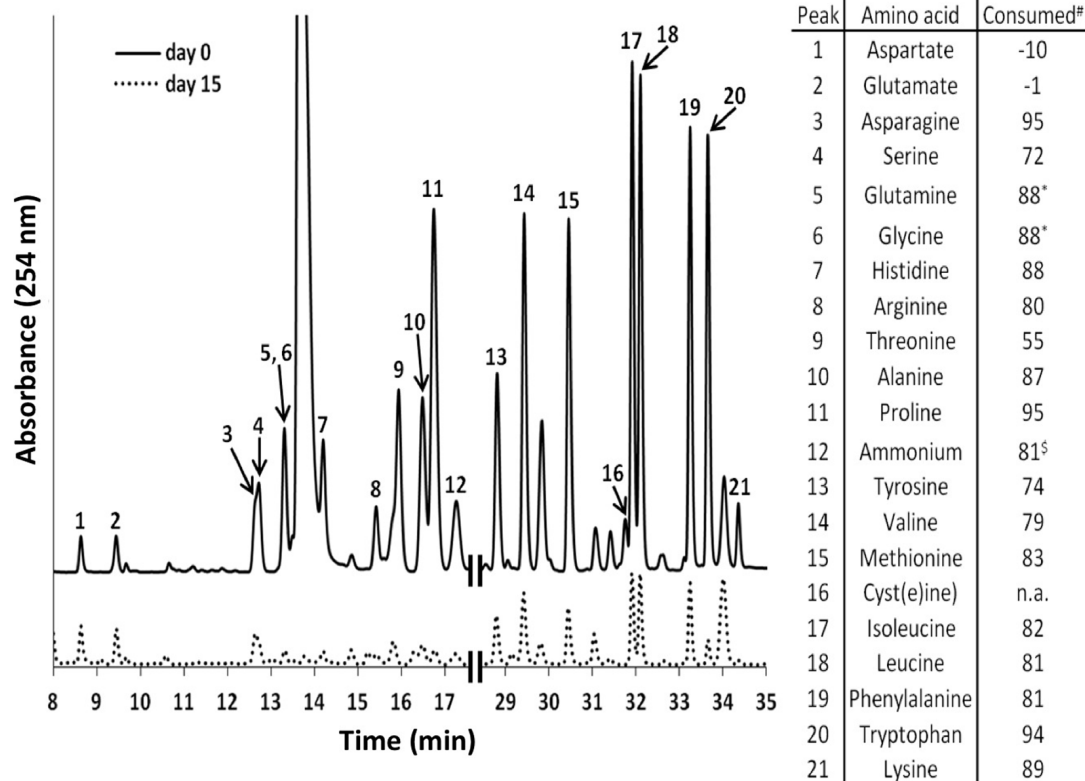
FIGURE 4 | High performance liquid chromatography elution profiles of amino acids (AA) in the growth medium of *C. thermophilum* strain B^T. The solid line shows the HPLC elution profile of the reference medium with all 20 AAs (5 mg L^{-1} of each) just prior to inoculation with *C. thermophilum* cells. The dashed line shows an elution profile for the spent medium after 15 days of growth when 19 AAs (all except L-cysteine) was added. The numbered peaks are identified in the table at the right, which shows the amount of each AA that had been consumed after 15 days. The table at the right shows the consumption or production (negative numbers for aspartate and glutamate) of each AA, which are based upon the peak heights differences between day 0 and day 15. Glutamine and glycine were not separated in our standard elution protocol and thus were calculated together. n.a., not added. $, free ammonium. Note that all AAs added were consumed except aspartate and glutamate.

compounds did not differ significantly from those of control cultures. We thus conclude that *C. thermophilum* is a nutritional specialist that is restricted to only a few carbon sources, principally AAs, and that it requires four AAs specifically as noted above.

Nitrogen Sources

The *C. thermophilum* genome does not contain genes for nitrogenase or assmililatory nitrate or nitrite reduction, and thus, not surprisingly, no growth occurred with dinitrogen or nitrate as sole N-source. Unexpectedly, *C. thermophilum* was also unable to grow with ammonium as the sole N-source. This was surprising because the genome encodes a gene predicted to encode an ammonium transporter (*amtB*; Cabther_A0161). In fact, *C. thermophilu*m cells lysed in the presence of 1 mM ammonium. Because genome analyses also predicted putative transporters for branched chain AAs, we started to use yeast extract (100 mg L^{-1}) as the nitrogen source in the growth medium. Together with the identification of a suitable S-source (thioglycolate) and carbon source (bicarbonate and 2-oxoglutarate), this N-source enabled the growth of *C. thermophilum* on plates. *C.*

thermophilum could easily be separated from the *Meiothermus* sp. by using the improved medium, but *C. thermophilum* still grew in tight association with *Anoxybacillus* sp. (**Figure 6B**), which could be eliminated after discovering the oxygen sensitivity of *C. thermophilum* (see discussion of Oxygen Relationships below).

C. thermophilum grew well on plates and in liquid culture when yeast extract was replaced by 100 mg L^{-1} peptone, which confirmed that *C. thermophilum* requires AAs for growth. In order to produce a completely defined growth medium for *C. thermophilum,* and to determine which AAs are utilized by *C. thermophilum,* we conducted growth experiments with axenic cultures, in which we tested different combinations of AAs, e.g., alpha-keto acids of the branched chain AAs plus ammonium, branched chain AAs only, and AAs without branched chain AAs. We also tested the essentiality of all 20 AAs individually. Growth of *C. thermophilum* is strictly dependent upon the branched chain AAs, L-isoleucine, L-leucine, and L-valine (**Figure 7**), and interestingly, L-lysine is also essential. When these four AAs were omitted from any growth medium, minimal or no growth of *C. thermophilum* occurred after the first

FIGURE 5 | Growth curves of *C. thermophilum* strain B[T] from sugar utilization tests in liquid culture. Growth is plotted as a function of BChl *c* absorbance at 667 nm over time.
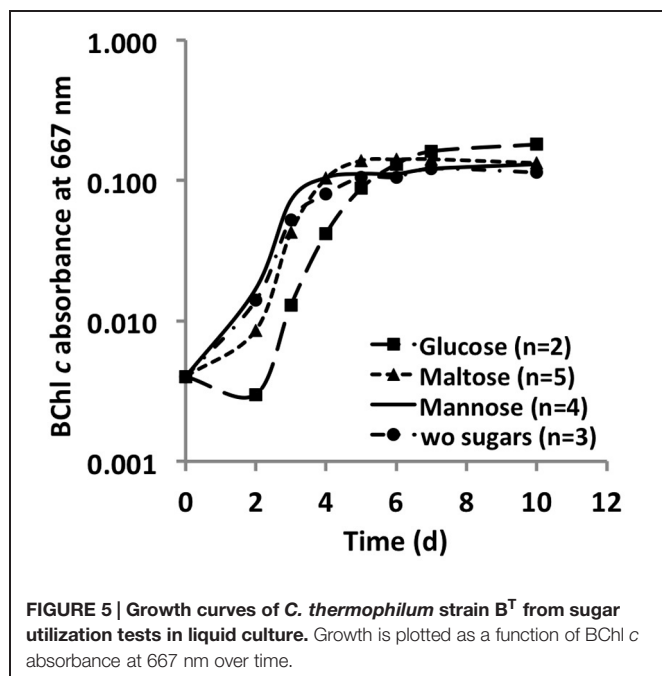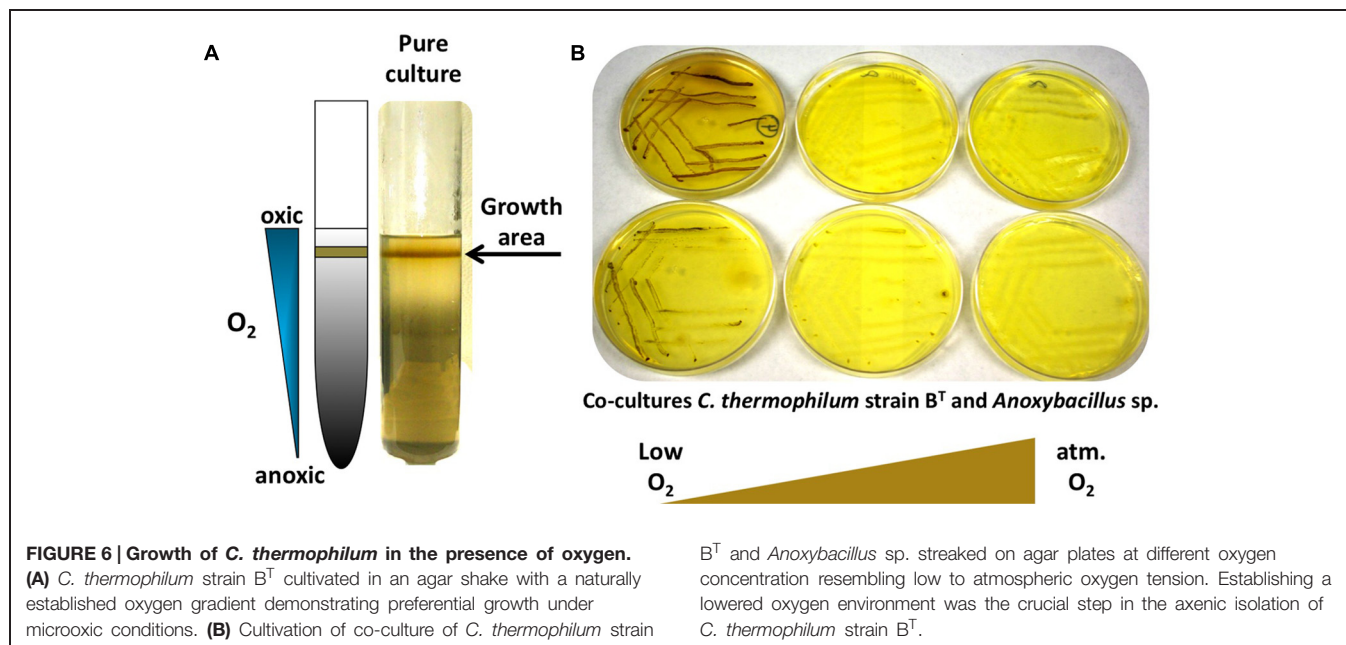
transfer. The essentiality of L-isoleucine, L-leucine, L-lysine, and L-valine is consistent with the genomic sequence data that had predicted that *C. thermophilum* would be unable to synthesize these four AAs (Garcia Costas et al., 2012a). On the other hand *C. thermophilum* still grew after the third transfer into medium that contained these four AAs as sole nitrogen source, confirming that these AAs are capable of providing all nitrogen required for growth. All other AAs can be synthesized by *C. thermophilum,* which could be demonstrated in experiments in which one or more of the other sixteen AAs were

omitted. In addition, HPLC analyses of AA utilization revealed that *C. thermophilum* is able to metabolize at least 18 of the 20 common AAs (**Figure 4**). Over a period of 14 days all AAs were consumed to varying extents except aspartic acid and glutamate, which actually seemed to be produced and excreted rather than being consumed. AA uptake ranged from a minimal value of 55% (L-threonine) to ∼95% (L-proline). As observed for the S- and C-sources, the response to N-source of *C. thermophilum* showed an oligotrophic behavior that is common among Acidobacteria (Eichorst et al., 2007; Fierer et al., 2007). Growth was not improved by simply adding a higher concentration of AAs to the medium at the start of cultivation. However, cultures produced higher biomass when they were supplemented with AAs over the time course for growth (**Figure 8**) Although they are common, natural, nitrogen-rich substances, putrescine, betaine, and DNA were not utilized as N-sources by *C. thermophilum*.

Oxygen Relationship

Establishing the growth relationship of *C. thermophilum* to oxygen was one of the biggest challenges, but studies with oxygen provided one of the key insights that led to an axenic culture. Growth tests clearly demonstrated that *C. thermophilum* preferred low oxygen concentrations for growth and maintenance in the laboratory. This correlates well with knowledge that *C. thermophilum* does not grow near the surface of the microbial mat community from which it is derived but that it grows near the bottom of the photic layer of these mats (Liu et al., 2012). No growth occurred under anoxic conditions (in a chamber with an atmosphere of H_2, CO_2, and N_2 (10:10:80 vol/vol/vol)), or in fully oxygenated cultures that were vigorously shaken. *C. thermophilum* showed the typical growth pattern observed for microaerophiles in agar deeps (**Figure 6A**). Growth only occurred in the narrow interface between the oxic and anoxic

FIGURE 6 | Growth of *C. thermophilum* in the presence of oxygen. **(A)** *C. thermophilum* strain B[T] cultivated in an agar shake with a naturally established oxygen gradient demonstrating preferential growth under microoxic conditions. **(B)** Cultivation of co-culture of *C. thermophilum* strain B[T] and *Anoxybacillus* sp. streaked on agar plates at different oxygen concentration resembling low to atmospheric oxygen tension. Establishing a lowered oxygen environment was the crucial step in the axenic isolation of *C. thermophilum* strain B[T].

Essential and non-essential amino acids for *C. thermophilum* strain B[T]

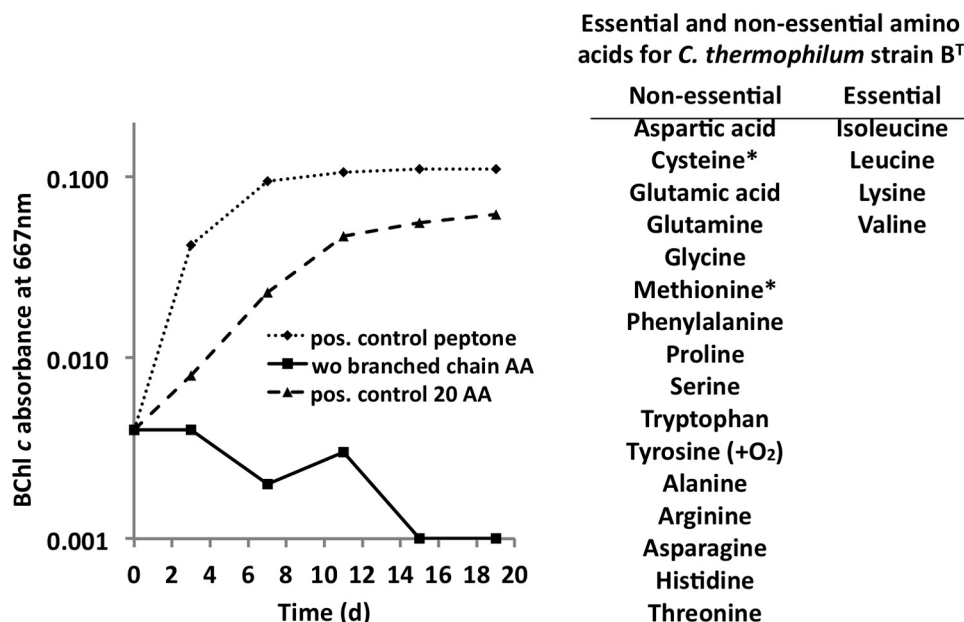| Non-essential	Essential
Aspartic acid	Isoleucine
Cysteine*	Leucine
Glutamic acid	Lysine
Glutamine	Valine
Glycine	
Methionine*	
Phenylalanine	
Proline	
Serine	
Tryptophan	
Tyrosine (+O2)	
Alanine	
Arginine	
Asparagine	
Histidine	
Threonine	

FIGURE 7 | Branched chain AAs are essential for growth of *C. thermophilum*. Growth of *C. thermophilum* strain B[T] with peptone, 20 common AAs and common AA without the branched chain AAs L-isoleucine, L-leucine, and L-valine. Note that no growth occurred in the absence of branched chain AAs. The asterisks for cysteine and methionine indicate that one of these AAs is essential in the absence of a reduced sulfur source. Tyrosine is essential under very low oxygen concentrations.
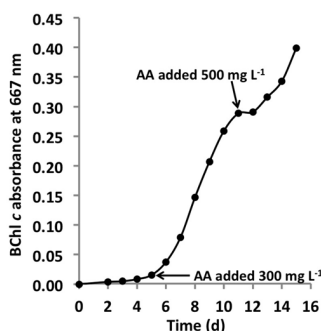

FIGURE 8 | Growth stimulation by addition of AAs. Growth of *C. thermophilum* strain B[T] in CTM-medium that was supplemented two times with a mixture of the 20 common AA, at concentrations of 300 mg L^{-1} and 500 mg L^{-1} as indicated by the arrows. Note, that each addition of AAs (arrows enhanced) growth. *C. thermophilum* strain B[T] typically reaches stationary phase at BChl *c* absorbance values of 0.1–0.15 without supplemental feeding.

FIGURE 9 | Vitamin requirements of *C. thermophilum* strain B[T]. The chart shows the growth with different combinations of vitamins after second transfer. Cultures containing Vitamin B$_{12}$ show normal growth. Cultures in medium lacking vitamin B$_{12}$ already show very weak growth after the second transfer. Thus, *C. thermophilum* only requires vitamin B$_{12}$.

regions of the agar deep. Interestingly, *C. thermophilum* survived long-term exposures to both fully oxic and anoxic conditions, but survival was distinctly longer in oxygenated medium. Because of the technical difficulty of providing alternating microoxic and anoxic conditions, we did not test if *C. thermophilum* could grow better under alternating oxygen concentrations, as occurs over each diel cycle in its natural habitat (Liu et al., 2011, 2012). In contrast to other bacteria that contain homodimeric type-1 reaction centers (green sulfur bacteria and heliobacteria; Bryant and Liu, 2013), which have reaction centers that are highly sensitive to oxygen, *C. thermophilum* requires oxygen for the biosynthesis of (B)Chls and carotenoids and also requires oxygen for the synthesis of tyrosine from phenylalanine (Garcia Costas et al., 2012a).

Vitamins

The genome predicts that *C. thermophilum* requires vitamin B$_{12}$ for L-methionine synthesis (Garcia Costas et al., 2012a), and because most of the genes for vitamin B$_{12}$ synthesis are missing in the genome, it was not surprising to establish that vitamin
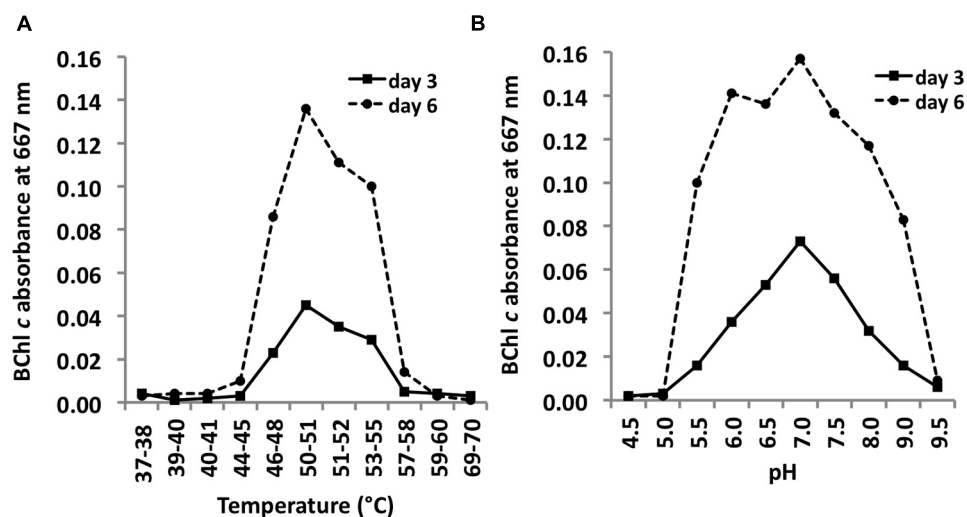
FIGURE 10 | Temperature and pH optima for growth. Temperature **(A)** and pH ranges **(B)** for growth of *C. thermophilum* strain B[T]. Cells were grown in CTM-Medium, pH 8.5 for 6 days, and relative cell yields on day 3 and day 6 were determined by the amount of BChl *c* that was synthesized (see Materials and Methods). Optimal growth was observed at 51–52°C; a broad pH optimum centered at about pH 7 was observed.

FIGURE 11 | Appearance of liquid cultures of *C. thermophilum* strain B[T] and strain E. Note the cell aggregates and clumpy growth of strain E compared to the homogeneous cell suspension of strain B[T]. Both cultures were shaken prior imaging. Cells of *C. thermophilum* do not float during growth but settle to the bottom of growth vessel.

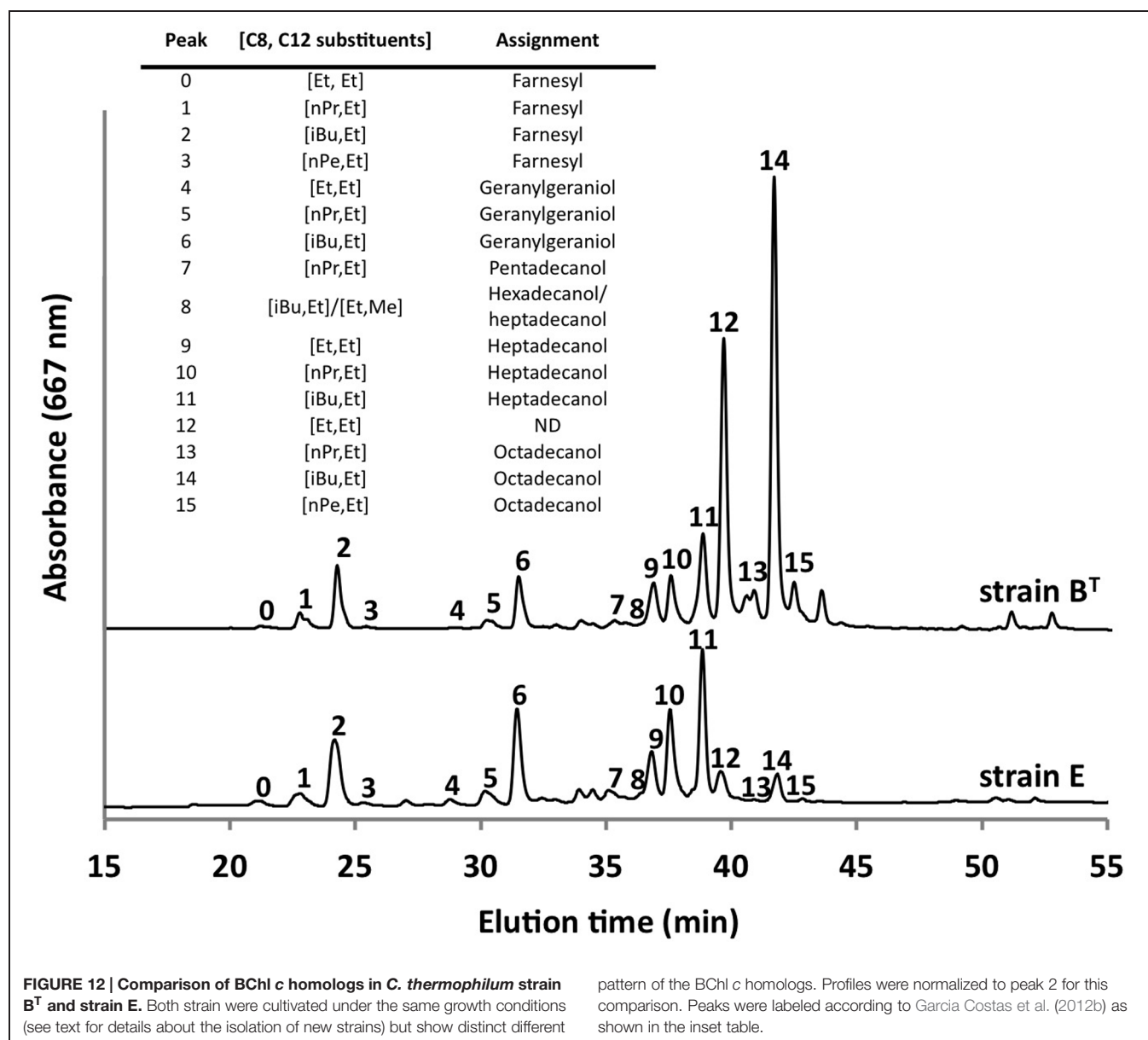B12 was essential for maintaining the growth of *C. thermophilum*. Reduced growth in the absence of vitamin B12 was obvious after the second transfer of cells into medium free of vitamin B12 (**Figure 9**). Cells starved for vitamin B12 exhibited very weak fluorescence from BChls by epifluorescence microscopy (data not shown). This is consistent with the important role that S-adenosylmethionine plays in Chl biosynthesis in general and biosynthesis of BChl *c* methylation homologs specifically (Bryant et al., 2007; Garcia Costas et al., 2011). When the 13-vitamin mix was omitted from the growth medium no obvious effect on growth was noted after four serial transfers into medium free of vitamins other than vitamin B12. This observation confirms that vitamin B12 is the only vitamin required for growth of *C. thermophilum*.

Light Requirements

C. thermophilum is a phototrophic bacterium that produces chlorosomes, and it showed the best growth under low continuous irradiance, ∼20–50 μmol photons m^{-2} s^{-1} (from a tungsten source). Light intensities above 50 μmol photons m^{-2} s^{-1} led to cell lysis. Very weak growth occurred in medium containing 2.5 mM 2-oxoglutarate and 2.5 mM mannose (together with AAs) in the dark, but otherwise, growth was dependent upon light in all experiments. Several explanations for the preference for lower irradiance can be given. Like green sulfur bacteria and green filamentous anoxygenic phototrophs like *Chloroflexus* sp., *C. thermophilum* produces chlorosomes (Bryant et al., 2012). Chlorosomes are the most efficient light harvesting organelles in nature (Frigaard and Bryant, 2006), and they have evolved to allow bacteria to obtain sufficient energy for phototrophic metabolism under very low irradiance conditions. Because *C. thermophilum* obviously does not generate biomass from CO2 fixation and does not fix nitrogen, both of which are energy intensive processes, the energy needed from light is not particularly

FIGURE 12 | Comparison of BChl *c* homologs in *C. thermophilum* strain B[T] and strain E. Both strain were cultivated under the same growth conditions (see text for details about the isolation of new strains) but show distinct different pattern of the BChl *c* homologs. Profiles were normalized to peak 2 for this comparison. Peaks were labeled according to Garcia Costas et al. (2012b) as shown in the inset table.

high and the metabolism of organic molecules may also serve as an energy source via respiration. All genes required to produce an aerobic respiration chain are found in the genome (Garcia Costas et al., 2012a). A third reason may be that too many reactive oxygen species are produced at high irradiance conditions, and these could damage the phototrophic apparatus of *C. thermophilum*.

Temperature and pH Range and Optima

C. thermophilum strain B[T] (Tank and Bryant, 2015) grew at temperatures between 44 and 58°C with a T_{opt} of ~51°C (**Figure 10A**). This relatively narrow temperature range supports the findings from Miller et al. (2009) that different ecotypes of *C. thermophilum* are adapted to specific temperatures. Miller et al. (2009) found *C. thermophilum* growing at temperatures of 38–68°C in White Creek and Garcia Costas (2010) found

C. thermophilum in several mat samples collected at temperatures from 34 to 68°C in Yellowstone National Park, WY, USA. *C. thermophilum* can be classified as a moderately thermophilic bacterium.

C. thermophilum grew between pH 5.5 and 9.5 and exhibited a broad optimum at circum-neutral pH (**Figure 10B**). *C. thermophilum* is well adapted to the pH value of its *in situ* habitat, which ranges from circum-neutral in the morning to around 9.5 in the late afternoon in the microbial mats at Mushroom and Octopus Spring (Jensen et al., 2011).

Isolation of New *C. thermophilum* Strains

From previous studies conducted at different hot springs in Yellowstone National Park, it was known that different representatives of *C. thermophilum* with different temperature

FIGURE 13 | Comparison of carotenoids in *C. thermophilum* strain B^T and strain E. Both strains were cultivated under the same growth conditions (see text for details about the isolation of new strains), but they show differences in carotenoid content. The HPLC profile was normalized to peak 5 to facilitate the comparison. Peaks were labeled according to Garcia Costas et al. (2012b) as indicated in the inset table.

optima occur in these chlorophototrophic microbial mats (Miller et al., 2009; Garcia Costas, 2010; Ross et al., 2012). More distantly related sequences 16S rRNA sequences have been recovered from microbial communities associated with hot springs in Tibet and Thailand (Kanokratana et al., 2004; Yim et al., 2006; Lau et al., 2009). These facts encouraged us to test the suitability of our defined CTM-medium, together with the experimental growth conditions used for lab strain B^T, to isolate new *C. thermophilum* strains directly from the environment.

Using the optimized growth medium and growth conditions reported here, it was possible to grow new *C. thermophilum* strains and isolate representatives from mat samples taken at 52 and 60°C samples in pure culture. In liquid medium the new strains form large cell aggregates and clumps, whereas strain B^T grows as a homogenous cell suspension (**Figure 11**). 16S rRNA sequence analyses showed that one new isolate is 100% identical to strain B^T, whereas other representatives were only 99% identical in sequence to the lab strain B^T (data not shown). Pigment analyses and growth tests with strains only showing 99% similarity to the lab strain B^T, respectively, also showed other differences from strain B^T. Under the same growth conditions, CTM-medium 52°C, pH 7.0, 50 μmol photons m^{-2} s^{-1} (from a tungsten source) and microoxic growth conditions, the new strains showed a distinctly different composition of BChl *c* homologs (**Figure 12**) and carotenoids (**Figure 13**) compared to the type strain B^T. Consistent with the findings of Miller et al. (2009) and Garcia Costas (2010), initial experiments confirmed

that different temperature ecotypes of *C. thermophilum* may exist in nature. The new strain can grow at temperatures as low as 39°C whereas strain B^T is unable to grow at temperatures below 44°C. Both strains have a similar growth optimum, ~50°C, and a similar upper limit of ~58°C under the conditions tested. In addition to ongoing growth and physiological testing, we plan to sequence the genome of at least one new strain and compare it with that of strain B^T in future studies. Besides the finding that the CTM-medium can be used for direct isolation of *C. thermophilum* from the environment, the new isolates will help us understand this unusual bacterium better.

Conclusion

Axenic growth of *C. thermophilum* required the identification of essential nutrients on the one hand and the appropriate oxygen concentration on the other. The present study demonstrates that *C. thermophilum* is a microaerophilic, moderately thermophilic, anoxygenic, photoheterotrophic eubacterium. Its growth is absolutely dependent on several medium components, which include a reduced sulfur source, bicarbonate, all branched chain AAs, L-lysine and vitamin B$_{12}$.

This study is an excellent example of how classical microbiology, in combination with modern –omics methods (Bryant et al., 2007; Liu et al., 2011, 2012; Garcia Costas et al., 2012a), led to the discovery and eventually to a fairly comprehensive characterization of this previously unknown bacterium. It was possible to develop a nutritionally defined medium to isolate *C. thermophilum* in axenic culture (Tank and Bryant, 2015). The growth conditions were further refined by determining the optimum pH, temperature and light intensities of *C. thermophilum*. The preceding –omics studies served as basis for developing specific hypotheses about the possible physiology of *C. thermophilum*, which were then tested with classical microbiology methods. This interplay of two different approaches led to the stepwise elimination of the co-culture contaminants, and in parallel we learned more about the physiology of *C. thermophilum*. Retrospectively, *C. thermophilum* is not a particularly extraordinary bacterium concerning its nutritional requirements and growth conditions. The major interest in *C. thermophilum* is due to it being the first chlorophototrophic member of the Acidobacteria. Furthermore, it shows characteristics that are usually either found in organisms that live under oxic or obligately anoxic conditions—combined into one organism. *C. thermophilum* uses AAs, peptides or proteins as a nearly universal supplier of major elements required for life. AAs can serve as carbon, nitrogen, sulfur and perhaps even energy source for *C. thermophilum*.

The ecological importance of *C. thermophilum* in the microbial mats it inhabits is still speculative, and this aspect was not a part of this study. According to previous metagenomic studies (Liu et al., 2011) and other molecular surveys (Miller et al., 2009), *C. thermophilum* represents only 5–10% of the chlorophototrophic mat communities in the alkaline hot spring mats it inhabits in Yellowstone National Park. *C. thermophilum* certainly benefits from other members of the mat community, because it relies on substrates produced by organisms in the mat,

including AAs, reduced sulfur compounds, and CO_2. In addition, *Synechococcus* spp. and *Roseiflexus* spp. are known to synthesize vitamin B_{12}, which is essential for growth of *C. thermophilum* (Garcia Costas et al., 2011). On the other hand, *C. thermophilum* shows no biotin auxotrophy and could potentially provide this vitamin to the mat-dominant *Synechococcus* spp. for which biotin is essential. Future experiments will test this hypothesis by co-culturing *C. thermophilum* and *Synechococcous* sp. in a medium lacking biotin.

Now that *C. thermophilum* can easily be cultivated in axenic culture, it can be used in further studies to answer questions about the photosynthetic apparatus and about its ecological role in mats. This study nicely demonstrates that one can sometimes cultivate previously uncultivated organisms in axenic culture if one knows or can demonstrate the specific physiological and nutritional needs of the organism. It was clearly shown here that *C. thermophilum* has very simple nutrient requirements and that oxygen concentration played a crucially important role in its growth. The reason for the specific oxygen needs is the balance between oxygen-sensitive traits, for example, the type-1 photosynthetic reaction center, ferredoxin, 2-oxoglutarate ferredoxin oxidoreductase (Cabther_B0326), and possibly other enzymes containing iron-sulfur clusters on one hand, and components whose synthesis is oxygen-dependent, for example, Chl and BChls, oxygen-containing ketocarotenoids or tyrosine biosynthesis on the other.

An unexpected requirement for the cultivation of *C. thermophilum* was the need for bicarbontate, which possibly originates from the presence of CO_2-incorporating enzymes, for example, the 2-oxoglutarate ferredoxin oxidoreductase (KDO, Cabther_B0326), phosphoenolpyruvate carboxylase (Cabther_A2240) or other enzymes involved in anaplerotic reactions. Heterotrophic CO_2 fixation is also known to occur in aerobic anoxygenic phototrophic purple bacteria, which can produce up to 10% to 15% of their biomass from this process (Tang et al., 2009; Hauruseu and Koblížek, 2012). The contribution of bicarbonate/CO_2 to the growth of *C. thermophilum* is not yet known, but it could be substantial. We hypothesize that succinyl-CoA is a major product of the degradation of branched chain AAs, and that this metabolite is then carboxylated by KDO to produce 2-oxoglutarate, which is a key precursor metabolite for the synthesis of proteins and (B)Chls. Because of the likely importance of this route for the production of metabolic precursors, it would be interesting to determine the oxygen sensitivity of KDO, which might partly explain the preference of this organism for microoxic conditions. The findings reported here concerning the oxygen relations of *C. thermophilum* suggest that oxygen concentration could prove to be a key factor in the isolation and cultivation of many other bacteria and archaea that have not yet been grown axenically.

Acknowledgments

References

Allewalt, J. P., Bateson, M. M., Revsbech, N.-P., Slack, K., and Ward, D. M. (2006). Effect of temperature and light on growth and photosynthesis of *Synechococcus* isolates typical of those predominating in the Octopus spring microbial mat community of Yellowstone National Park. *Appl. Environ. Microbiol.* 72, 544–550. doi: 10.1128/AEM.72.1.544-550.2006

Anderson, K. L., Tayne, T. A., and Ward, D. M. (1987). Formation and fate of fermentation products in hot-spring cyanobacterial mats. *Appl. Environ. Microbiol.* 53, 2343–2352.

Barns, S. M., Takala, S. L., and Kuske, C. R. (1999). Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment. *Appl. Environ. Microbiol.* 65, 1731–1737.

Bateson, M. M., and Ward, D. M. (1988). Photoexcretion and fate of glycolate in a hot-spring cyanobacterial mat. *Appl. Environ. Microbiol.* 54, 1738–1743.

Bhaya, D., Grossman, A. R., Steunou, A.-S., Khuri, N., Cohan, F. M., Hamamura, N., et al. (2007). Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. *ISME J.* 1, 703–713. doi: 10.1038/ismej.2007.46

Bryant, D. A., and Liu, Z. (2013). "Green bacteria: insights into green bacterial evolution through genomic analyses," in *Advances in Botanical Research*, Vol. 66, *Genome Evolution of Photosynthetic Bacteria*, ed. J. T. Beatty (New York, NY: Elsevier), 99–150. doi: 10.1016/B978-0-12-397923-0.00004-7

Bryant, D. A., Garcia Costas, A. M., Maresca, J. A., Chew, A. G., Klatt, C. G., Bateson, M. M., et al. (2007). *Candidatus* Chloracidobacterium thermophilum: an aerobic phototrophic acidobacterium. *Science* 317, 523–526. doi: 10.1126/science.1143236

Bryant, D. A., Liu, Z., Li, T., Zhao, F., Costas, G., Klatt, C. G., et al. (2012). "Comparative and functional genomics of anoxygenic green bacteria from the taxa Chlorobi, Chloroflexi, and Acidobacteria," in *Advances in Photosynthesis and Respiration,* Vol. 35, *Functional Genomics and Evolution of Photosynthetic Systems*, eds R. L. Burnap and W. Vermaas (Dordrecht: Springer), 47–102. doi: 10.1007/978-94-007-1533-2_3

Crowe, M. A., Power, J. F., Morgan, X. C., Dunfield, P. F., Lagutin, K., Rijpstra, W. I., et al. (2014). *Pyrinomonas methylaliphatogenes* gen. nov., sp. nov.,

a novel group 4 thermophilic member of the phylum Acidobacteria from geothermal soils. *Int. J. Syst. Evol. Microbiol.* 64, 220–227. doi: 10.1099/ijs.0.055079-0

Eichorst, S. A., Breznak, J. A., and Schmidt, T. M. (2007). Isolation and characterization of soil bacteria that define *Teniglobus* gen. nov., in the phylum Acidobacteria. *Appl. Environ. Microbiol.* 73, 2708–2717. doi: 10.1128/AEM.02140-06

Fierer, N., Bradford, M. A., and Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology* 88, 1354–1364. doi: 10.1890/05-1839

Frigaard, N.-U., and Bryant, D. A. (2006). "Chlorosomes: antenna organelles in green photosynthetic bacteria," in *Microbiology Monographs, Vol. 2, Complex Intracellular Structures in Prokaryotes,* ed. J. Shively (Berlin: Springer), 79–114. doi: 10.1007/7171_021

Foesel, B. U., Rohde, M., and Overmann, J. (2013). Blastocatella fastidiosa gen. nov., sp. nov., isolated from semiarid savanna soil-the first described species of Acidobacteria subdivision 4. *Syst. Appl. Microbiol.* 36, 82–89. doi: 10.1016/j.syapm.2012.11.002

Garcia Costas, A. M. (2010). *Isolation and Characterization of Candidatus Chloracidobacterium thermophilum.* Ph. D. thesis, dissertation, The Pennsylvania State University, University Park, PA, USA.

Garcia Costas, A. M., Liu, Z., Tomsho, L. P., Schuster, S. C., Ward, D. M., and Bryant, D. A. (2012a). Complete genome of *Candidatus* Chloracidobacterium thermophilum, a chlorophyll-based photoheterotroph belonging to the phylum Acidobacteria. *Environ. Microbiol.* 14, 177–190. doi: 10.1111/j.1462-2920.2011.02592.x

Garcia Costas, A. M., Tsukatani, Y., Rijpstra, W. I., Schouten, S., Welander, P. V., Summons, R. E., et al. (2012b). Identification of the bacteriochlorophylls, carotenoids, quinones, lipids, and hopanoids of "*Candidatus* Chloracidobacterium thermophilum." *J. Bacteriol.* 194, 1158–1168. doi: 10.1128/JB.06421-11

Garcia Costas, A. M., Tsukatani, Y., Romberger, S. P., Oostergetel, G. T., Boekema, E. J., Golbeck, J. H., et al. (2011). Ultrastructural analysis and identification of envelope proteins of "*Candidatus* Chloracidobacterium thermophilum" chlorosomes. *J. Bacteriol.* 193, 6701–6711. doi: 10.1128/JB.06124-11

Hall, T. A. (1999). BioEdit: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* 41, 95–98.

Hauruseu, D., and Koblížek, M. (2012). Influence of light on carbon utilization in aerobic anoxygenic phototrophs. *Appl. Environ. Microbiol.* 78, 7414–7419. doi: 10.1128/AEM.01747-12

Huber, K. J., Wust, P. K., Rhode, M., Overmann, J., and Foesel, B. U. (2014). *Aridibacter famidurans* and *Aridibacter kavangonensis*, 2 novel species of Acidobacteria subdivision 4 isolated from semiarid savanna soil. *Int. J. Syst. Evol. Microbiol.* 64, 1866–1875. doi: 10.1099/ijs.0.06023610

Jensen, S. I., Steunou, A.-S., Bhaya, D., Kühl, M., and Grossman, A. R. (2011). In situ dynamics of O2, pH and cyanobacterial transcripts associated with CCM, photosynthesis and detoxification of ROS. *ISME J.* 5, 317–328. doi: 10.1038/ismej.2010.131

Kanokratana, P., Chanapan, S., Pootanakit, K., and Eurwilaichitr, L. (2004). Diversity and abundance of bacteria and Archaea in the Bor Khlueng hot spring in Thailand. *J. Basic Microbiol.* 44, 430–444. doi: 10.1002/jobm.200410388

Kim, Y.-M., Nowack, S., Olsen, M. T., Becraft, E. D., Wood, J. M., Thiel, V., et al. (2015). Diel metabolomics analysis of a hot spring chlorophototrophic microbial mat leads to new hypotheses of community member metabolisms. *Front. Microbiol.* 6:209. doi: 10.3389/fmicb.2015.00209

Lau, M. C., Aitchison, J. C., and Pointing, S. B. (2009). Bacterial community composition in thermophilic microbial mats from five hot springs in central Tibet. *Extremophiles* 13, 139–149. doi: 10.1007/s00792-008-0205-3

Liu, Z., Klatt, C. G., Ludwig, M., Rusch, D. B., Jensen, S. I., Kuhl, M., et al. (2012). 'Candidatus Thermochlorobacter aerophilum:' an aerobic chlorophotoheterotrophic member of the phylum Chlorobi defined by metagenomics and metatranscriptomics. *ISME J.* 6, 1869–1882. doi: 10.1038/ismej.2012.24

Liu, Z., Klatt, C. G., Wood, J. M., Rusch, D. B., Ludwig, M., Wittekindt, N., et al. (2011). Metatranscriptomic analyses of chlorophototrophs of a hot-spring microbial mat. *ISME J.* 5, 1279–1290. doi: 10.1038/ismej.2011.37

Miller, S. R., Strong, A. L., Jones, K. L., and Ungerer, M. C. (2009). Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Appl. Environ. Microbiol.* 75, 4565–4572. doi: 10.1128/AEM.02792-08

Ross, K. A., Feazel, L. M., Robertson, C. E., Fathepure, B. Z., Wright, K. E., Turk-Macleod, R. M., et al. (2012). Phototrophic phylotypes dominate mesothermal microbial mats associated with hot springs in Yellowstone National park. *Microb. Ecol.* 64, 162–170. doi: 10.1007/s00248-012-0012-3

Tang, K.-H., Feng, X., Tang, Y. J., and Blankenship, R. E. (2009). Carbohydrate metabolism and carbon fixation in *Roseobacter denitrificans* OCh114. *PLoS ONE* 4:e7233. doi: 10.1371/journal.pone.0007233

Tang, K.-H., Tang, Y. J., and Blankenship, R. E. (2011). Carbon metabolic pathways in phototrophic bacteria and their broader evolutionary implications. *Front. Microbiol.* 2:165. doi: 10.3389/fmicb.2011.00165

Tank, M., and Bryant, D. A. (2015). *Chloracidobacterium thermophilum* gen. nov., sp. nov.: an anoxygenic, microaerophilic and chlorophotoheterotrophic acidobacterium. *Int. J. Syst. Evol. Microbiol.* doi: 10.1099/ijs.000113 [Epub ahead of print].

Tsukatani, Y., Romberger, S. P., Golbeck, J. H., and Bryant, D. A. (2012). Isolation and characterization of homodimeric type-I reaction center complex from *Candidatus* Chloracidobacterium thermophilum, an aerobic chlorophototroph. *J. Biol. Chem.* 287, 5720–5732. doi: 10.1074/jbc.M111.323329

Tsukatani, Y., Wen, J., Blankenship, R. E., and Bryant, D. A. (2010). Characterization of the FMO protein from the aerobic chlorophototroph, *Candidatus* Chloracidobacterium thermophilum. *Photosynth. Res.* 104, 201–209. doi: 10.1007/s11120-009-9517-0

Wen, J., Tsukatani, Y., Cui, W., Zhang, H., Gross, M. L., Bryant, D. A., et al. (2011). Structural model and spectroscopic characteristics of the FMO antenna protein from the aerobic chlorophototroph, *Candidatus* Chloracidobacterium thermophilum. *Biochim. Biophys. Acta* 1807, 157–164. doi: 10.1016/j.bbabio.2010.09.008

Yim, L. C., Hongmei, J., Aitchison, J. C., and Pointing, S. B. (2006). Highly diverse community structure in a remote central Tibetan geothermal spring does not display monotonic variation to thermal stress. *FEMS Microbiol. Lett.* 57, 80–91. doi: 10.1111/j.1574-6941.2006.00104.x

Zeng, Y., Feng, F., Medová, H., Dean, J., and Koblizek, M. (2014). Functional type-2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7795–7800. doi: 10.1073/pnas.1400295111

Diel metabolomics analysis of a hot spring chlorophototrophic microbial mat leads to new hypotheses of community member metabolisms

Young-Mo Kim[1], Shane Nowack[2,3], Millie T. Olsen[2], Eric D. Becraft[2†], Jason M. Wood[2], Vera Thiel[4], Isaac Klapper[3,5], Michael Kühl[6,7], James K. Fredrickson[1], Donald A. Bryant[4,8], David M. Ward[2] and Thomas O. Metz[1*]

[1] Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA, [2] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA, [3] Department of Mathematical Sciences, Montana State University, Bozeman, MT, USA, [4] Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA, [5] Department of Mathematics, Temple University, Philadelphia, PA, USA, [6] Marine Biological Section, Department of Biology, University of Copenhagen, Helsingør, Denmark, [7] Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, Ultimo, NSW, Australia, [8] Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA

Dynamic environmental factors such as light, nutrients, salt, and temperature continuously affect chlorophototrophic microbial mats, requiring adaptive and acclimative responses to stabilize composition and function. Quantitative metabolomics analysis can provide insights into metabolite dynamics for understanding community response to such changing environmental conditions. In this study, we quantified volatile organic acids, polar metabolites (amino acids, glycolytic and citric acid cycle intermediates, nucleobases, nucleosides, and sugars), wax esters, and polyhydroxyalkanoates, resulting in the identification of 104 metabolites and related molecules in thermal chlorophototrophic microbial mat cores collected over a diel cycle in Mushroom Spring, Yellowstone National Park. A limited number of predominant taxa inhabit this community and their functional potentials have been previously identified through metagenomic and metatranscriptomic analyses and *in situ* metabolisms, and metabolic interactions among these taxa have been hypothesized. Our metabolomics results confirmed the diel cycling of photorespiration (e.g., glycolate) and fermentation (e.g., acetate, propionate, and lactate) products, the carbon storage polymers polyhydroxyalkanoates, and dissolved gasses (e.g., H_2 and CO_2) in the waters overlying the mat, which were hypothesized to occur in major mat chlorophototrophic community members. In addition, we have formulated the following new hypotheses: (1) the morning hours are a time of biosynthesis of amino acids, DNA, and RNA; (2) photo-inhibited cells may also produce lactate via fermentation as an alternate metabolism; (3) glycolate and lactate are exchanged among *Synechococcus* and *Roseiflexus* spp.; and (4) fluctuations in many metabolite pools (e.g., wax esters) at different times of day result from species found at different depths within the mat responding to temporal differences in their niches.

Keywords: gas chromatography-mass spectrometry, metabolomics, microbial mats, polyhydroxyalkanoates, *Roseiflexus*, *Synechococcus*, wax esters

Introduction

Microbial communities inhabiting extreme environments in Yellowstone National Park (YNP) have been investigated for more than half a century (Brock, 1972, 1998). In particular, chlorophototrophic (i.e., chlorophyll-based phototrophs) microbial mat communities present in the effluent channels of Octopus Spring and Mushroom Spring within the Lower Geyser Basin have been intensively studied (Brock, 1978; Ward et al., 2012). As a result of metagenomic (Klatt et al., 2011) and metatranscriptomic (Liu et al., 2011, 2012; Klatt et al., 2013) analyses, an objective and more complete understanding of the major taxa inhabiting the upper 2 mm of the 60–65°C regions of the Mushroom Spring mat, in terms of their contribution to the gene pool and their functional potentials, has emerged (**Table 1**). Cyanobacteria from the genus *Synechococcus* are the predominant primary producers driving metabolism in these communities via oxygenic photosynthesis (Klatt et al., 2011; Liu et al., 2011). *Synechococcus* spp. fix CO_2 and synthesize, and possibly excrete, metabolites that are then consumed by (photo)-heterotrophic members of the community, including several Chloroflexi, especially *Roseiflexus* spp. (**Table 1**), which were formerly thought to be exclusively photoheterotrophs. However, genomics, metagenomics, and metatranscriptomics analyses have revealed that *Roseiflexus* spp. also have the genetic potential to fix CO_2 (Klatt et al., 2007; Van Der Meer et al., 2010). Collectively, cyanobacteria and *Roseiflexus* spp. account for the majority of the biomass of the upper 0–2 mm portion of the mat community (**Table 1**), and thus they should have the greatest influence on the metabolites in this portion of the mat. Two additional Chloroflexi, *Chloroflexus* spp. and a novel, apparently phototrophic, Anaerolineae-like taxon, and two aerobic/microaerophilic, anoxygenic photoheterotrophs, Chloracidobacterium thermophilum (Bryant et al., 2007; Garcia Costas et al., 2012) and "*Candidatus* Thermochlorobacter aerophilum" (Liu et al., 2012), also occur in the upper photic layer of the mat. Non-chlorophyllous, heterotrophic bacteria have been detected in the upper mat community, but they are much less abundant (Liu et al., 2011), and are unlikely to strongly influence mat metabolites. Heterotrophs, together with the photoheterotrophic and photomixotrophic community members, can be considered potential consumers of metabolites produced by cyanobacteria and possibly other mat inhabitants.

Studies performed by Konopka (1992) and Nold and Ward (1996) showed that CO_2-fixing chlorophototrophic community members undergo diel metabolic switching. Recently, metatranscriptomics analyses have provided a comprehensive view of diel transcription patterns in predominant mat taxa (Liu et al., 2011, 2012; Klatt et al., 2013), and have led to new hypotheses about *Synechococcus* spp. and *Roseiflexus* spp. metabolisms. For instance, *Synechococcus* spp. express genes involved in photosynthesis diurnally and have the genetic potential to produce glycogen, which they accumulate during the day (Van Der Meer et al., 2007). Extremely high irradiance during the day leads to O_2 supersaturation combined with CO_2 depletion (as indicated by elevated pH), causing production and possible accumulation of toxic levels of glycolate, a common product of photorespiration (Bateson and Ward, 1988). *Synechococcus* spp. also have the

genetic potential to conduct fermentation with production of lactate, acetate, ethanol and formate (Bhaya et al., 2007). When photosynthesis declines in the evening, O_2 uptake by aerobically respiring community members exceeds O_2 production and the mat becomes anoxic, except within the upper ∼150 μm. Fermentation genes, as well as genes involved in N_2 fixation, are expressed at this time, consistent with measured N_2 fixation driven by fermentative metabolism at night and by light in the early morning (Steunou et al., 2006, 2008).

Diurnal transcription patterns of the genes involved in CO_2 fixation suggested that *Roseiflexus* spp. can conduct photomixotrophic metabolism, in which they combine CO_2 fixation with assimilation of low-molecular weight organic compounds, possibly produced by *Synechococcus* spp. (Klatt et al., 2013). Other transcription patterns suggested that *Roseiflexus* spp. construct and decompose intracellular polymers, including glycogen, polyhydroxyalkanoates (PHAs) and possibly wax esters (genomic and metagenomic analyses show that *Synechococcus* spp. lack the ability to synthesize PHAs (Bhaya et al., 2007; Klatt et al., 2011). Because external reductants such as H_2 and H_2S are not present in the oxic mid-day photic layers of the mat, it was further hypothesized that utilization of these intracellular storage polymers may provide reductants and organic intermediates for photomixotrophic CO_2 incorporation during the day. As suggested by Bauld and Brock (1973), organic compounds produced by CO_2-fixing community members might be cross-fed to (photo)-heterotrophic or mixotrophic mat community members. Little is known about metabolite exchange in the mat, although it has been shown that acetate, butyrate, ethanol, glycolate, lactate, and propionate are photoassimilated into filamentous community members (Anderson et al., 1987; Bateson and Ward, 1988).

Metabolomics has been successfully applied to characterize the metabolic responses of diverse organisms, both qualitatively and quantitatively, under various growth conditions (Koek et al., 2011). These measurements are increasingly used to study microbial communities (Mosier et al., 2013; Xie et al., 2013). In the current study, a combination of untargeted and targeted metabolomics analyses was performed to quantify five groups of metabolites. Volatile organic acids, polar metabolites, wax esters, and PHAs were measured in the mat, while selected dissolved gasses and inorganic ions were quantified in the overflowing water. Measurements of acetate, propionate, and glycolate in the mat, as well as H_2, CO_2, and CH_4 in the water, were performed to test hypotheses regarding the production of these products during different parts of the diel cycle. Similarly, targeted measurements of wax esters and PHAs were performed to characterize these molecules as intracellular carbon and energy storage polymers that should undergo diel cycling if photomixotrophy occurs as hypothesized in *Roseiflexus* spp. (Klatt et al., 2013). Finally, untargeted metabolomics measurements were performed to identify and quantify polar metabolites extracted from the mat samples (intracellular) and interstitial fluids (extracellular) to identify additional metabolites that are changing during the diel cycle and that may be available for possible metabolic exchange among mat community members, respectively. In addition to

TABLE 1 | Major members of Mushroom Spring mat photosynthetic layer and their metabolisms.

Organism type based on metagenome cluster[a]	% of rRNA[b]	Mode of primary carbon assimilation in culture[c]	Hypothesized mode of primary carbon assimilation in mat[c]	Potential carbonaceous substrates consumed based on evidence of[c]		Use by representative isolates[d]	Potential metabolites inferred from genomic/metagenomic analyses		Storage polymers inferred from genomic/metagenomic analyses
				Transporter genes	Pathway genes		From photorespiration	From fermentation	
Synechococcus spp.	27.0	autotrophy	autotrophy	CO$_2$, urea, carbohydrates, amino acids	CO/bicarbonate, urea, carbohydrates	CO$_2$/bicarbonate	glycolate, glyoxylate, oxalate, glycerate, formate, glycine, serine[e]	lactate, formate, acetate, ethanol	glycogen, cyanophycin
Roseiflexus spp.	33.4	heterotrophy	mixotrophy	lactate, fructose, citrate, ornithine, amino acids/peptides, putrescine, spermidine, glucosamine, aminoglycoside, mannitol, trehalose, acetylmuramate, diacetyl chitobiose, mellibiose, dimethoxytoluene	CO$_2$/bicarbonate, glycolate, lactate, citrate, fructose, amino acids/peptides, ornithine, putrescine, spermidine, glucosamine	fructose, glucose, acetate, pyruvate, lactate, propionate, succinate, malate[f]		lactate, acetate, propionate	glycogen, PHAs, wax esters
Chloracidobacterium spp.	10.0	heterotrophy	heterotrophy	oligopeptides, spermidine, putrescine, glutamate, ribose, branched chain amino acids, mannose	mannose, amino acids, glutamate	2−oxo−glutarate, amino acids[h]		lactate, alcohols	glycogen
Candidatus Thermochlorobacter spp.	6.0	heterotrophy	heterotrophy	oligopeptides, sugars, oxalate, formate, maltose, acetate, methionine, bicarbonate, glutamate, hydroxymethyl-pyrimidine	amino acids, carbohydrates, formate, acetate, methionine, glutamate, hydroxymethyl-pyrimidine			lactate, acetate, CO$_2$	glycogen
Chloroflexus spp.	2.5	autotrophy and heterotrophy[g]	mixotrophy	oligopeptides, branched chain amino acids, glutamate, maltose, hydroxymethyl-pyrimidine, ribose, spermidine, putrescine, acetylglucosamine, glycerol−3−phosphate, rhamnose, arabinose	CO$_2$, acetate, glycolate, glyoxylate, lactate, glucose, fructose, sucrose, mannose, lactose, galactose, maltose, maltodextrin, trehalose	CO$_2$		lactate	glycogen, PHAs, wax esters

(Continued)

TABLE 1 | Continued

Organism type based on metagenome cluster[a]	% of rRNA[b]	Mode of primary carbon assimilation in culture[c]	Hypothesized mode of primary carbon assimilation in mat[c]	Potential carbonaceous substrates consumed based on evidence of[c]			Potential metabolites inferred from genomic/metagenomic analyses		Storage polymers inferred from genomic/metagenomic analyses
				Transporter genes	Pathway genes	Use by representative isolates[d]	From photorespiration	From fermentation	
Anaerolineae–like	3.8	unknown	unknown						
Heterotroph 1	ND	ND	heterotrophy		glycolate, acetate				
Heterotroph 2	ND	ND	heterotrophy						

[a]From Klatt et al. (2011).
[b]From Liu et al. (2012).
[c]Gene presence is based on annotations of metagenomic clusters and genomes of representative isolates (Klatt et al., 2011, 2013; Liu et al., 2012). Other nutritional requirements, such as those for vitamins, are not considered here.
[d]Only strains known to be representative of Mushroom Spring populations were considered, including Roseiflexus sp. strains RS1 and RS2 and Chloracidobacterium thermophilum.
[e]Possible photorespiration products (Bauwe et al., 2010).
[f]From (Van Der Meer et al., 2010).
[g]Chloroflexus sp. strain OK-70-fl was reported to grow autotrophically on sulfide and bicarbonate (Madigan and Brock, 1975), and a Chloroflexus sp. isolate from Mushroom Spring can grow on sulfide and carbon dioxide (Thiel et al., unpublished).
[h]Tank and Bryant, 2015.
ND, not determined.

evaluating the above hypothesized metabolisms, these data were collectively used to formulate new hypotheses of community metabolisms and metabolite exchange.

Materials and Methods

Chemicals and Materials

All chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. A mixture of fatty acid methyl esters (FAMEs; C8–C28) dissolved in hexane was prepared for use as a retention index standard. PHA polymers were purchased from Sigma-Aldrich or were provided as a gift by Prof. Alexander Steinbüchel at University of Münster, Germany. Deionized and purified water was used to prepare buffer and standard solutions (Milli-Q System Advantage A10, Merck Millipore, Billerica, MA). All solvents and chemicals were obtained in the highest purity available.

Sample Collection
Mat Samples
For whole-mat (i.e., intracellular and extracellular metabolites combined) analyses of volatile organic acids, polar metabolites, wax esters, and PHAs, core samples were collected in from a microbial mat in the effluent channel of Mushroom Spring in the Lower Geyser Basin (YNP, WY) where the temperature of water in the sampling area varied from 58 to 62°C during the diel cycle. A cork-borer with a 8 mm inner diameter was used to collect the same volume of mat sample, and a razor blade was used to separate the top 5 mm of each mat core such that the analyses were focused on the top green phototrophic layer and the red-orange undermat layers in the zone that contain most of the biological activity (Ward et al., 1987) (**Figure 1**). The samples were transferred to microcentrifuge tubes and immediately frozen in a Dewar containing liquid nitrogen. Mat samples were collected in triplicate at 14 time points between 13:30 h on September 21, 2012 and 11:00 h the following day.

For analyses of extracellular metabolites, mat core samples were collected at 03:00, 09:00, 13:00, and 19:00 h ($n = 6$, each) during the same diel cycle. Once collected, three core samples from each time point were immediately frozen as described above for use as unrinsed controls, while the remaining three samples were transferred to 15 mL Falcon tubes containing 1 mL of spring water that had been filtered through a 0.2-μm filter. Since the 68°C source pool of Mushroom Spring is lined with mat, in order to avoid metabolites that might have diffused from the mat to overflowing water, we used water from the source pool (92°C) of chemically similar Octopus Spring (Papke et al., 2003), which is well upstream of photosynthetic mats (maximum range of 72–74°C). This water did not contain significant levels of any of the organic compounds detected in this study. The re-suspended mat cores were then quickly disrupted onsite by vigorous shaking, and the biomass and rinse water were then immediately separated using a centrifuge ($16,025 \times g$ for 5 min). The supernatant was transferred to a clean microcentrifuge tube and the rinsed biomass and the rinse water samples were immediately frozen with liquid nitrogen. All samples were stored at −80°C until further processing. This process did not

FIGURE 1 | Core sampling using a cork-borer in Mushroom Spring phototrophic mat (A) and longitudinal view of core sample and brief analytical scheme (B).
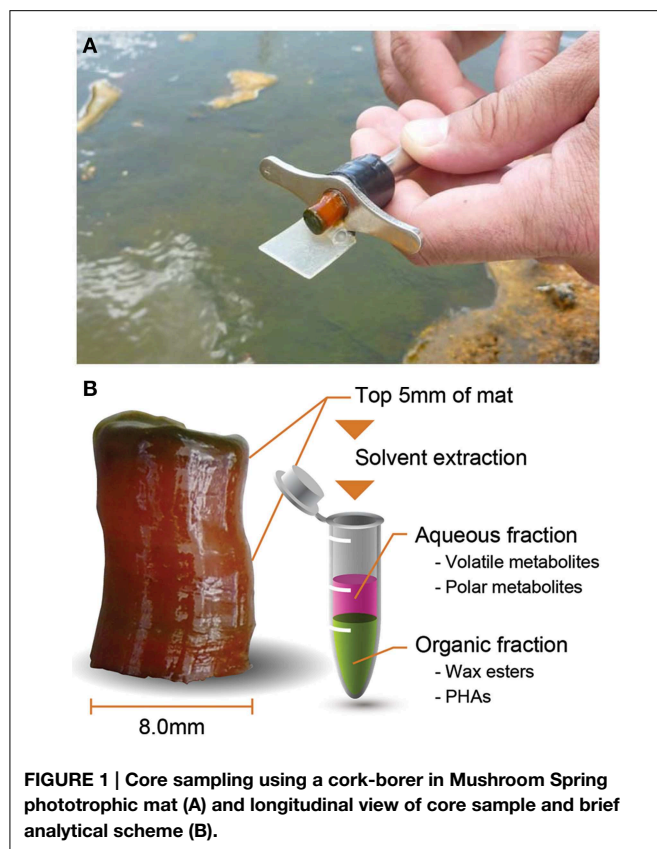
result in release of metabolites identified in analyses of biomass, suggesting that it did not cause leakage of constituents from intact cells.

Water Samples

Duplicate water samples were collected at 03:00, 07:00, 09:00, 11:00, 13:00, 15:00, 17:00, 19:00, and 23:00 h, during the same diel cycle. The temperature at the collection site was approximately 60°C in the main effluent channel. Channel water was filtered through 0.4 μm HTTP Isopore™ polycarbonate membrane filters, collected in 160-mL serum bottles, and then after several exchanges of the serum-bottle volume, sealed with butyl-acetate stoppers (without head-space).

Metabolite Extraction

A single metabolite extraction protocol was used for the analysis of the various classes of metabolites described herein. Frozen mats were thawed at room temperature and 100 μL each of Nanopure™ water and zirconia-silica beads (0.1 mm size; Biospec Products; Bartlesville, OK) were added, respectively, to the samples and vigorously vortexed for 2 min. This bead-beating process was repeated after the samples were maintained at room temperature for 5 min. A mixture of chloroform/methanol (400 μL; 2:1, v/v) spiked with 20 μg of ^{13}C-labeled acetate (Sigma-Aldrich catalog number 282022-250) was added to each disrupted mat sample, and the mixtures were repeatedly vortexed to ensure thorough mixing. The samples were centrifuged at 15,000 × g

for 5 min at 4°C to separate aqueous and organic layers from precipitated proteins.

For analysis of acetate and propionate, aliquots (50 μL) of the aqueous layer from each extract were transferred to glass vials equipped with glass inserts for direct GC-MS analysis without chemical derivatization. Because acetate and propionate are volatile molecules, all samples were immediately analyzed after extraction.

For untargeted analysis of polar metabolites, aliquots (150 μL) of the remaining aqueous layer from each sample extract were transferred to glass vials and completely dried *in vacuo*. The dried extracts were stored at -20°C until chemical derivatization.

For analysis of wax esters, aliquots (200 μL) of the organic layer from each extract were analyzed directly using GC-MS without chemical derivatization.

For analysis of PHAs, the remaining organic layer from each extract was combined with the corresponding protein pellet, and the combined extract and pellet were completely dried *in vacuo*. The samples were hydrolyzed using a modification of the method reported by Lageveen et al. (1988). Briefly, the dried pellets were dissolved in methanol containing 15% H_2SO_4 (v/v) and incubated at 100°C for 15 h. The resulting PHA monomers were extracted with chloroform and analyzed by GC-MS.

Metabolomics Analyses

An Agilent 7890A gas chromatograph coupled with a single quadrupole 5975C mass spectrometer (Agilent Technologies, Inc.) was used for all analyses. Samples were analyzed in duplicate by optimized GC-MS methods, which varied according to the classes of molecular targets as described below.

Acetate and propionate were quantified in a targeted fashion using ^{13}C-labeled acetate as an internal standard. Briefly, mixtures of unlabeled acetate and proprionate at different concentrations were combined with constant amounts of ^{13}C-labeled acetate in order to construct calibration curves. ^{13}C-labeled acetate was then spiked into microbial mat lysates prior to extraction of metabolites, and the measured ratios of unlabeled acetate and propionate to labeled internal standard were used to accurately quantify the target molecules. A polar column (HP-FFAP; 30 m × 0.250 mm × 0.250 μm; Agilent Technologies, Santa Clara) was used. The temperature of the GC inlet was maintained at 200°C, and samples (1 μL) were injected in splitless mode with a helium gas flow rate of 1.0 mL min^{-1}. A temperature gradient from 40 to 200°C over 20 min was used, and data were collected over the mass range 20–300 *m/z*. To reduce any carry over arising from the direct injection of the aqueous layers (a mixture of methanol and water) from the metabolite extraction procedure, pure methanol blanks were analyzed between each sample.

For untargeted analysis of polar metabolites, extracted metabolites in the dried aqueous layers were chemically derivatized to trimethylsilyl esters as previously described (Kim et al., 2013). Metabolite extracts were dried *in vacuo* again to remove any residual moisture. To protect carbonyl groups and reduce the number of tautomeric isomers, methoxyamine (20 μL of a 30 mg mL^{-1} stock in pyridine) was added to each sample, followed by incubation at 37°C with shaking for 90 min. To derivatize hydroxyl and amine groups to trimethylsilylated (TMS)

forms, *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) (80 μL) was added to each vial, followed by incubation at 37°C with shaking for 30 min. The samples were allowed to cool to room temperature and were analyzed on the same day. A HP-5MS column (30 m × 0.25 mm × 0.25 μm; Agilent Technologies) was used for untargeted analyses. Samples (1 μL) were injected in splitless mode, and the helium gas flow rate was determined by the Agilent Retention Time Locking function based on analysis of deuterated myristic acid (Agilent Technologies, Santa Clara, CA). The injection port temperature was held at 250°C throughout the analysis. The GC oven was held at 60°C for 1 min after injection, and the temperature was then increased to 325°C by 10°C/min, followed by a 5 min hold at 325°C. Data were collected over the mass range 50–550 *m/z*. A mixture of FAMEs (C8–C28) was analyzed together with the samples for retention index alignment purposes during subsequent data analysis.

For analysis of wax esters, aliquots of the organic layer from the metabolite extracts were directly injected into the GC-MS. For analysis of PHA monomers, the acid-hydrolyzed samples were analyzed. Wax esters and PHA monomers were chromatographically separated using the same HP-5MS column as described above. Samples (1 μL) were injected in splitless mode. The GC oven was held at 60°C for 5 (wax esters) or 10 (PHA monomers) min after injection, and the temperature was then increased to 325°C by 10°C/min, followed by a 1 (PHA monomers) or 5 (wax esters) min hold at 325°C. The helium gas flow rate was 1.0 mL/min and the injection port temperature was held at 250°C throughout the analysis. Data were collected over the mass range 50–600 (PHA monomers) or 50–700 (wax esters) *m/z*.

All GC-MS raw data will be made available via the Metabo-Lights metabolomics data repository (http://www.ebi.ac.uk/metabolights/) under study identifier MTBLS187.

Metabolomics Data Analysis

The relative amounts of acetate and propionate in the mat samples were quantified by isotope dilution mass spectrometry. Standard curves for acetate and propionate were constructed as described above, and the integrated peak areas of acetate, propionate, and ^{13}C-acetate in mat samples were determined using the corresponding extracted ion chromatograms (EICs; acetate, *m/z* 60; propionate, *m/z* 74; and ^{13}C-acetate, *m/z* 62). The peak areas of endogenous acetate and propionate were divided by that of ^{13}C-acetate to obtain ratios of unlabeled/labeled target molecules.

GC-MS raw data files from untargeted analyses of polar metabolites were processed using MetaboliteDetector (Hiller et al., 2009). Retention indices (RI) of detected metabolites were calculated based on the analysis of the FAME standard mixture, followed by their chromatographic alignment across all analyses after deconvolution. Metabolites were then identified by matching GC-MS features (characterized by measured retention indices and mass spectra) to an augmented version of the Agilent Fiehn Metabolomics Retention Time Locked (RTL) Library (Kind et al., 2009), which contains spectra and validated retention indices for over 700 metabolites. All metabolite identifications were manually validated to reduce deconvolution errors during automated

data-processing and to eliminate false identifications. The NIST 08 GC-MS library was also used to cross-validate the spectral matching scores obtained using the Agilent library. A heat-map analysis was also carried out after *z*-score transformation of the obtained signal intensities and with K-means clustering ($K = 5$, Distance metric: Euclidean) using DanteR (Taverner et al., 2012).

For wax ester analysis, the 10 most abundant species from >30 detected and quantified in the microbial mat samples were selected based on a previous report (Dobson et al., 1988), and their abundances were determined by the EIC method described above. For PHA analysis, the monomers were also quantified using the EIC method described above. A common representative fragment ion (*m/z* 103) was used for quantifying both 3-hydroxybutyrate and 3-hydroxyvalerate.

Dissolved Gas Analysis

Dissolved gasses (CO_2, H_2, and CH_4) were determined using closed head-space GC as described (Inskeep et al., 2005).

Solar Irradiance Analysis

The incident downwelling irradiance was logged throughout the field campaign with a LI-1400 light meter equipped with a LI-192 quantum irradiance sensor (LI-COR, Lincoln, NE).

Results

Triplicate samples were taken from a 60°C region of Mushroom Spring mat at approximately 2-h intervals over a diel cycle (**Figure 1**). The top 5 mm was removed for solvent extraction and separate analyses of polar and volatile aqueous metabolites, PHAs and wax esters.

Polar Metabolites in the Mat

Untargeted metabolomics analyses were performed to identify fluctuations in polar metabolites. This analysis resulted in identification of 58 metabolites that were reproducibly detected in the 42 samples over the diel cycle. The time-course abundance patterns of these 58 metabolites are shown individually in Supplemental Figure S1. K-means clustering was used to categorize these patterns of temporal changes in relative abundances, resulting in five clusters of metabolites (**Table 2**) that each contained metabolites sharing similar patterns of abundance fluctuation over the diel cycle (**Figure 2**).

Metabolites detected in Cluster A accumulated in the predawn and early morning and included 3-hydroxybutyrate and 3-hydroxyvalerate, the monomeric units of PHA. These two compounds showed similar patterns of relative abundance over the diel cycle, as well as to the monomers liberated from acid hydrolysis of intact PHA polymers (see below). Many of the metabolites of Cluster A were lowest in relative abundance in the afternoon and began to increase at 03:00 h, peaking by 09:00 to 11:00 h. Adenine, ornithine (indistinguished from arginine during GC-MS analysis), dihydroxyacetone phosphate, α-hydroxyglutaric acid, sophorose and phosphoinositol showed similar diel profiles.

Cluster B contains metabolites that showed highest abundances in late morning. These included the majority of the

TABLE 2 | List of categorized metabolites showing diel cycling patterns.

Cluster	Pattern[a]	Metabolites
A (8)	Increase in early morning (03:00–11:00)	adenine dihydroxyacetone phosphate 3-hydroxybutyric acid α-hydroxyglutaric acid 3-hydroxyvaleric acid L-ornithine phosphoinositol* sophorose
B (28)	Increase in late morning (07:00–11:00)	L-asparagine L-cysteine fumaric acid D-glucose D-glucose-6-phsophate L-glutamic acid L-glutamine glycine hypoxanthine inosine lumazine L-lysine D-malic acid maltose maltotriose methylcitric acid* nicotinic acid L-phenylalanine phosphate ion L-pyroglutamic acid ribose L-serine succinic acid L-threonine thymine L-tyrosine uracil L-valine
C (9)	Increase in afternoon (11:00–15:30)	carbonate ion citric acid glyceric acid glycolic acid L-homoserine oxalic acid 2-oxo-glutaric acid phosphoenolpyruvic acid pyruvic acid
D (11)	Increase in late afternoon/early evening (14:30–22:00)	adenosine benzoic acid glycerol-3-phosphate L-(+) lactic acid D-(+) melezitose 1-methyl nicotinamide 3-phosphoglyceric acid pyrophosphate D-(+) trehalose urea xylopyranose
E (2)	Increase at night (19:00–22:00)	fructose sucrose

[a]*Clusters are the same as those shown in **Figure 2**. Numbers in parentheses correspond to the numbers of metabolites comprising the cluster.*
Metabolites identified by the NIST spectral library only.

amino acids identified in the mat, precursors for the synthesis of nucleic acids, such as hypoxanthine, inosine, phosphoric acid, ribose, thymine, and uracil, as well as intermediates in glycolysis (e.g., glucose and glucose-6-phosphate) and the citric acid cycle (e.g., fumaric, malic, and succinic acids). The abundance profile of maltose, an α-1,4 disaccharide of glucose, paralleled that of glucose, whereas maltotriose, an α-1,4 trisaccharide of glucose, initially increased in abundance, gradually declined and then remained low with minor oscillations throughout the afternoon. A number of metabolites (e.g., asparagine, glycine, malic acid, phenylalanine, succinic acid, threonine, tyrosine, and valine) showed maximal abundances at 11:00 h, followed by an abrupt decrease near mid-day, which was then followed by a secondary maximum around 14:00–15:00 h.

Metabolites assigned to Cluster C showed highest abundance in the early afternoon, a time that correlates to peak photosynthetic activity over the diel cycle (see Discussion). Organic acids such as citric, glyceric, glycolic, oxalic, 2-oxo-glutaric acid (α-ketoglutaric acid), and pyruvic acids were detected in highest abundance during the period of 12:00 to 16:00 h. In contrast to the proteinogenic amino acid serine, the abundance of homoserine was highest from 11:00 to 14:00 h, during which there was an abrupt decrease at mid-day.

Cluster D metabolites accumulated in the late afternoon. Among these, the amounts of lactate and urea dramatically increased from 12:30 to 15:30 h, then gradually decreased until mid-night. In contrast, benzoic acid, glycerol-3-phosphate and trehalose, an α,α-1,1 disaccharide of glucose, showed peak abundance in the early evening (17:00 h).

Only two metabolites, fructose and sucrose, were assigned to cluster E; these accumulated around 19:00 to 22:00 h, decreased at 23:00 h, and then exhibited a relatively low but constant abundance from midnight to noon.

Volatile Fatty Acids in the Mat

To quantify acetate and propionate in the mat accurately, ^{13}C-labeled acetate was spiked into samples as an internal standard before metabolite extraction, and the ratios of unlabeled acetate and propionate peak areas to ^{13}C-acetate peak area were compared to a calibration curve. Using this quantitative approach, the levels of acetate and propionate were observed to be highest at midnight, followed by a gradual decrease to 17:00 h (**Figure 3**). Overall, the abundances of acetate and propionate were similar to each other over the diel cycle.

Carbon Storage Polymers in the Mat

PHA was measured over the diel cycle in the form of the major components 3-hydroxybutyric acid (3-HB) and 3-hydroxyvaleric acid (3-HV) (**Figure 4A**). 3-HV was three times more abundant than 3-HB, with both fluctuating over the diel cycle, although the general trend was for accumulation from 19:00 to 10:00 h followed by a decrease between 10:00 and 19:00 h.

The mat contained a mixture of C_{30}-C_{35} *n,n* and *i,n* wax esters. In total, 42 species were identified (Supplemental Table S1), although we present data for the 10 most abundant species here (representative data shown in **Figures 4B,C**; all data shown in Supplemental Figure S2). The abundances of these wax esters
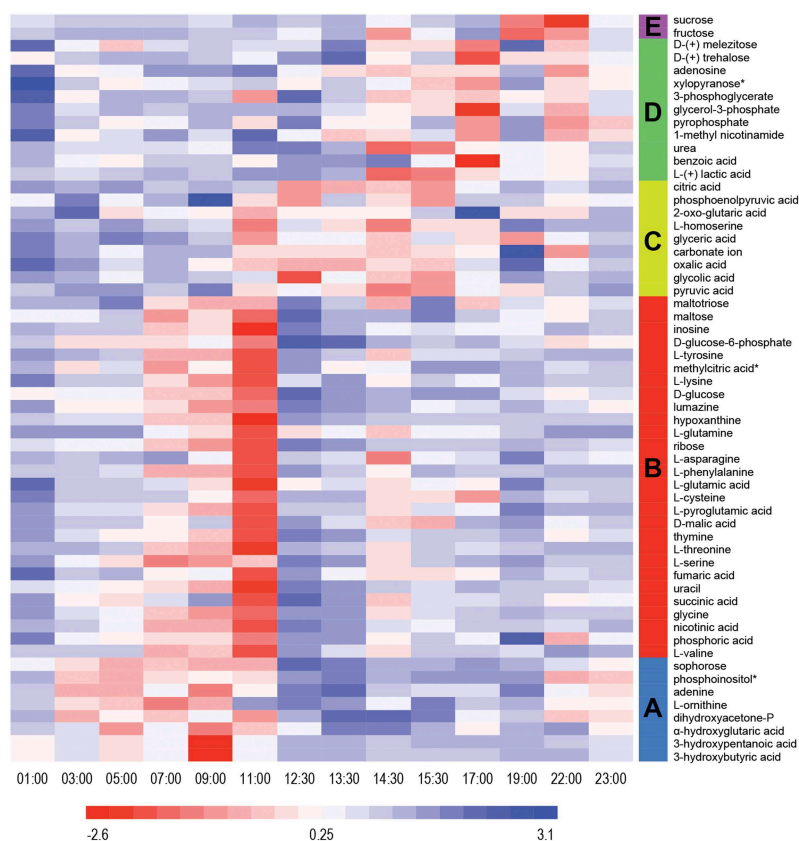
FIGURE 2 | Heat map view of z-scored polar metabolite abundances over a diel cycle. A total of 58 metabolites were reproducibly detected (*n* = 3) and confidently identified in the mat samples. The scale bar indicates the z-score transformed average intensity values of metabolites. K-means clustering was performed to categorize the metabolites based on their diel abundance patterns. Cluster A—increase in early morning; Cluster B—increase in late morning; Cluster C—increase in afternoon; Cluster D—increase in late afternoon/early evening; Cluster E—increase at night.

changed throughout the diel cycle in a complex pattern. In general, the wax ester abundances showed decreases from mid-night to mid-day, except for increases in morning and afternoon, followed by an increase again in the evening. Interestingly, *i,n* forms of C_{31}, C_{32}, and C_{35} wax esters increased before *n,n* forms.

Metabolite Partitioning in the Mat

To evaluate the potential for metabolite exchange among members of the community, we analyzed additional mat core samples that were collected during four time points over the diel cycle and measured metabolites that were excreted or were otherwise extracellular. For this experiment, warm, filtered hot spring water was used to rinse the mat samples on site to avoid release of metabolites due to osmotic shock. Glycolate and lactate were the only metabolites confidently identified in the rinse waters within the detection limits of our instrumentation (data not shown). We compared the levels of these metabolites between the rinsed and control mats (**Figures 5A,B**) at 4 time points during the diel cycle. The highest level of glycolate in control and rinsed mat samples occurred at 13:00 h. Otherwise, the level of glycolate was relatively the same at 03:00, 09:00, and 19:00 h. In contrast, the abundance of lactate in the rinsed

mat samples was equal across the four time points. The levels of lactate in the control mat were much higher than in the rinsed mat and increased from 03:00 to 19:00 h. **Figure 5C** shows data for glycolate and lactate over the full diel in unrinsed mat samples. The diel trends for glycolate in **Figure 5A** (control mat) and **Figure 5C** (unrinsed samples from the full diel sampling) clearly track each other. Similarly, the data for lactate in control mat from the rinsing experiment (**Figure 5B**) shows a rise in lactate abundance beginning at 13:00 h, which matches the time of the rise in lactate in the unrinsed samples from the full diel sample collection (**Figure 5C**). However, while the lactate abundance continues to rise to 19:00 h in the control mat from the rinsing experiment, it has begun to decline by 15:30 h in the unrinsed samples from the full diel experiment.

Gases in the Overflowing Water

The amounts of three gaseous molecules in the water overflowing the 60°C mat—CO_2, H_2, and CH_4—were also measured over the diel cycle (**Figure 6**). While the levels of hydrogen and carbon dioxide were lower during the day, methane abundance fluctuated, with maxima at 07:00, 12:30, and 23:00 h.

FIGURE 3 | Volatile metabolites over the diel cycle. Solar irradiance (solid gray) and acetate and propionate concentrations in unrinsed mat samples are shown. Metabolite values plotted are mean ± standard error (*n* = 3).
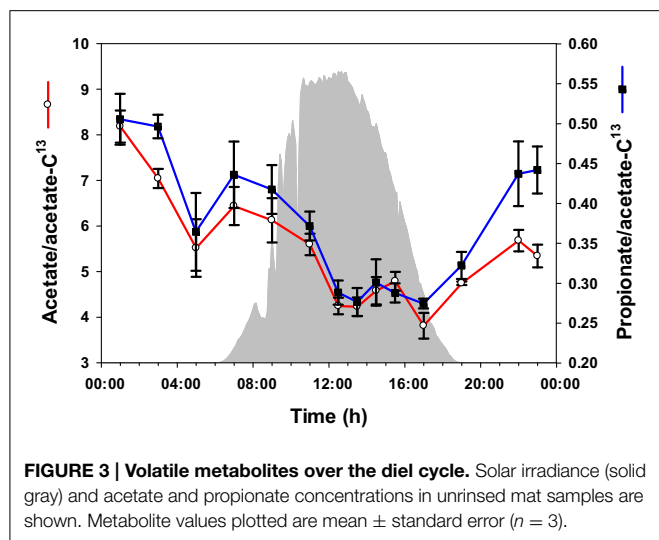
Discussion

The application of systems biology approaches is expanding from lab-cultured samples to complex environmental communities. In this way, integrated studies are becoming more common for understanding biological sy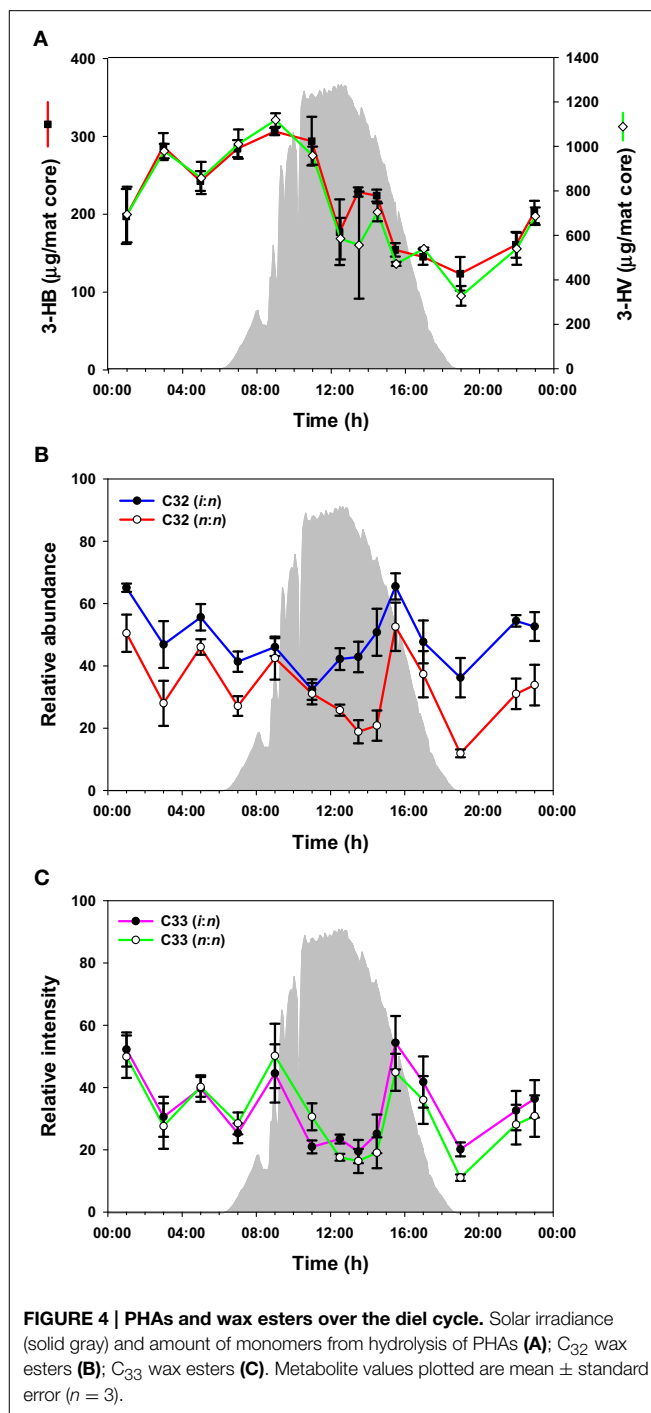stems through the combination of data from metagenomics, metatranscriptomics, metaproteomics, and metametabolomics analyses. Interpreting data from metabolomics analyses of a complex microbial community is challenging because many taxa may contribute to metabolite pools and because they may do so at different times during a diel cycle. Furthermore, metabolite concentrations represent pools that are influenced by production and consumption, as well as by diffusion, and all three factors are closely coupled in aquatic microbial mats. Thus, metabolite fluctuations with time likely represent periods of net production/accumulation or consumption/diffusion. Nevertheless, the data obtained in this study supported existing hypothesized metabolisms of major taxa in the mat and led to new hypotheses based on novel observations, as discussed below.

Integration of Metabolomics and Gene Expression Data: Support of Existing Hypotheses on *Synechococcus* spp. and *Roseiflexus* spp. Metabolisms within the Mat Community

In this section, we interpret certain metabolomics results in the context of hypotheses generated from previous diel metatranscriptomics studies (Liu et al., 2011, 2012; Klatt et al., 2013). Although the metatranscriptomics results are from a different year (September 2009), the long-term stability of the mat community, its composition and structure (Ramsing et al., 2000; Ferris et al., 2003; Ward et al., 2006; Becraft et al., 2011; Melendrez et al., 2011), processes conducted during diel cycles by phototrophic community members based on O_2 concentration profiles, and expression of *Synechococcus* photosynthesis and N_2 fixation genes (Ramsing et al., 2000; Ward et al., 2006; Steunou et al., 2008; Jensen et al., 2011; Liu et al., 2011, 2012) as measured between 1996 and the present, make comparisons of

FIGURE 4 | PHAs and wax esters over the diel cycle. Solar irradiance (solid gray) and amount of monomers from hydrolysis of PHAs **(A)**; C_{32} wax esters **(B)**; C_{33} wax esters **(C)**. Metabolite values plotted are mean ± standard error (*n* = 3).

data collected at comparable temperature sites and times of the year valid. Indeed, comparison of solar irradiance and glycolate levels over a diel cycle in mat samples collected in 2011 showed very similar abundance profiles as the data presented here (Supplemental Figure S3).

Synechococcus spp.

Based on diel changes in glycogen (Van Der Meer et al., 2007) and metatranscriptomics analyses (Liu et al., 2012), we

FIGURE 5 | Measured levels of glycolate (A) and lactate (B) from the rinsed and control (un-rinsed) mats. The difference between the two conditions is regarded as a portion biologically available by excretion and diffusion in the mats, which can be taken up by other heterotrophic bacteria in the communities. Glycolate and lactate profiles in the unrinsed mat over the full diel cycle (C) are shown as a reference. The glycolate and lactate abundances were z-score transformed (i.e., normalized), and the values plotted are mean ± standard error (n = 3). Solar irradiance is shown in solid gray.

FIGURE 6 | Gaseous metabolites over the diel cycle. Solar irradiance (solid gray) and CO_2, H_2, and CH_4 levels in water overflowing the mat are shown. Metabolite values plotted are mean ± standard error (n = 3).

hypothesized that *Synechococcus* spp. shift from daytime photosynthesis and the production of glycogen to nighttime glycogen fermentation (Van Der Meer et al., 2007). Consistent with this hypothesis, fermentation products that mat *Synechococcus* populations have the genetic potential to produce (e.g., acetate and lactate) accumulated during the afternoon and night (**Figures 3, 5C**).
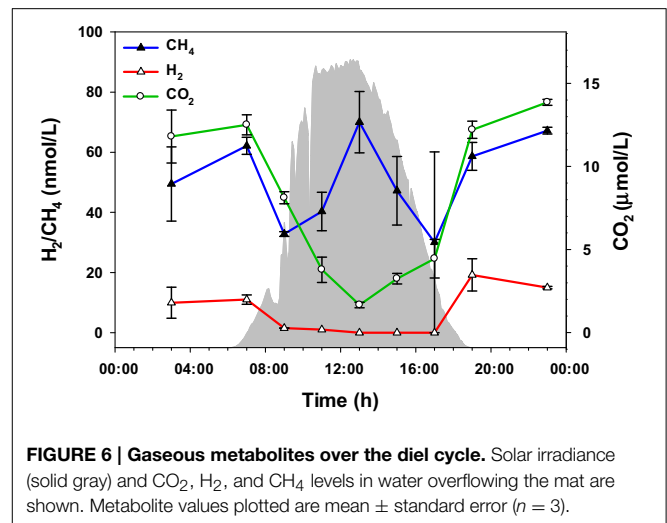
Mid-day extremes of light and O_2 concentration, as well as CO_2 depletion (suggested by a rise in pH, which shifts the carbonate equilibrium) have been shown to lead to photorespiratory production of glycolate (Bateson and Ward, 1988). Thus, we hypothesized that *Synechococcus* spp. in the mat experience photorespiration during periods of high light irradiance. Supporting this hypothesis, CO_2 in the water flowing over the mat decreased during the day (**Figure 6**), and glycolate accumulated between ∼12:00 and ∼16:00 h (**Figure 5C**). Production of glycolate at peak solar irradiance correlated with the expression of *Synechococcus* spp. genes encoding photosynthesis machinery (Liu et al., 2012).

Additionally, nighttime and early morning N_2 fixation by *Synechococcus* has been demonstrated (Steunou et al., 2006, 2008), and because mat *Synechococcus* lack an uptake hydrogenase, we hypothesized that H_2 accumulation should temporally follow N_2 fixation. Diel patterns of H_2 concentration in the water above the mat (**Figure 6**) are consistent with this prediction.

Roseiflexus spp.
Noting the diel cycling of transcript abundances encoding enzymes associated with the 3-hydroxypropionate pathway and the production and consumption of polymers known to be produced by *Roseiflexus* spp., Klatt et al. (2013) hypothesized that *Roseiflexus* spp. shift from a photomixotrophic metabolism leading to glycogen synthesis during the day to nighttime fermentation of glycogen, coupled with nighttime synthesis of PHA and/or wax esters, whose breakdown during the day could in turn provide the necessary metabolites for photomixotrophy. Consistent with this hypothesis, diel glycogen cycling was previously demonstrated by Van Der Meer et al. (2007). Also consistent with the hypothesis, levels of CO_2 in the water overflowing the mat and of intracellular fermentation products known to be used by *Roseiflexus* [e.g., acetate, propionate, and lactate; based

on genomic (Van Der Meer et al., 2010; Bryant et al., 2012) and metagenomic (Klatt et al., 2011) analyses and on laboratory growth experiments (Hanada et al., 2002)] are lower during the day.

In addition, PHAs, measured by their constituent monomers (e.g., 3-HB and 3-HV) after acid hydrolysis of the polymers, were relatively higher at night and in the early morning, followed by a decrease during the day (**Figure 4A**). The accumulation of 3-HB and 3-HV as free monomers in the morning (**Figure 2**, Supplemental Figure S1), together with methyl-citrate, an intermediate in the oxidation of propionate (which could be derived from 3-HV), provides evidence that PHAs are being degraded in the early morning. These observations are consistent with previous metatranscriptomics data on expression of *Roseiflexus* spp. PHA biosynthesis genes, and our previous hypothesis that these molecules might be used for mixotrophic metabolism by filamentous anoxygenic phototrophic bacteria (Klatt et al., 2013).

Wax esters generally cycled in a manner consistent with the expression patterns of *Roseiflexus* genes associated with their production and degradation, supporting their hypothesized involvement in photomixotrophy. However, these compounds fluctuated in a complex manner, possibly reflecting differences due to the timing of metabolisms of different *Roseiflexus* species (see below).

Novel Observations Leading to New Hypotheses

In this section, we highlight novel observations of metabolism in the Mushroom Spring microbial mat community with respect to metabolites identified or measured for the first time, as well as to the time of day at which certain metabolites showed peaks in accumulation. These observations were then used as the basis upon which new hypotheses have been formulated.

Detection and Accumulation of Previously Unreported Metabolites

The accumulation of CH_4 in the mat at mid-day was unexpected (**Figure 6**), since methanogenesis is an anaerobic process that should only occur in the anoxic nighttime mat (Ward, 1978; Sandbeck and Ward, 1981). However, genomic and metagenomic analyses indicate that *Synechococcus* spp. have the potential to metabolize phosphonate (Gomez-Garcia et al., 2011), which can also lead to methane production. We therefore hypothesize that the mid-day peak in methane concentration is a result of *Synechococcus* spp. metabolism of phosphonates.

Metabolites in cluster B accumulated specifically in the morning and in general reached their highest levels at 11:00 h. The metabolites present in this cluster (most amino acids, hypoxanthine, inosine, phosphoric acid, ribose, thymine, and uracil) imply that amino and nucleic acid biosynthesis occur maximally during the early morning period. Interestingly, all of these nitrogen-rich compounds reached peak levels shortly after the maximal period of N_2 fixation by *Synechococcus* spp., which occurred between 06:00 and 10:00 h in the morning (Steunou et al., 2008). This period also corresponded to the time when total mRNA levels increased sharply in members of the major phototrophic taxa that occur in the mats Liu et al., 2011, 2012; Klatt

et al., 2013). While not unexpected, these collective observations lead to the hypothesis that the morning hours represent a time when RNA, DNA, and protein biosynthesis rates are maximal for major taxa in the mat.

At midday (11:00 to 12:00 h) there is an abrupt decline in all metabolites of cluster B, when metabolites of cluster C, including glycolate, oxalate, carbonate, citrate, and phosphoenolpyruvate, accumulated (**Figure 2**). The accumulation of glycolate (as discussed above), glycerate, and oxalate is likely due to photorespiration by *Synechococcus* spp. (Bateson and Ward, 1988; Bauwe et al., 2010). Interestingly, the abundance of carbonate ion also increased at this time, consistent with extreme CO_2 consumption and elevated pH during peak periods of photosynthesis shifting the equilibrium of dissolved inorganic carbon (Revsbech and Ward, 1984; Jensen et al., 2011). Also of interest is the observation that peak production of glycolate coincides with the abrupt decrease in levels of certain metabolites (asparagine, glycine, malic acid, phenylalanine, succinic acid, threonine, tyrosine, and valine) in cluster B, suggesting a decrease in activity in these metabolic pathways possibly due to photoinhibition. At the same time as the abrupt decrease in abundances of cluster B metabolites and just after the peak in glycolate abundance (∼12:00 h), the levels of lactate in the mat begin to increase, with maximal abundance at ∼15:00 h and correlating with a second peak in glycolate abundance (**Figure 5C**). We hypothesize that *Synechococcus* spp. may be a source of the peak in lactate abundance at this time via fermentation either as an alternative metabolism for photoinhibited cells closest to the mat surface, or because cells deeper in the mat experience a shorter period of peak solar irradiance (Becraft et al., this issue; Olsen et al., this issue), or both.

Metabolite Exchange

Metabolic interactions among community members are key features stabilizing the composition and function of microbial communities. In a chlorophototrophic microbial community, organic compounds produced and excreted by CO_2-fixing taxa could be used as nutrients by (photo)-heterotrophic or mixotrophic mat community members. Indeed, diurnal transcription patterns of the genes involved in CO_2 fixation have suggested that *Roseiflexus* spp. in the Mushroom Spring mat community can conduct photomixotrophic metabolism, presumably using organic compounds produced and excreted by other community members. In this section, we discuss the potential for metabolic exchange between *Synechococcus* and *Roseiflexus* spp.

Two metabolites, the photorespiration product glycolate and the fermentation product lactate, were identified in the extracellular fractions of the rinsing experiment and were therefore available as nutrients for members of the mat community. Glycolate was most abundant in the mat during the early afternoon (**Figures 2**, **5C**, and Supplemental Figure S1). At 13:00 h, the amount of glycolate associated with mat biomass was much lower (25–30%) in the rinsed compared to the unrinsed control samples, suggesting that glycolate is excreted into the intracellular milieu (**Figure 5A**). At other time points examined, the amounts of glycolate were similar in rinsed or unrinsed samples, suggesting a balanced consumption and production or

that photorespiration is less active at lower irradiance levels. Although, Klatt et al. (2013) did not observe significant changes in transcription patterns in *Roseiflexus* spp. during the same time period as the peak in mat glycolate abundance, these organisms are still the most likely consumers of glycolate because glyoxylate derived from glycolate by oxidation can readily be assimilated by the 3-hydroxypropionate bi-cycle (Klatt et al., 2007). In contrast, a very sharp and large increase (∼60-fold above the minimum) in transcript abundance for lactate permease in *Roseiflexus* sp. at approximately 18:00 h has been observed (Bryant et al., unpublished data), just after the afternoon increase in lactate abundance in unrinsed vs. rinsed mat samples in our experiment (**Figure 5B**). This observation suggests that *Roseiflexus* sp. might utilize a significant proportion of the lactate produced. Indeed, lactate levels declined in the early evening hours after the spike in lactate permease transcripts occurred. As with glycolate, the lower levels of lactate during the night may indicate an efficient balance between production and consumption. Based on these observations, we hypothesize that glycolate, and possibly lactate (as discussed above), are mostly produced and excreted by the cyanobacteria (i.e., *Synechococcus* spp.) during the early afternoon and are available to other mat inhabitants, particularly *Roseiflexus* spp., as a carbon and energy source.

It is interesting that only glycolate and lactate were identified in the rinse water. We have considered several possible explanations for this observation. It is possible that other extracellular metabolites (e.g., volatile fatty acids, ethanol) may have been lost during the *in vacuo* drying of the rinse water samples, as previous analyses have shown that these compounds accumulate in the aqueous fraction during dark, anaerobic incubation of mat samples (Anderson et al., 1987). Alternatively, our sampling of extracellular metabolites, which occurred at 03:00, 09:00, 13:00, and 19:00 h, may not have occurred during the peak times of metabolite excretion. A third possibility is that certain metabolites are rapidly scavenged from the extracellular milieu as soon

as they are excreted. The last possibility is that there were no other metabolites that were excreted.

Depth- or Temporally-Resolved Metabolisms

As mentioned above, a complex pattern of wax ester abundances was observed, with peak abundances in the predawn, morning, and afternoon periods, and differential timing of *i,n*- and *n,n*-forms of the same wax esters. Such complexity might arise because of contributions from multiple taxa capable of wax ester synthesis with different diel timing. As shown in **Table 3**, although *Roseiflexus* wax esters are a better match to wax esters found in the mat, *Chloroflexus* also makes *n,n* forms of C_{31}, C_{32}, and C_{35} wax esters, and the different abundances of these forms might relate to differential timing of wax ester synthesis in members of these two genera. Such could also be the case for different *Roseiflexus* species. Zeng et al. (1992) showed that the ratio of *i,n*- to *n,n*-forms of C_{31-35} wax esters increased nearly 5-fold in mat layers 4–5 mm below the surface of the highly similar Octopus Spring mat, raising the question of whether different species of *Roseiflexus*, with different vertical distributions, experience different light regimes and have different timing of wax ester synthesis and degradation. Taxon-related and/or depth-related differences in metabolisms may be generally important, because similar small-scale fluctuations were observed in PHA, glycolate, and fermentation products. Furthermore, a number of metabolites (e.g., asparagine, glycine, malic acid, phenylalanine, succinic acid, threonine, tyrosine, and valine) showed maxima in abundances at 11:00 h, followed by an abrupt decrease near mid-day, which was then followed by a secondary maximum around 14:00–15:00 h. Metabolomics analyses were conducted on the top 5 mm region of the mat, whereas the transcription results of Klatt et al. (2013) were from the top 2 mm region. Different taxa (and/or different species within these taxa) inhabit different vertical regions of the mat (Ramsing et al., 2000; Becraft et al., 2011), and we hypothesize that they exhibit maximal metabolic rates for specific processes at different times during the diel cycle.

TABLE 3 | Wax ester content of mat and *Chloroflexi* and timing of abundance changes in the mat.

Wax ester	*Roseiflexus*[a]	*Chloroflexus*[a]	mat[a]	Evening rise start	Night peak	Morning peak	Afternoon rise start	Afternoon peak
C31 *n,n*	+		++	19:00	01:00	05:00	14:30	15:30
C31 *i,n*	+		+	17:00	22:00	03:00	11:30	14:30
C32 *n,n*	++	+	+++	19:00	01:00	05:00, 09:00	14:30	15:30
C32 *i,n*	+++		++	19:00	01:00	05:00, 09:00	12:30	15:30
C33 *n,n*	+	+	+++	19:00	01:00	05:00, 09:00	14:30	15:30
C33 *i,n*	+		++	19:00	01:00	05:00, 09:00	14:30	15:30
C34 *n,n*	+++	++	++	19:00	01:00	05:00, 09:00	13:30	15:30
C34 *i,n*	+++		+	19:00	01:00	05:00, 09:00	13:30	15:30
C35 *n,n*	+	++	+	19:00	01:00	05:00, 09:00	15:00	17:00
C35 *i,n*	+		+	19:00	01:00	05:00, 09:00	13:00	17:00
C36 *n,n*		++						
C37 *n,n*		+						

[a]*Relative abundances of wax esters in Roseiflexus spp., Chloroflexus spp., and the Mushroom Spring microbial mat are indicated by "+" for low abundance, "++" for moderate abundance, and "+ + +" for high abundance. Data for Roseiflexus spp. and Chloroflexus spp. are from Van Der Meer et al. (2010).*

Indeed, evidence that *Synechococcus* species with different depth distributions (Becraft et al., this issue), light adaptations (Nowack et al., this issue) and gene expression timing (Olsen et al., this issue), strongly supports this hypothesis.

Acknowledgments

Supplementary Material

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fmicb.2015.00209/abstract

References

Anderson, K. L., Tayne, T. A., and Ward, D. M. (1987). Formation and fate of fermentation products in hot spring cyanobacterial mats. *Appl. Environ. Microbiol.* 53, 2343–2352.

Bateson, M. M., and Ward, D. M. (1988). Photoexcretion and fate of glycolate in a hot spring cyanobacterial mat. *Appl. Environ. Microbiol.* 54, 1738–1743.

Bauld, J., and Brock, T. (1973). Ecological studies of *Chloroflexis*, a gliding photosynthetic bacterium. *Arch. Mikrobiol.* 92, 267–284. doi: 10.1007/BF00409281

Bauwe, H., Hagemann, M., and Fernie, A. R. (2010). Photorespiration: players, partners and origin. *Trends Plant Sci.* 15, 330–336. doi: 10.1016/j.tplants.2010.03.006

Becraft, E. D., Cohan, F. M., Kühl, M., Jensen, S. I., and Ward, D. M. (2011). Fine-Scale distribution patterns of *Synechococcus* ecological diversity in microbial mats of Mushroom Spring, Yellowstone National Park. *Appl. Environ. Microbiol.* 77, 7689–7697. doi: 10.1128/AEM.05927-11

Bhaya, D., Grossman, A. R., Steunou, A.-S., Khuri, N., Cohan, F. M., Hamamura, N., et al. (2007). Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. *ISME J.* 1, 703–713. doi: 10.1038/ismej.2007.46

Brock, T. D. (1972). "One hundred years of algal research in Yellowstone National Park," in *Proceedings of the International Symposium on Taxonomy and Biology of Blue-Green Algae*, ed T. V. Desikachary (Madras), 393–405.

Brock, T. D. (1978). *Thermophilic Microorganisms and Life at High Temperatures*. New York, NY: Springer-Verlag. doi: 10.1007/978-1-4612-6284-8

Brock, T. D. (1998). Early days in Yellowstone microbiology. *ASM News* 64, 137–140.

Bryant, D. A., Costas, A. M. G., Maresca, J. A., Chew, A. G. M., Klatt, C. G., Bateson, M. M., et al. (2007). *Candidatus* Chloracidobacterium thermophilum: an aerobic phototrophic acidobacterium. *Science* 317, 523–526. doi: 10.1126/science.1143236

Bryant, D. A., Liu, Z., Li, T., Zhao, F., Garcia Costas, A. M., Klatt, C. G., et al. (2012). "Comparative and functional genomics of anoxygenic green bacteria from the taxa *Chlorobi*, *Chloroflexi*, and *Acidobacteria*," in *Advances in Photosynthesis and Respiration: Functional Genomics and Evolution of Photosynthetic Systems*, eds R. L. Burnap and W. Vermaas (Dordrecht: Springer), 47–102.

Dobson, G., Ward, D. M., Robinson, N., and Eglinton, G. (1988). Biogeochemistry of hot spring environments: extractable lipids of a cyanobacterial mat. *Chem. Geol.* 68, 155–179. doi: 10.1016/0009-2541(88)90093-9

Ferris, M. J., Kühl, M., Wieland, A., and Ward, D. M. (2003). Cyanobacterial ecotypes in different optical microenvironments of a 68°C hot spring mat community revealed by 16S-23S rRNA internal transcribed spacer region variation.

Appl. Environ. Microbiol. 69, 2893–2898. doi: 10.1128/AEM.69.5.2893-2898.2003

Garcia Costas, A. M., Tsukatani, Y., Rijpstra, W. I. C., Schouten, S., Welander, P. V., Summons, R. E., et al. (2012). Identification of the bacteriochlorophylls, carotenoids, quinones, lipids, and hopanoids of "*Candidatus* Chloracidobacterium thermophilum." *J. Bacteriol.* 194, 1158–1168. doi: 10.1128/JB.06421-11

Gomez-Garcia, M. R., Davison, M., Blain-Hartnung, M., Grossman, A. R., and Bhaya, D. (2011). Alternative pathways for phosphonate metabolism in thermophilic cyanobacteria from microbial mats. *ISME J.* 5, 141–149. doi: 10.1038/ismej.2010.96

Hanada, S., Takaichi, S., Matsuura, K., and Nakamura, K. (2002). *Roseiflexus castenholzii* gen. nov., sp. nov., a thermophilic, filamentous, photosynthetic bacterium that lacks chlorosomes. *Int. J. Syst. Evol. Microbiol.* 52, 187–193.

Hiller, K., Hangebrauk, J., Jäger, C., Spura, J., Schreiber, K., and Schomburg, D. (2009). MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis. *Anal. Chem.* 81, 3429–3439. doi: 10.1021/ac802689c

Inskeep, W. P., Ackerman, G. G., Taylor, W. P., Kozubal, M., Korf, S., and Macur, R. E. (2005). On the energetics of chemolithotrophy in nonequilibrium systems: case studies of geothermal springs in Yellowstone National Park. *Geobiology* 3, 297–317. doi: 10.1111/j.1472-4669.2006.00059.x

Jensen, S. I., Steunou, A.-S., Bhaya, D., Kuhl, M., and Grossman, A. R. (2011). *In situ* dynamics of O_2, pH and cyanobacterial transcripts associated with CCM, photosynthesis and detoxification of ROS. *ISME J.* 5, 317–328. doi: 10.1038/ismej.2010.131

Kim, Y.-M., Schmidt, B. J., Kidwai, A. S., Jones, M. B., Kaiser, B. L. D., Brewer, H. M., et al. (2013). *Salmonella* modulates metabolism during growth under conditions that induce expression of virulence genes. *Mol. Biosyst.* 9, 1522–1534. doi: 10.1039/c3mb25598k

Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S., et al. (2009). FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal. Chem.* 81, 10038–10048. doi: 10.1021/ac9019522

Klatt, C. G., Bryant, D. A., and Ward, D. M. (2007). Comparative genomics provides evidence for the 3-hydroxypropionate autotrophic pathway in filamentous anoxygenic phototrophic bacteria and in hot spring microbial mats. *Environ. Microbiol.* 9, 2067–2078. doi: 10.1111/j.1462-2920.2007.01323.x

Klatt, C. G., Liu, Z., Ludwig, M., Kuhl, M., Jensen, S. I., Bryant, D. A., et al. (2013). Temporal metatranscriptomic patterning in phototrophic *Chloroflexi* inhabiting a microbial mat in a geothermal spring. *ISME J.* 7, 1775–1789. doi: 10.1038/ismej.2013.52

Klatt, C. G., Wood, J. M., Rusch, D. B., Bateson, M. M., Hamamura, N., Heidelberg, J. F., et al. (2011). Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. *ISME J.* 5, 1262–1278. doi: 10.1038/ismej.2011.73

Koek, M., Jellema, R., Greef, J., Tas, A., and Hankemeier, T. (2011). Quantitative metabolomics based on gas chromatography mass spectrometry: status and perspectives. *Metabolomics* 7, 307–328. doi: 10.1007/s11306-010-0254-3

Konopka, A. (1992). Accumulation and utilization of polysaccharide by hot-spring phototrophs during a light-dark transition. *FEMS Microbiol. Lett.* 102, 27–32. doi: 10.1111/j.1574-6968.1992.tb05792.x

Lageveen, R. G., Huisman, G. W., Preusting, H., Ketelaar, P., Eggink, G., and Witholt, B. (1988). Formation of polyesters by *Pseudomonas oleovorans*: effect of substrates on formation and composition of poly-(R)-3-hydroxyalkanoates and poly-(R)-3-hydroxyalkenoates. *Appl. Environ. Microbiol.* 54, 2924–2932.

Liu, Z., Klatt, C. G., Ludwig, M., Rusch, D. B., Jensen, S. I., Kuhl, M., et al. (2012). '*Candidatus* Thermochlorobacter aerophilum:' an aerobic chlorophotoheterotrophic member of the phylum Chlorobi defined by metagenomics and metatranscriptomics. *ISME J.* 6, 1869–1882. doi: 10.1038/ismej.2012.24

Liu, Z., Klatt, C. G., Wood, J. M., Rusch, D. B., Ludwig, M., Wittekindt, N., et al. (2011). Metatranscriptomic analyses of chlorophototrophs of a hot-spring microbial mat. *ISME J.* 5, 1279–1290. doi: 10.1038/ismej.2011.37

Madigan, M. T., and Brock, T. D. (1975). Photosynthetic sulfide oxidation by *Chloroflexus aurantiacus*, a filamentous, photosynthetic, gliding bacterium. *J. Bacteriol.* 122, 782–784.

Melendrez, M. C., Lange, R. K., Cohan, F. M., and Ward, D. M. (2011). Influence of molecular resolution on sequence-based discovery of ecological diversity among *Synechococcus* populations in an alkaline siliceous hot spring microbial mat. *Appl. Environ. Microbiol.* 77, 1359–1367. doi: 10.1128/AEM.02032-10

Mosier, A. C., Justice, N. B., Bowen, B. P., Baran, R., Thomas, B. C., Northen, T. R., et al. (2013). Metabolites associated with adaptation of microorganisms to an acidophilic, metal-rich environment identified by stable-isotope-enabled metabolomics. *MBio* 4, e00484–e00412. doi: 10.1128/mBio.00484-12

Nold, S. C., and Ward, D. M. (1996). Photosynthate partitioning and fermentation in hot spring microbial mat communities. *Appl. Environ. Microbiol.* 62, 4598–4607.

Papke, R. T., Ramsing, N. B., Bateson, M. M., and Ward, D. M. (2003). Geographical isolation in hot spring cyanobacteria. *Environ. Microbiol.* 5, 650–659. doi: 10.1046/j.1462-2920.2003.00460.x

Ramsing, N. B., Ferris, M. J., and Ward, D. M. (2000). Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter-thick photic zone of a hot spring cyanobacterial mat. *Appl. Environ. Microbiol.* 66, 1038–1049. doi: 10.1128/AEM.66.3.1038-1049.2000

Revsbech, N. P., and Ward, D. M. (1984). Microelectrode studies of interstitial water chemistry and photosynthetic activity in a hot spring microbial mat. *Appl. Environ. Microbiol.* 48, 270–275.

Sandbeck, K. A., and Ward, D. M. (1981). Fate of immediate methane precursors in low-sulfate, hot-spring algal-bacterial mats. *Appl. Environ. Microbiol.* 41, 775–782.

Steunou, A.-S., Bhaya, D., Bateson, M. M., Melendrez, M. C., Ward, D. M., Brecht, E., et al. (2006). *In situ* analysis of nitrogen fixation and metabolic switching in unicellular thermophilic cyanobacteria inhabiting hot spring microbial mats. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2398–2403. doi: 10.1073/pnas.0507513103

Steunou, A.-S., Jensen, S. I., Brecht, E., Becraft, E. D., Bateson, M. M., Kilian, O., et al. (2008). Regulation of nif gene expression and the energetics of N$_2$ fixation over the diel cycle in a hot spring microbial mat. *ISME J.* 2, 364–378. doi: 10.1038/ismej.2007.117

Tank, M., and Bryant, D. A. (2015). Nutrient requirements and growth physiology of the photoheterotrophic Acidobacterium, *Chloracidobacterium thermophilum. Front. Microbiol.* 6:226. doi: 10.3389/fmicb.2015.00226

Taverner, T., Karpievitch, Y. V., Polpitiya, A. D., Brown, J. N., Dabney, A. R., Anderson, G. A., et al. (2012). DanteR: an extensible R-based tool for quantitative analysis of -omics data. *Bioinformatics* 28, 2404–2406. doi: 10.1093/bioinformatics/bts449

Van Der Meer, M. T., Schouten, S., Damste, J. S., and Ward, D. M. (2007). Impact of carbon metabolism on ^{13}C signatures of cyanobacteria and green non-sulfur-like bacteria inhabiting a microbial mat from an alkaline siliceous hot spring in Yellowstone National Park (USA). *Environ. Microbiol.* 9, 482–491. doi: 10.1111/j.1462-2920.2006.01165.x

Van Der Meer, M. T. J., Klatt, C. G., Wood, J., Bryant, D. A., Bateson, M. M., Lammerts, L., et al. (2010). Cultivation and genomic, nutritional, and lipid biomarker characterization of *Roseiflexus* strains closely related to predominant *in situ* populations inhabiting Yellowstone hot spring microbial mats. *J. Bacteriol.* 192, 3033–3042. doi: 10.1128/JB.01610-09

Ward, D. M. (1978). Thermophilic methanogenesis in a hot-spring algal-bacterial mat (71 to 30 degrees C). *Appl. Environ. Microbiol.* 35, 1019–1026.

Ward, D. M., Bateson, M. M., Ferris, M. J., Kühl, M., Wieland, A., Koeppel, A., et al. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 1997–2008. doi: 10.1098/rstb.2006.1919

Ward, D. M., Castenholz, R. W., and Miller, S. R. (2012). "Cyanobacteria in Geothermal Habitats," in *Ecology of Cyanobacteria II*, ed B. A. Whitton (Rotterdam: Springer), 39–63. doi: 10.1007/978-94-007-3855-3_3

Ward, D., Tayne, T., Anderson, K., and Bateson, M. (1987). "Community Structure and Interactions Among Community Members in Hot Spring Cyanobacterial Mats," in *Symposia of the Society for General Microbiology* (Cambridge).

Xie, G., Zhang, S., Zheng, X., and Jia, W. (2013). Metabolomics approaches for characterizing metabolic interactions between host and its commensal microbes. *Electrophoresis* 34, 2787–2798. doi: 10.1002/elps.201300017

Zeng, Y. B., Ward, D. M., Brassell, S. C., and Eglinton, G. (1992). Biogeochemistry of hot-spring environments. 3. Apolar and polar lipids in the biologically-active layers of a cyanobacterial mat. *Chem. Geol.* 95, 347–360. doi: 10.1016/0009-2541(92)90021-V

Conflict of Interest Statement: The Guest Associate Editor, William P. Inskeep, declares that although he has promoted collaboration in this research topic, he is not directly involved with the research reported in this paper, and has no relationship with the independent reviewers who provided comments to the manuscript. He confirms that the review process was handled objectively and that no conflict of interest exists. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The molecular dimension of microbial species: 1. Ecological distinctions among, and homogeneity within, putative ecotypes of *Synechococcus* inhabiting the cyanobacterial mat of Mushroom Spring, Yellowstone National Park

Eric D. Becraft[1,2]*, Jason M. Wood[1], Douglas B. Rusch[3], Michael Kühl[4,5], Sheila I. Jensen[4,6], Donald A. Bryant[7,8], David W. Roberts[9], Frederick M. Cohan[10] and David M. Ward[1]

[1] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA, [2] Single Cell Genomics Center, Bigelow Laboratory for Ocean Sciences, East Boothbay, ME, USA, [3] J. Craig Venter Institute, Rockville, MD, USA, [4] Marine Biological Section, Department of Biology, University of Copenhagen, Helsingør, Denmark, [5] Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, Ultimo, NSW, Australia, [6] The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hellerup, Denmark, [7] Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA, [8] Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA, [9] Department of Ecology, Montana State University, Bozeman, MT, USA, [10] Department of Biology, Wesleyan University, Middletown, CT, USA

Based on the Stable Ecotype Model, evolution leads to the divergence of ecologically distinct populations (e.g., with different niches and/or behaviors) of ecologically interchangeable membership. In this study, pyrosequencing was used to provide deep sequence coverage of *Synechococcus psaA* genes and transcripts over a large number of habitat types in the Mushroom Spring microbial mat. Putative ecological species [putative ecotypes (PEs)], which were predicted by an evolutionary simulation based on the Stable Ecotype Model (Ecotype Simulation), exhibited distinct distributions relative to temperature-defined positions in the effluent channel and vertical position in the upper 1 mm-thick mat layer. Importantly, in most cases variants predicted to belong to the same PE formed unique clusters relative to temperature and depth in the mat in canonical correspondence analysis, supporting the hypothesis that while the PEs are ecologically distinct, the members of each ecotype are ecologically homogeneous. PEs responded differently to experimental perturbations of temperature and light, but the genetic variation within each PE was maintained as the relative abundances of PEs changed, further indicating that each population responded as a set of ecologically interchangeable individuals. Compared to PEs that predominate deeper within the mat photic zone, the timing of transcript abundances for selected genes differed for PEs that predominate in microenvironments closer to upper surface of the mat with

spatiotemporal differences in light and O_2 concentration. All of these findings are consistent with the hypotheses that *Synechococcus* species in hot spring mats are sets of ecologically interchangeable individuals that are differently adapted, that these adaptations control their distributions, and that the resulting distributions constrain the activities of the species in space and time.

Introduction

Across a great diversity of microbial habitats, closely related populations have frequently shown distinct distributions along environmental gradients (West and Scanlan, 1999; Béjà et al., 2001; Ward and Cohan, 2005; Johnson et al., 2006; Walk et al., 2007; Hunt et al., 2008; Manning et al., 2008; Lau et al., 2009; Martiny et al., 2009; Miller et al., 2009; Connor et al., 2010; Denef et al., 2010; Shapiro et al., 2012; Kashtan et al., 2014). These patterns have prompted the hypothesis that some microorganisms might exist as ecological species occupying distinct niches (Cohan, 1994; Ward, 1998; Ward and Cohan, 2005; Smith et al., 2006; Sikorski, 2008; Vos, 2011), as theorized in the Stable Ecotype Model of species and speciation (Cohan and Perry, 2007). However, there are two fundamental unresolved issues with these kinds of studies. First, it is unclear whether the molecular variation that has been observed is sufficient to resolve populations at the species level, especially in studies based on variation in highly conserved molecular markers, such as 16S rRNA. Second, most studies address only the first of two expectations of the Stable Ecotype Model, that ecological species be distinct from one another and thus able to coexist indefinitely. The model also has the expectation that the membership within each species must be ecologically homogeneous, so that lineages within a species cannot coexist indefinitely (Koeppel et al., 2013; Kopac et al., 2014). This second quality has presented a challenge to microbial ecology because it requires sufficient phylogenetic resolution to detect individuals within the most newly divergent, ecologically distinct populations. Until recently, investigating the ecological interchangeability of individuals within a naturally occurring population has not been a priority of microbial speciology (Kopac et al., 2014).

This is the first of a three-paper series concerned with the degree of molecular resolution needed to observe microbial species. Here we report high-resolution, theory-based molecular identification of *Synechococcus* species in a well-studied hot spring microbial mat community, and we test the ecological distinctions of these species by investigating their distributions, responses to perturbation, and differences in gene expression. Our observations have led to the prediction that there are *Synechococcus* species with distinct light adaptations. Our second paper reports on the cultivation of strains representative of some of these species and demonstrates their adaptations (Nowack et al., 2015, this issue). The third paper compares the genomes of these strains, which have begun to demonstrate possible mechanisms underlying their light adaptations, as well as unsuspected adaptations to other environmental parameters

Olsen et al. (2015, this issue). As will be shown, these strains exhibited identical or nearly identical 16S rRNA gene sequences.

Earlier molecular studies of microbial mat communities in alkaline siliceous hot springs (Mushroom Spring and Octopus Spring, Yellowstone National Park, WY, USA) were limited by the molecular resolution of slowly evolving sequences. 16S rRNA gene analyses revealed several cyanobacterial (*Synechococcus*) variants identified as A″, A′, A, B′ and B, which exhibited different distribution patterns between 72–74°C and ~50°C along the effluent channel flow path (Ferris and Ward, 1997), or in vertical position within the photic zone of the mat (Ramsing et al., 2000). Others have found similar results using 16S rRNA variation to study hot spring cyanobacterial distributions (Jing et al., 2006; Lau et al., 2009; Miller et al., 2009). However, subsequent studies of Mushroom Spring and Octopus Spring mats employing the more rapidly evolving 16S–23S rRNA internal transcribed region revealed the existence of ecologically distinct populations with the same 16S rRNA sequence and prompted concerns about the possible need for yet higher molecular resolution to detect ecological species populations (Ferris et al., 2003). Indeed, population diversity analyses (Melendrez et al., 2011) and fine-scale distribution analyses of *Synechococcus* sequence diversity along effluent flow and vertical gradients (Becraft et al., 2011) using protein-encoding loci showed that molecular resolution influences the identification of ecological species populations in the Mushroom Spring mat.

In these most recent studies, we used an evolutionary simulation algorithm based on the Stable Ecotype Model of speciation to predict ecological species, rather than using an arbitrary molecular divergence cutoff. The algorithm Ecotype Simulation (Koeppel et al., 2008) hypothesizes from neutral single-gene or multi-locus sequence variation which variants should be grouped into different ecological species, each of which represents a set of ecologically interchangeable individuals. Each of these populations is considered a putative ecotype (PE), because the output of the algorithm is a set of hypothesized ecological species, whose predicted properties of ecological distinctness among and interchangeability within groups have not yet been demonstrated. Ecotype Simulation models the sequence diversity within a phylogenetic lineage as the evolutionary result of net ecotype formation, periodic selection (Koch, 1974), and drift, yielding the number of PEs within an environmental sample and demarcating sequence variants grouped into each PE.

In Becraft et al. (2011), genetic variation at the *psaA* locus (encoding a major Photosystem I reaction center protein subunit)

was analyzed because (i) it offers 7–10 times more molecular resolution than the 16S rRNA locus, (ii) it is known to be highly expressed *in situ* (Liu et al., 2012), (iii) it exists as a single copy in the genomes of mat *Synechococcus* isolates (Bhaya et al., 2007), (iv) the protein subunit it encodes (PsaA) is essential for photosynthesis, (v) it exhibits no evidence of recombination in the region studied and (vi) its structure is known (Jordan et al., 2001). While this gene might underlie some ecological distinctions among PEs, *psaA* is for our present purposes like any gene in the genomes of the populations being investigated. That is, the individuals within PEs are expected to accumulate neutral sequence divergence *in every gene* (including *psaA*), regardless of what physiological properties make PEs ecologically distinct. Thus, sequence clustering in any gene is expected to reveal ecotypes. To the extent that there is also adaptive divergence in *psaA*, this could supply additional resolution for distinguishing newly divergent ecotypes.

In Becraft et al. (2011), partial *psaA* sequences obtained by PCR amplification, cloning, and sequencing from samples collected along the effluent flow channel were used to demarcate *Synechococcus* PEs. The 320 cloned sequences analyzed suggested that each PE contained a dominant variant sequence that could, in most cases, be separated by denaturing gradient gel electrophoresis (DGGE) to demonstrate the distinctness of PE distributions along flow and vertical gradients. However, co-migration of DGGE bands made band purification and sequencing difficult, and without band sequences it was impossible to know whether different bands corresponded to the same or different PEs. Furthermore, the low number of sequences (320) and habitats sampled (25) limited the ability to identify different variants within PEs, making it impossible to examine the ecological interchangeability of individuals within a PE.

In this study we used pyrosequencing of PCR-amplified *psaA* genes and transcripts, which resulted in a more complete (>500 times greater coverage than in Becraft et al., 2011), sequence-based view of the genetic and ecological diversity within the *Synechococcus* populations in Mushroom Spring microbial mats. Pyrosequencing can overestimate the diversity within microbial populations due to a high rate of sequencing errors (Reeder and Knight, 2009), but because reference clone library sequences and genomes of representative *Synechococcus* isolates were available (Allewalt et al., 2006; Bhaya et al., 2007; Becraft et al., 2011), it was possible to recognize such artifacts by alignment to known reading frames. To test the ecological distinctions of PEs we (i) extended fine-scale distribution studies to include four temperatures and up to 12 depth intervals at multiple temperatures, (ii) investigated the responses of *Synechococcus* populations to perturbations of temperature and irradiance, and (iii) examined PE-specific transcription patterns over a diel cycle. Additionally, because the greater depth of coverage led to recovery of multiple sequence variants within PEs, it was possible to test for ecological homogeneity within PEs, based on the expectation that all members of an ecological species population would be co-distributed along environmental gradients and change uniformly in response to environmental perturbations. We provide evidence that most of the abundant

PEs predicted by Ecotype Simulation are ecologically distinct from one another, and that the individuals of a given PE are ecologically homogeneous.

Materials and Methods

Sampling

This study focused on the mat community inhabiting the $\sim$60–68°C region of the major effluent channel of Mushroom Spring, an alkaline siliceous hot spring in the Lower Geyser Basin, Yellowstone National Park, WY, USA. Samples were collected from sites spread over a distance of $\sim$10 m along the main flow path and were defined by temperatures measured at the time of collection (see Becraft et al., 2011 Supplementary Information). Mat samples for distribution studies were collected in duplicate on 12, 13, 14, and 15 September 2008 (60, 63, and 65°C), and 12 and 13 September 2009 (68°C) using a #2 cork borer (19.6 mm^2) and were immediately frozen in liquid N$_2$, or, in the case of cores for vertical analysis, in isopentane cooled with liquid N$_2$ to minimize decomposition of nucleic acids and to preserve core integrity (Ramsing et al., 2000). The latter cores were subsequently dissected at $\sim$80 μm intervals using a cryotome, as described in Becraft et al. (2011). Samples for transcript analysis over a diel cycle were taken at hourly intervals starting at 1700 h on 11 September 2009, and continuing until 1600 h on 12 September 2009 using a #4 cork borer (38.5 mm^2). Duplicate samples were collected from within a $\sim$1 m^2 area at a 60°C site and pooled.

Perturbation Experiments
Temperature Shift

Samples for temperature-shift experiments were collected in duplicate using a #2 cork borer, placed in their original vertical orientation into 3-ml glass vials that were filled with spring water, capped, and suspended in the effluent channel at a higher temperature site by aluminum wires attached to wooden stakes on either side of the channel (see Supplementary Figure S1A). Duplicate samples were retrieved 2 and 4 days after the disturbance was initiated on 27 October 2008. Previous studies showed that confining samples in this way did not alter the initial population structure of samples when incubated at the collection site (i.e., not shifted in temperature; Ruff-Roberts et al., 1994). Some bleaching was noticed in later stages of the incubation period, likely due to the lack of flow in the closed vials, which severely impacts diffusion, or possibly to other inhibiting factors, such as the accumulation of toxic metabolites.

Light Alteration

On 27 October 2008, a 254 cm^2 rectangular wooden frame covered with four layers of stretched muslin (Supplementary Figure S1B) was placed $\sim$2.5 cm above the mat surface at a $\sim$63°C site to reduce ambient solar irradiance by $\sim$92%. Duplicate samples were retrieved 2 and 4 days after the disturbance was initiated on 27 October 2008. All samples were immediately frozen on dry ice (-78.5°C) in the field and kept frozen at -80°C until analysis.

Molecular Methods

DNA was extracted and purified as described in Becraft et al. (2011), and RNA was extracted and purified as described in Liu et al. (2011). Primers for the amplification of *Synechococcus* A/B-lineage *psaA* genes (psaAcenterforward: 5′-TTCCACTACCACAAGCGGGCTCC-3′, psaAreverse: 5′-CAGGCCACCCTTGAAGGTG-3′) were designed to yield a 324 bp segment to maximize the number of single-nucleotide polymorphisms (SNPs) that could be used to differentiate PEs, and were tested for specificity to A/B′-like *Synechococcus* sequences as described in Becraft et al. (2011). The SNPs in this region of the *psaA* gene are unlikely to be under positive selection, as described in Supplementary Information Section I. The reduced sequence length in this study eliminated nucleotide diversity in some cases. This caused sequences representative of subclades PE A1-3 and A1-4, identified in Becraft et al. (2011) on the basis of 523-nt sequences, to be combined into a single high-frequency sequence variant (PE A1 in this study). Also, sequences representative of subclade PE A′9-2 were combined with those of PE A′9 [previously labeled A9 in Becraft et al. (2011); see **Table 1**, which compares A-like and A′-like PEs from this and the previous study]. Thus, it is possible that PEs A1 and A′9 each include >1 ecotype and that the sequence analyzed is too conserved to identify these populations accurately.

To estimate ecotype-specific expression patterns, environmental nucleic acids were treated with DNase, and *psaA* cDNA was synthesized from RNA using the psaAreverse primer (see above) with the SuperScript III First-Strand Synthesis Supermix (Invitrogen, Carlsbad, CA, USA) according to the instructions of the manufacturer, and the product was amplified by PCR according to protocols described in Becraft et al. (2011). Controls containing only RNA (i.e., no reverse transcription step) were used during amplification to insure that all DNA had been degraded. Barcoding and Ti454 sequencing were completed at the J. Craig Venter Institute according to the GS FLX Titanium Series Rapid Library Preparation Method Manual. DNA was sheared using the Covaris S2 System, and qPCR was used to estimate accurately the number of molecules needed for emulsion PCR[1]. Sequences were submitted to MG-RAST (4613896.3-4614007.3).

Identification of Sequences Containing Homopolymeric Artifacts

The resulting sequence data were first processed to remove homopolymeric errors (base pair insertions or gaps following strings of identical nucleotides) and sequences of poor quality. To generate alignments, we employed a Perl script from the Pigeon package (homopolymer-extinguisher.pl[2]) that uses ClustalW (Larkin et al., 2007) to align each raw sequence (plus its complement, reverse, and reverse complement sequences) with an in-frame consensus reference sequence of all *psaA* sequences from Becraft et al. (2011). Nucleotides in the raw sequences that caused a gap to form in the reference sequence reading frame

with the best alignment of these sequences were assumed to result from erroneous homopolymers (Quince et al., 2009, 2011; Reeder and Knight, 2009), and these nucleotides were excised. This step also trimmed the raw sequences to the same length as the reference sequence. Any trimmed raw sequences containing gaps relative to the consensus reference sequence were removed from further analysis. This yielded 164,467 sequences, with an average of 1,713 (SD of 251) for each of 96 unique environmental samples. Separately, 17 total cDNA samples yielded, on average, 360 sequences per sample (SD of 120). cDNA sequences were trimmed to a 247 bp segment to obtain the maximum number of sequences for analysis.

Putative Ecotype Demarcation

It was computationally too challenging to analyze all sequence variation using the current version of Ecotype Simulation. Thus, a Perl script from the Pigeon package (hfs-finder.pl[2]) was used to generate a list of high-frequency sequences by counting the number of occurrences of each unique sequence. Unless otherwise specified, high-frequency sequences were defined as unique sequence types with >50 identical representatives across all samples (96 DNA and 17 cDNA), which included the dominant variants of each PE identified in Becraft et al. (2011). High-frequency sequences showed no evidence of recombination in the region studied using the RDP3 software package (Martin et al., 2010). Each high-frequency sequence was assigned as either *Synechococcus* A-like or B′-like if it was ≥95% identical to the respective *psaA* genomic homologs in *Synechococcus* strain A (JA-3-3Ab) or B′ (JA-2-3B′a (2-13)) (Bhaya et al., 2007). Because no genome sequence was available for the A′ lineage [as defined by 16S rRNA sequence variation (Ferris and Ward, 1997)], we classified *psaA* sequences to the A′ lineage if they showed the highest similarity to 68°C metagenomic library sequences, which contain predominantly A′-like variants (see Supplementary Information in Becraft et al., 2011; Klatt et al., 2011), and these were included in the A-like phylogeny. Separate analyses were performed on the pool of A′-like plus A-like high-frequency sequences and for B′-like high-frequency sequences using Ecotype Simulation (with 1.5x sorting) to predict the number of PEs (Cohan and Perry, 2007; Koeppel et al., 2008). Neighbor-joining trees were constructed and uploaded into Ecotype Simulation as Newick files for ecotype demarcation. Two methods for Ecotype Simulation ecotype demarcation have been developed (Becraft et al., 2011; Melendrez et al., 2011; Francisco et al., 2012; Kashtan et al., 2014), both of which were used. The more conservative approach tends to yield PEs that are more inclusive clades; here a PE is demarcated as the most inclusive phylogenetic group for which the confidence interval for the number of predicted ecotypes includes the value 1. In our alternative fine-scale demarcation, PEs are demarcated as the largest groups for which the maximum likelihood solution for the number of predicted PEs equals the value 1. The Ecotype Simulation software and instructions for its use are freely available online[3].

---

[1]http://454.com/downloads/my454/documentation/gs-flx-plus/Rapid-Library-Preparation-Method-Manual_XLPlus_May2011.pdf

[2]https://github.com/sandain/pigeon

[3]http://fcohan.web.wesleyan.edu/ecosim/

TABLE 1 | Summary of non-singleton fine-scale demarcated A and A'-like putative ecotype (PE) average population percentages for individual habitat samples corresponding to distributions, expressing timing and responses to environmental perturbations.

PE Ti454	# CCA clusters	Becraft et al. (2011) designation[1]	Percent population and vertical orientation[2]												Perturbation response[3]	
			2006 cloning & sequencing					2008/09 Ti454 bar code								
			60°C	V	63°C	V	65°C	60°C	V	63°C	V	65°C	V	68°C	60→65°C	Light reduction
A1	1	A1 (A1-3, A1-4)	9.1	→	43.8	↔	6.2	2.6	→	2.9		26.8	↔	4.7	←	←
A2		A2	0.4		0.0		3.3	0.2		0.3		0.4		2.0	–	→
A3		A3	1.1		0.0		0.0	0.1		0.1		0.5		0.0	–	–
A4	1	A4	4.2		0.0	→	0.0	2.7	→	0.5		3.6	→	0.6	←	→
A5		A5	1.5		0.0	→	0.0	1.5		1.0		8.1	→	0.6	←	←
A6	1	A6	3.8		5.6		1.6	6.4	→	3.2		13.2	↔	1.9	→	→
A7	1	A7	1.1		50.0	←	1.6	0.0		0.2		4.5	↔	3.2	–	←
A'8		A'8	0.0		0.0		18.0	0.0		0.0		0.0		10.7	–	–
A'9		A'9 (A'9-2)	0.0		0.0		52.5	0.0		0.0		0.3		23.1	–	–
A10		A10						0.3		0.2		0.8		0.5	–	→
A11		A11						0.1		0.1		1.4		0.1	–	→
A12	1	A1-2	0.0		16.7		0.0	0.1		0.3		12.6	↔	1.8	–	←
A13								0.0		0.0		0.1		0.2	–	–
A14	1	A1-1	3.8		0.0		0.0	2.0	→	1.3		10.7	↔	0.6	←	←
A15								0.1		0.0		0.7		0.1	↑	–
A'16								0.0		0.1		0.2		8.3	–	–
A'17		A'9-1	0.0		0.0		11.5	0.0		0.0		0.0		14.6	–	–
A'18								0.0		0.0		1.2		0.4	–	←

Putative ecotypes without identified vertical distributions or perturbation responses (arrows) were too low in relative abundance to track with confidence. Highlighted rows correspond to predominant PEs discussed in the main text.
[1]*Putative ecotype subclades that were demarcated as a single ecotype in Becraft et al. (2011) that were separately demarcated in barcode analyses are indicated with hyphenated numbers next to PE designations. PE subclades that were not detected in barcode analysis due to decreased sequence length are indicated in parentheses next to the PE they were demarcated with in Becraft et al. (2011).*
[2]*'V' indicates vertical positioning at a specified temperature. Arrows under 'V' indicate highest relative abundance in upper green mat layer (↑), highest relative abundance in lower portion of the upper green mat layer (↓), highest relative abundance in center layers (↔), and relative abundance through all layers (↕).*
[3]*Perturbation responses are indicated by ↑ for increase in relative abundance, ↓ for decrease, and – for no change.*

Determination of All Variants within Putative Ecotype Populations

The set of high-frequency sequences was de-replicated (i.e., to include only one example of each sequence). Two neighbor-joining trees were created, one for all unique A-like plus A′-like sequences, and the other for B′-like sequences, using Phylip's dnadist and neighbor-joining programs (Felsenstein, 1989). When a low-frequency sequence fell within a clade of high-frequency sequences that were demarcated to a single PE, it was classified to that PE. When a low-frequency sequence did not fall within a clade of high-frequency sequences, it was tentatively assigned to the most closely related PE comprised of high-frequency sequences. All high-frequency sequences (and associated low-frequency sequences) were combined to obtain the total number of variants detected for each PE in each sample (PE sequences = dominant variant sequence + all other high-frequency sequences + all low-frequency sequences). The percentage contribution of each PE in each environmental sample was calculated by dividing the total number of sequences within a PE detected in a sample by the total number of *psaA* sequences in that sample [(PE sequences in sample/all associated sequences in sample) × 100]. Absolute quantitation of each PE's abundance could not be achieved in these analyses because dispersed SNP patterns prevented the development of SNP-specific primers or probes, and the PCR and sequencing method employed only sampled proportions of populations in the data.

Gene Expression Analyses

Transcripts obtained in sequence analysis were aligned with *psaA* gene sequences, which had been demarcated into PEs, to determine the number of transcripts per PE in each sample (no mismatches were allowed). This was divided by the total number of *psaA* transcripts retrieved in a sample, and the results are expressed as percent relative abundance at each time point.

Diel metatranscriptomic datasets described by Liu et al. (2012) were analyzed by using the Burrows–Wheeler Aligner (BWA) to identify transcripts for specific photosynthesis and N$_2$ fixation genes (see Liu et al., 2012 for genes analyzed) associated with the genomes of either *Synechococcus* A (JA-3-3Ab) or B′ (JA-2-3B′a (2-13)) (Bhaya et al., 2007). We used the methods described in Liu et al. (2011, 2012), except that we recruited transcripts using genomes, instead of metagenomic assemblies. Because the transcript sequences were only 50 nt in length, we allowed up to 5 SNP differences to capture transcripts that matched with the diverse variants of B′-like and A-like populations. Raw transcript counts were normalized by the total number of transcripts recruited by that genome at each time point and then by the geometric mean of normalized transcript counts for all time points (Liu et al., 2011, 2012).

Canonical Correspondence Analysis

The distributions of high-frequency *psaA* sequence variants in vertical subsections of duplicate cores, which had been collected from 60, 63, and 65°C sites, were analyzed using canonical correspondence analysis (CCA; Ter Braak, 1986; Legendre and Legendre, 1998) from the Vegan R package (Oksanen et al., 2013).

A custom version of ordtest from the labdsv R package (Roberts, 2013) was used to test each PE for nonrandom distribution, and a custom plotting function was written to display the data (CCA-plot.R, Supplemental Information). This permitted us to evaluate whether high-frequency sequence variants of the same PE showed restricted distributions in the ordination space, and whether high-frequency sequence variants of different PEs showed disjunct distributions. The variation in distribution among variants was analyzed with respect to temperature and depth as linear predictors.

Statistical Analyses

G-tests were conducted on *psaA* sequence counts along all environmental gradients for abundant PEs, and for A-like and B′-like PEs separately, to determine whether there was statistical significance for heterogeneity of PE distributions (see Supplementary Table S1). Analysis of co-variance (ANCOVA) tests were conducted for perturbation experiments to determine whether PE populations differed in their response to environmental change (see Supplementary Table S2). *G*-tests and ANCOVA analyses were done in Stata (StataCorp LP, Stata Statistical Software: Release 13, College Station, TX, USA). Additionally, over the course of the light-reduction and temperature-shift experiments, we tested whether PEs changed their relative frequencies, and we tested whether the membership of high-frequency sequences (and associated low-frequency sequences) in a given PE changed in unison. This was done by evaluating whether the percentage of each high-frequency sequence within a PE, and the ratio of low-frequency sequences to high-frequency sequences within a PE, changed separately over time using a generalized linear model (a flexible generalization of ordinary linear regression that allows for binomial distributions; see Supplementary Table S3).

Microsensor Measurements

Diel sampling for molecular analyses was closely coordinated with simultaneous logging of the incident downwelling solar irradiance and *in situ* microsensor measurements of O$_2$ concentration profiles in the microbial mats at the sample site following the same calibration and measurement procedures described in detail in earlier studies (Becraft et al., 2011; Jensen et al., 2011). At the end of the diel measurements, ∼3 cm^2 of mat was sampled with a glass corer and transported to the laboratory in spring water for subsequent measurements (within 24 h) of spectral light penetration using fiber-optic scalar irradiance microsensors (see details in Kühl, 2005; Ward et al., 2006).

Scalar irradiance spectra measured at particular depths of the mat were normalized to the incident downwelling irradiance at the mat surface and expressed in percent of downwelling irradiance at the mat surface. Additionally, measured scalar irradiance spectra in each depth were integrated over 400–700 nm [i.e., photosynthetic active radiation (PAR)] and normalized to the downwelling irradiance of PAR at the mat surface. This depth profile of E$_0$(PAR) (in percent of downwelling irradiance) was then used together with the actual measured downwelling photon irradiance at the field site during the diel study to construct depth profiles of *in situ* E$_0$(PAR) in units of

μmol photons m^{-2} s^{-1} for all sampled time points in the diel study.

Isopleth diagrams of O_2 concentration and photon scalar irradiance, E_0(PAR) depth distribution over the diel cycle were constructed from depth profiles measured at different time points using the graphing software Origin Pro 8.5 (Origin Lab Corp., Northampton, MA, USA).

Results

Ecotype Simulation Analysis and Sequence Variation within and Among Putative Ecotypes

A total of 119 unique high-frequency sequences were found in the variation detected in 113 environmental samples (96 DNA; 17 cDNA), and these sequences were analyzed using Ecotype Simulation to predict PEs (**Figure 1**). Ecotype Simulation analysis of high-frequency sequences from the clade of A-like and A′-like

lineages predicted 22 PEs using fine-scale demarcation (15 and 7 PEs, respectively) and analysis of the B′-like lineage predicted 24 PEs in conservative demarcation analyses (**Tables 1** and **2,** and **Figure 1**). Results from the alternative demarcation approaches are also shown in **Figure 1**. The choice of demarcation approach was initially guided by PE distributions identified in Becraft et al. (2011) and high-frequency sequence distributions in the current study.

Many PEs predicted from sequences obtained by pyrosequencing (2008 collections) corresponded to PEs previously demarcated in cloning and sequencing studies (2006 collections; Becraft et al., 2011) and this correspondence is shown in **Tables 1** and **2**. All 22 of the newly discovered PEs were either based on a single sequence or were in low frequency (ranging from 0.05 to 3.8% of total sequences; average 0.8% $\pm$ 0.14 SE). Among these only PE B′15 had a frequency greater than 1% of the total sequences (Supplementary Tables S4–S13). Newly discovered PEs in pyrosequencing analyses were

FIGURE 1 | Comparison of putative ecotypes (PEs) demarcated from high-frequency *psaA* sequences by Ecotype Simulation using the fine-scale and conservative demarcation approaches in the (A) A/A′ lineage and the (B) B′ lineage. Black vertical bars indicate PEs demarcated using the conservative approach and red vertical bars indicate PEs demarcated using the fine-scale approach. Numbers indicate the total number of sequences sampled with a particular sequence across all 96 DNA samples. Large filled

circles represent PEs newly identified in the present sequencing analysis, and large open circles indicate previously identified ecotype subclades in cloning and sequencing studies now demarcated as separate PEs. Small filled circles indicate PEs based on single high-frequency sequences. Colored PE designations represent predominant populations discussed in the main text, and colors correspond across all figures. Scale bars represent **(A)** 0.005 and **(B)** 0.002 substitutions per site.
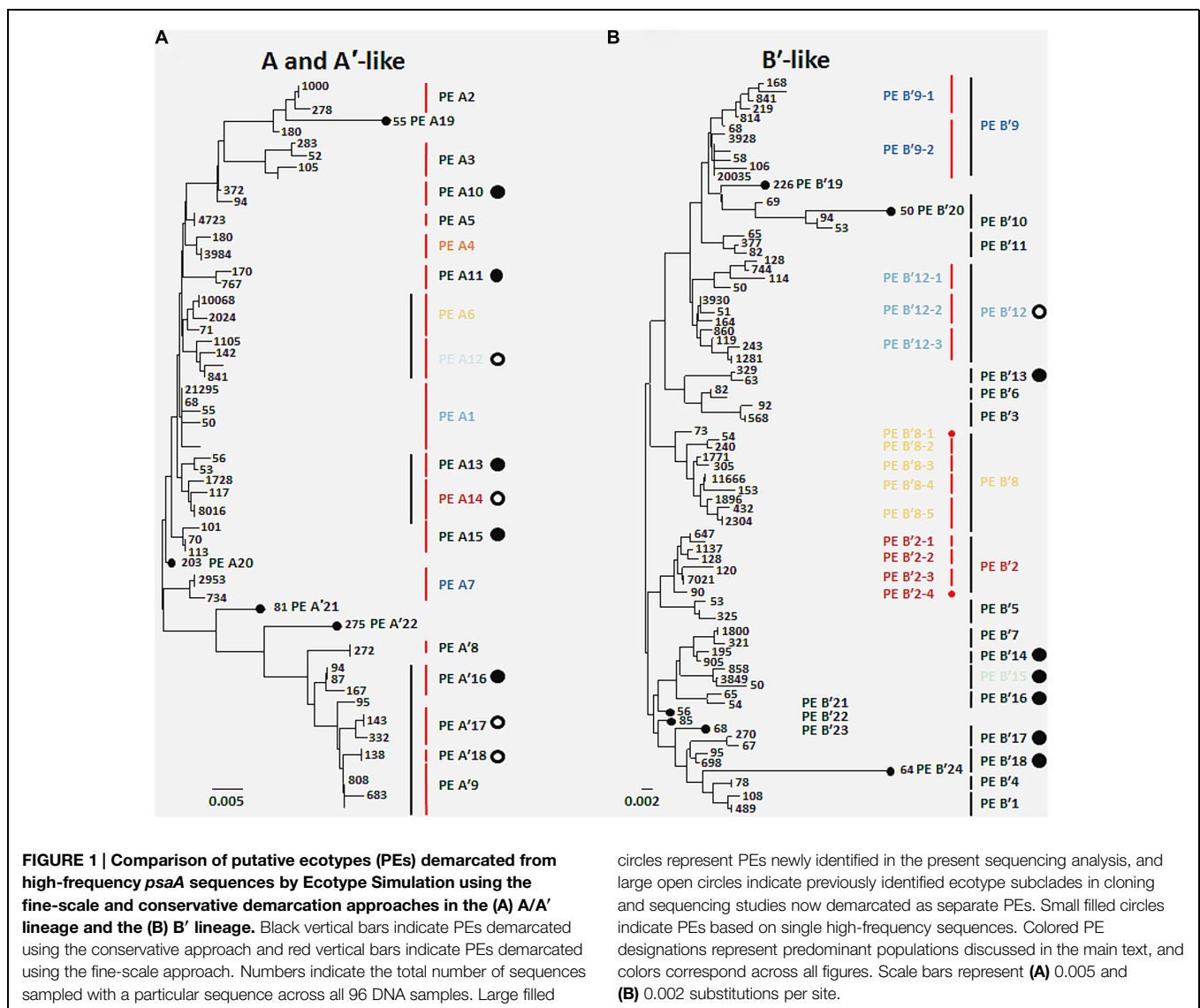
TABLE 2 | Summary of non-singleton conservatively demarcated B′-like PE average population percentages for individual habitat samples corresponding to distributions, expression timing and responses to environmental perturbations.

			Percent population and vertical position[2]												Perturbation response[3]	
			2006 cloning & sequencing					2008/09 Ti454 bar code								
PE Ti454	# CCA clusters	Becraft et al. (2011) designation[1]	60°C	V	63°C	V	65°C	60°C	V	63°C	V	65°C	V	68°C	60→65°C	Light reduction
B′1		B′1	6.9		0.0		0.0	0.4		0.3		0.0		0.4	–	–
B′2*	1	B′2	3.1		0.0		0.0	15.7		9.7		1.4		0.5	↑	–
B′3		B′3	2.3		0.0		0.0	0.1		0.3		0.0		0.9	–	–
B′4		B′4	4.2		0.0		0.0	0.0		0.2		0.0		0.0	–	–
B′5		B′5	1.1		0.0		0.0	0.0		0.3		0.1		0.0	–	–
B′6		B′6	1.1		0.0		0.0	0.1		0.0		0.0		0.8	–	–
B′7	1	B′7	8.0	↑	0.0		0.0	0.8		5.2	↑	0.1		0.2	–	–
B′8*	3	B′8	8.4		0.0		0.0	8.6	↑	2.5		1.6		1.0	↑	–
B′9*	1	B′9	29.9	↑	0.0	↑	0.0	23.5	↑	27.9	↑	3.4		3.2	↑	↑
none		B′10	1.1		0.0		0.0	0.2		0.0		0.2		0.0	–	–
none		B′11	1.1		0.0		0.0	0.6		1.0		0.1		0.1	–	–
none		B′12	0.8		0.0		0.0									
B′12*	2	B′9-1 and B′9-2	11.1		0.0		0.0	4.8	↑	7.8		1.0		1.1	↑	–
B′13		B′13						0.1		0.2		0.0		0.1	–	–
B′14		B′14						0.0		1.0		0.0		0.1	–	↓
B′15*	1	B′15						0.2		0.3	↑	5.1	↑	1.0	↓	–
B′16		B′16						0.0		0.4		0.0		0.0	–	–
B′17		B′17						0.9		0.8		0.0		0.0	–	–
B′18		B′18						0.8		0.7		0.1		0.0	–	–

PEs without identified vertical distributions or perturbation responses (arrows) were too low in relative abundance to track with confidence. Highlighted rows correspond to predominant PEs in main-text.

[1] Putative ecotype subclades that were demarcated as a single ecotype in Becraft et al. (2011) that were separately demarcated in barcode analyses are indicated with hyphenated numbers next to PE designations.

[2] 'V' indicates vertical positioning at a specified temperature. Arrows under 'V' indicate highest relative abundance in upper green mat layer (↑), highest relative abundance in lower portion of the upper green mat layer (↓), highest relative abundance in center layers (↔), and relative abundance through all layers (↕).

[3] Perturbation responses are indicated by ↑ for increase in relative abundance, ↓ for decrease, and – for no change.

*B′-like PEs that contained one or more high frequency sequences that did not correspond with the most abundant high-frequency sequence distributions in CCA analysis, and could be representative of multiple PEs.

labeled sequentially following the PEs previously identified and numbered in Becraft et al. (2011). Ecotype Simulation calculated the following rates of periodic selection (sigma) and ecotype formation (omega): 2.71 (sigma) and 0.94 (omega) for the A-like plus A′-like sequences, and 1.86 and 0.94 for B′-like sequences, parameterized as the number of events (periodic selection or ecotype formation) per nucleotide substitution in the 324 bp of the *psaA* gene. These results indicated that the rate of periodic selection was higher than the ecotype formation rate.

Each high-frequency sequence within a PE was associated with a set of low-frequency sequences (<50 identical sequences in the entire database), which typically exhibited randomly distributed differences in SNPs (examples are shown in Supplementary Figure S3). The number of unique high-frequency sequences in each well-sampled PE (i.e., average of 1713 total sequences) was lower in A-like PEs (range 1–4; average 2.67 ± 0.42 SE) than in B′-like PEs (range 3–11; average 7.8 ± 1.46 SE; $P = 0.0052$ for a two-tailed t-test). Because we classified many high-frequency sequences per PE, we had an opportunity to test whether high-frequency variants predicted to be members of a PE were ecologically interchangeable. For clarity, in-text figures focus on predominant PEs defined as those that were >5% of the population in any sample and well-represented in the barcode sampling (i.e., >3000 sequences). This includes PEs that were unresolved in Becraft et al. (2011; shaded in **Tables 1** and **2**); results for all PE populations are presented in Supplementary Tables S4–S13.

Distribution Along the Effluent Flow Path and with Depth within the Upper Green Mat Layer

Most predominant PEs exhibited strong and largely unique associations with environmental gradients. For instance, PEs from the A′-like, A-like and B′-like were clearly distributed to high, middle and low temperature sites, respectively (**Figure 2**). At a given temperature, predominant PEs were also distributed differently with respect to depth in the upper green layer of the mat. For instance, at 60–63°C we observed a stratification of PE B′9 above PE A1 above PE A4 above PEs A6 and A14 (**Figure 3**).

This stratification paralleled changes in the light environment over the upper 1 mm-thick green surface layer of the mat. Microsensor analyses of scalar irradiance demonstrated that both the intensity and spectral composition changed dramatically with depth in the upper green layer of the mat (**Figure 4**). The uppermost 0.2 mm layer showed moderate attenuation of visible and far-red light, while strong attenuation was found in mat layers 0.2–0.3 mm below the mat surface. The spectral minima indicated presence of chlorophyll *a* (Chl *a*), bacteriochlorophyll *c/d* (Bchl *c/d*) and phycobiliproteins (PBPs) as the major pigments. With increasing depth, absorption maxima in the spectral range 600–630 nm, presumably representing phycobiliproteins, showed a distinct blue-shift.

Canonical Correspondence Analyses

Where high-frequency sequences predicted to belong to the same PE were sufficiently abundant to analyze, they appeared

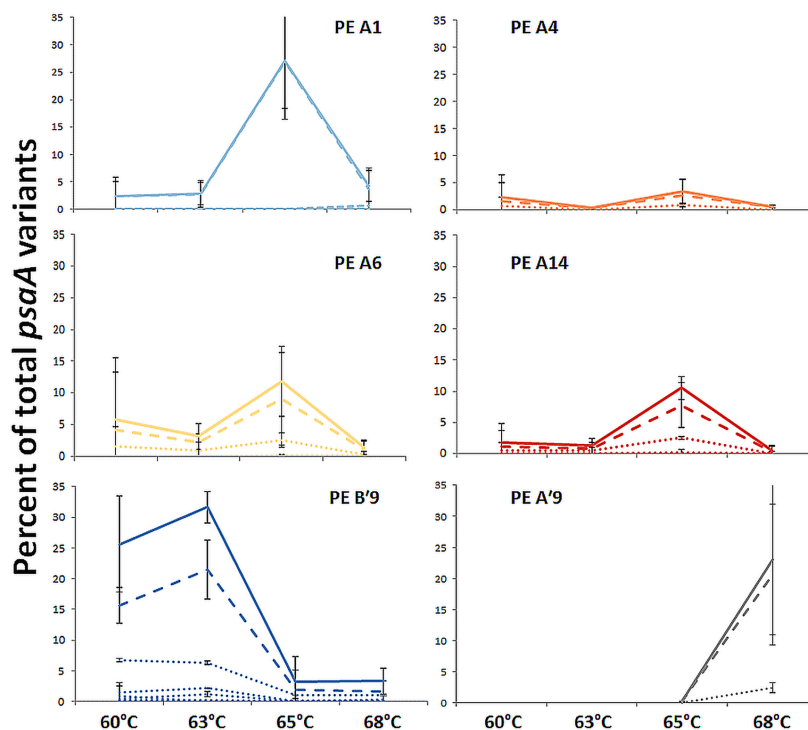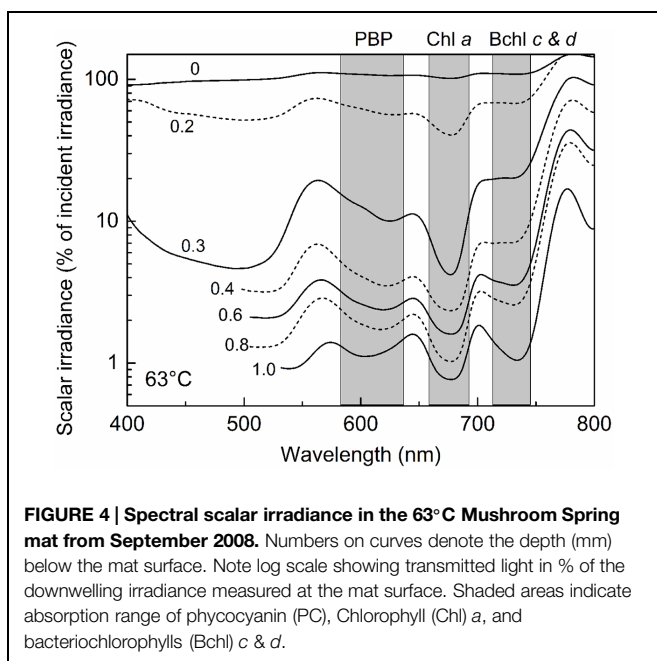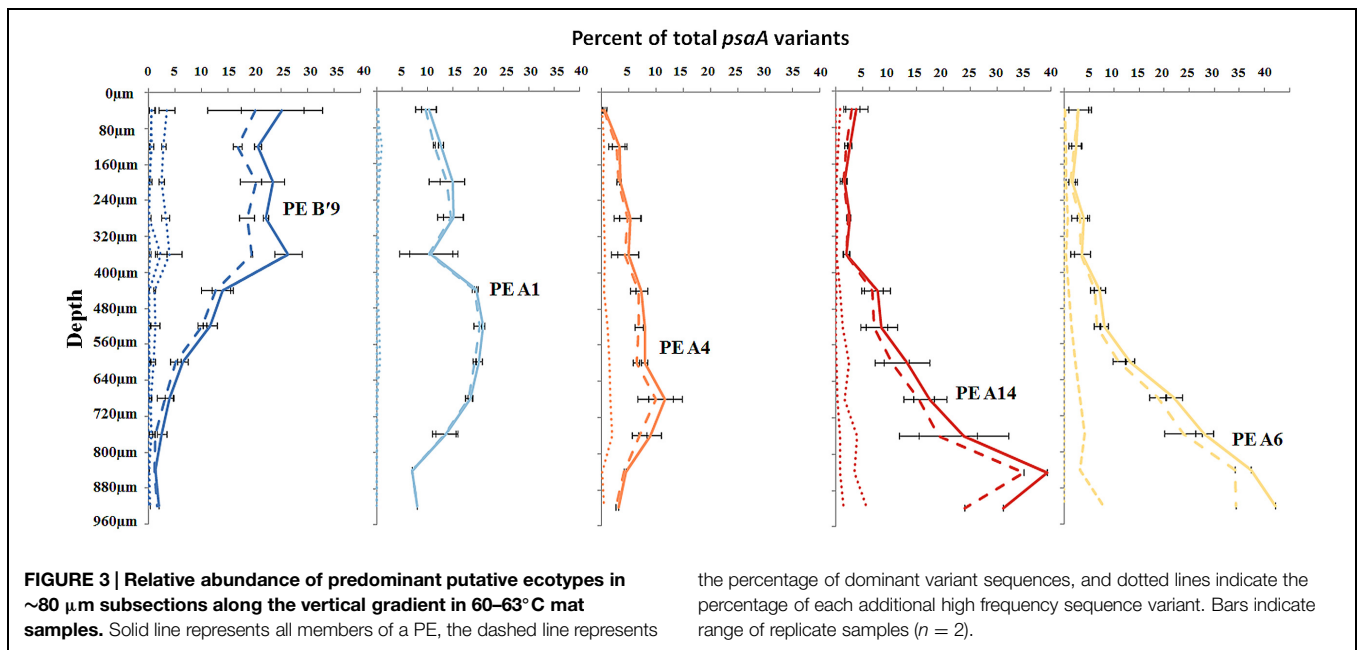**FIGURE 2 | Relative abundance of predominant PEs along the effluent flow path from 60 to 68°C.** Solid line represents all members of a PE, the dashed line represents the percentage of dominant variant sequences, and dotted lines indicate the percentage of each additional high frequency sequence variant. Bars indicate range of replicate samples (*n* = 2). Temperature distributions of other PEs are shown in **Tables 1** and **2**.
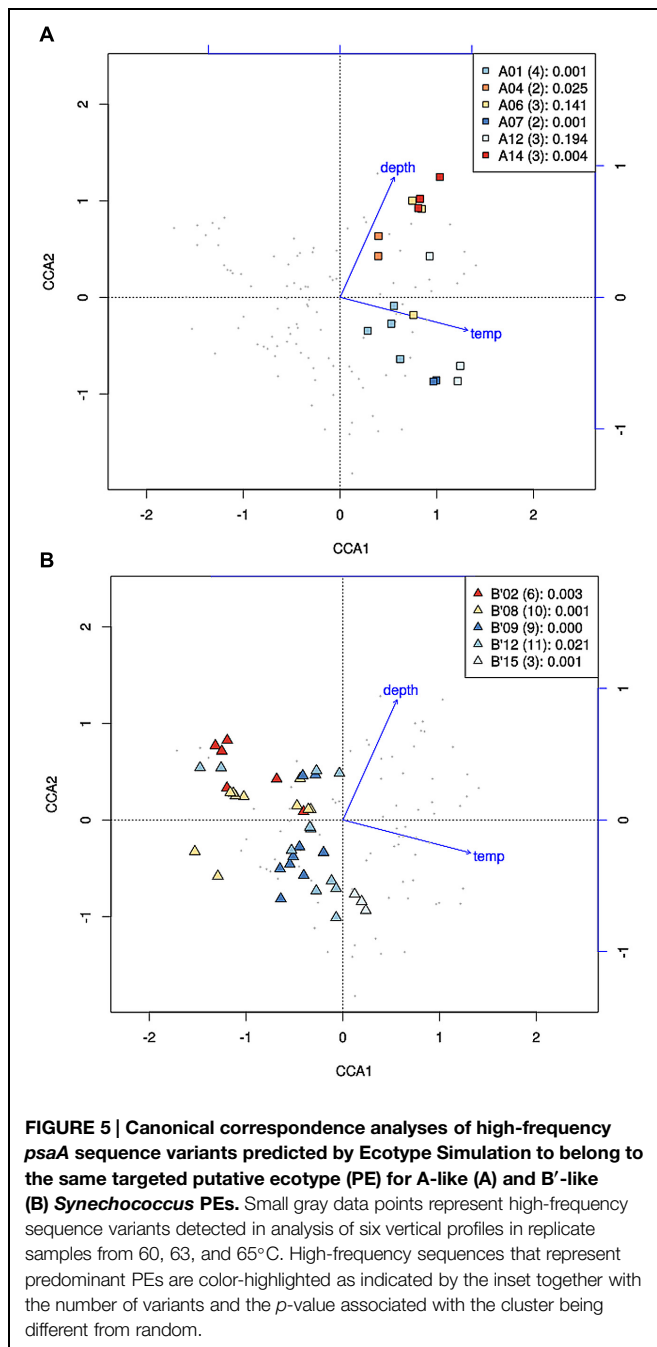
FIGURE 3 | Relative abundance of predominant putative ecotypes in ~80 μm subsections along the vertical gradient in 60–63°C mat samples. Solid line represents all members of a PE, the dashed line represents the percentage of dominant variant sequences, and dotted lines indicate the percentage of each additional high frequency sequence variant. Bars indicate range of replicate samples (n = 2).

FIGURE 4 | Spectral scalar irradiance in the 63°C Mushroom Spring mat from September 2008. Numbers on curves denote the depth (mm) below the mat surface. Note log scale showing transmitted light in % of the downwelling irradiance measured at the mat surface. Shaded areas indicate absorption range of phycocyanin (PC), Chlorophyll (Chl) *a*, and bacteriochlorophylls (Bchl) *c* & *d*.

to co-vary along temperature and depth transects (**Figures 2** and **3**). That is, each constituent sequence within a PE changed proportionately with the entire set of sequences from the PE. Environmental associations among the high-frequency sequences within A-like and B′-like PEs recovered from vertical sections of cores collected at 60, 63, and 65°C were tested in a single CCA analysis, and the results for predominant PEs are displayed separately for each lineage in **Figures 5A,B** (see Supplementary Figure S4 for non-predominant PEs, including A′ PEs). The analyses demonstrated strong evidence of clumped

distributions of the various high-frequency sequences within each predominant PE, with respect to temperature and depth (Ordtest analyses, $P < 0.001$).

High-frequency sequences assigned to most A-like PEs formed clusters based on their temperature and depth distributions that were unlikely to have formed by chance (significance levels for PEs were all $p \leq 0.025$). This suggested ecological homogeneity among variants within PEs A1, A4, A7, and A14. Variants predicted to comprise PEs A6 and A12 were not significantly clustered ($p = 0.141$ and 0.194), but these PEs each contained three high-frequency sequences, two of which formed a significant cluster to the exclusion of the third variant, which was much more rare (**Figures 6A,B** and see **Figure 1A**). The most abundant PE A6 variants were distributed at depths that overlapped with variants of PE 14. The most abundant PE A12 variants were distributed near the surface of the mat at higher temperatures. All other A-like PEs exhibited unique distributions. PEA7 was found at higher temperatures than PEs A1, A4, A6, and A14, which were distributed with depth in the same order as shown in analyses of all variants within PEs (**Figure 3**). Conservative demarcation lumped clades that were, in some cases (e.g., PEs A6 and A12), ecologically distinct (**Figure 1A**).

Each predominant B′-like PE demarcated using the conservative approach formed a cluster unlikely to have formed by chance ($p \leq 0.023$; **Figure 5B**), again suggesting ecological homogeneity among variants within all PEs. For example, PE B′15 variants formed a tight cluster distributed toward warmer surface samples. PE B′2 formed a looser cluster distributed in cooler and deeper samples. Fine-scale demarcations permitted us to explore the structure of conservatively demarcated B′-like PEs in greater detail. For instance, fine-scale demarcation predicted that PE B′2 contained three ecotypes (these are designated PE B′2-1, -2, and -3; **Figure 1B**), but CCA analyses did not provide statistical support for this (**Figure 6C**). PE B′9 clustered mainly

FIGURE 5 | Canonical correspondence analyses of high-frequency *psaA* sequence variants predicted by Ecotype Simulation to belong to the same targeted putative ecotype (PE) for A-like (A) and B′-like (B) *Synechococcus* PEs. Small gray data points represent high-frequency sequence variants detected in analysis of six vertical profiles in replicate samples from 60, 63, and 65°C. High-frequency sequences that represent predominant PEs are color-highlighted as indicated by the inset together with the number of variants and the *p*-value associated with the cluster being different from random.

split these PEs into five and three ecotypes, respectively (**Figure 1B**), many of which clustered uniquely (**Figures 6D,F**). The most abundant fine-scale B′8 PE, PE B′8-4, was distributed in warmer deeper samples and was separated from fine-scale PE B′8-3, which was distributed toward cooler, deeper samples, but the PE B′8-5 distribution overlapped with these two PEs. PE B′8-2 variants formed a cluster that was distributed in cooler surface environments. The one significantly distinct fine-scale B′12 PE, PE B′12-3, was distributed in samples collected in surface layers of warmer mats.

Although coverage was deep, the number of unique high-frequency sequence variants detected per PE was relatively low (i.e., 1–4 for A-like and 1–10 for B′-like PEs). Thus, we also performed CCA analyses in which we included all variants that occurred at least 10 times in the dataset (as opposed to 50 times), increasing the number of variants per PE 2- to 10-fold. Clustering of all predominant PEs remained significant with 54, 23, and 56 variants in the case of PEs A1, B′8-4 and B′9-2, respectively. The only change in results was that subsets of PEs A6 and A12 and fine-scale PE B′8-5 were no longer statistically supported, and subclade B′12-1 became statistically significant in CCA analyses.

CCA also found that high-frequency sequence variants within B′-like, A-like and A′-like PEs were progressively distributed from lower to higher temperatures, respectively (compare **Figures 5A,B** and see Supplementary Figure S4). The vertical stratification of PE B′9, above PE A1 above PE A4 above PEs A6 and A14 was also observed (compare **Figures 5A,B**). Temperature and vertical position explained ∼19 and ∼8% of the variation in the distribution of PEs among samples, respectively, which suggests that other parameters must also be important in defining the niches of these PEs.

Responses to Environmental Perturbations

Environmental perturbation studies were used to test whether differences in PE distributions reflect adaptations to different temperature and irradiance levels and whether the high-frequency sequences within a given PE respond homogenously. Results are summarized in **Tables 1** and **2**. As described in Supplementary Section III (Supplementary Figure S5), during similar light alteration experiments in 1996 populations, control samples remained stable over the period of time studied ($p = 0.35$; see Supplementary Table S2). The observed changes in relative abundance (up to fivefold in 2 days) are easily accommodated by observed growth rates of 1–3 doublings per day at optimal temperature and irradiance (Nowack et al., 2015, this issue).

Temperature Increase

The PE composition changed in the 4 days following a shift of samples collected at a 60°C site to a 65°C site (**Figure 7A**; Supplementary Table S13). The most abundant PEs in the temperature manipulation experiment were significantly different from one another in their responses to increased temperature (ANCOVA test of covariance with time, $P = 0.0155$, $F = 3.83$, $df = 5.18$; see Supplementary Table S2). In particular, PE B′9 declined, while PEs A1 and A14 increased in relative abundance. Effectively, the composition of abundant *Synechococcus* PEs within the sample shifted from that roughly
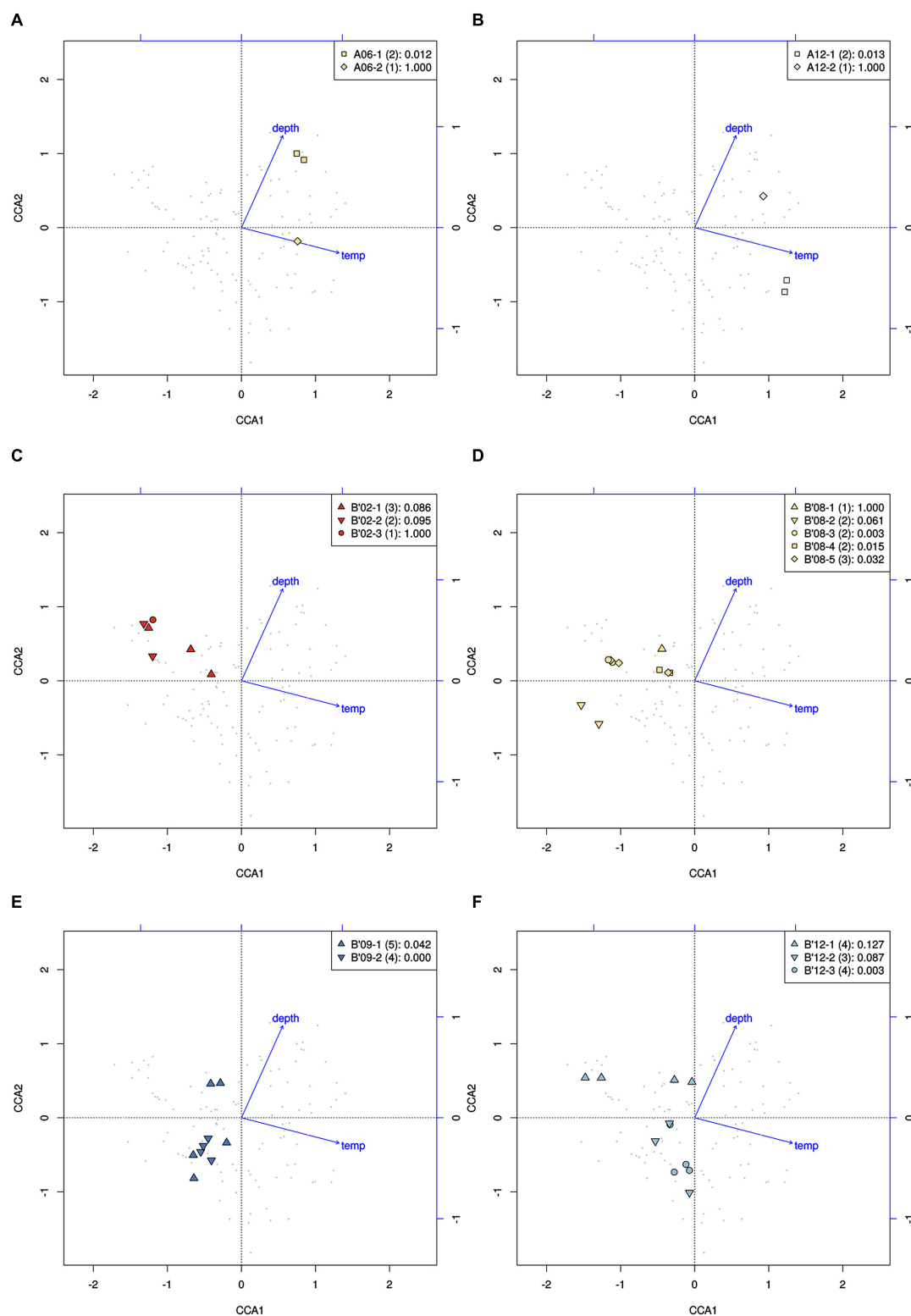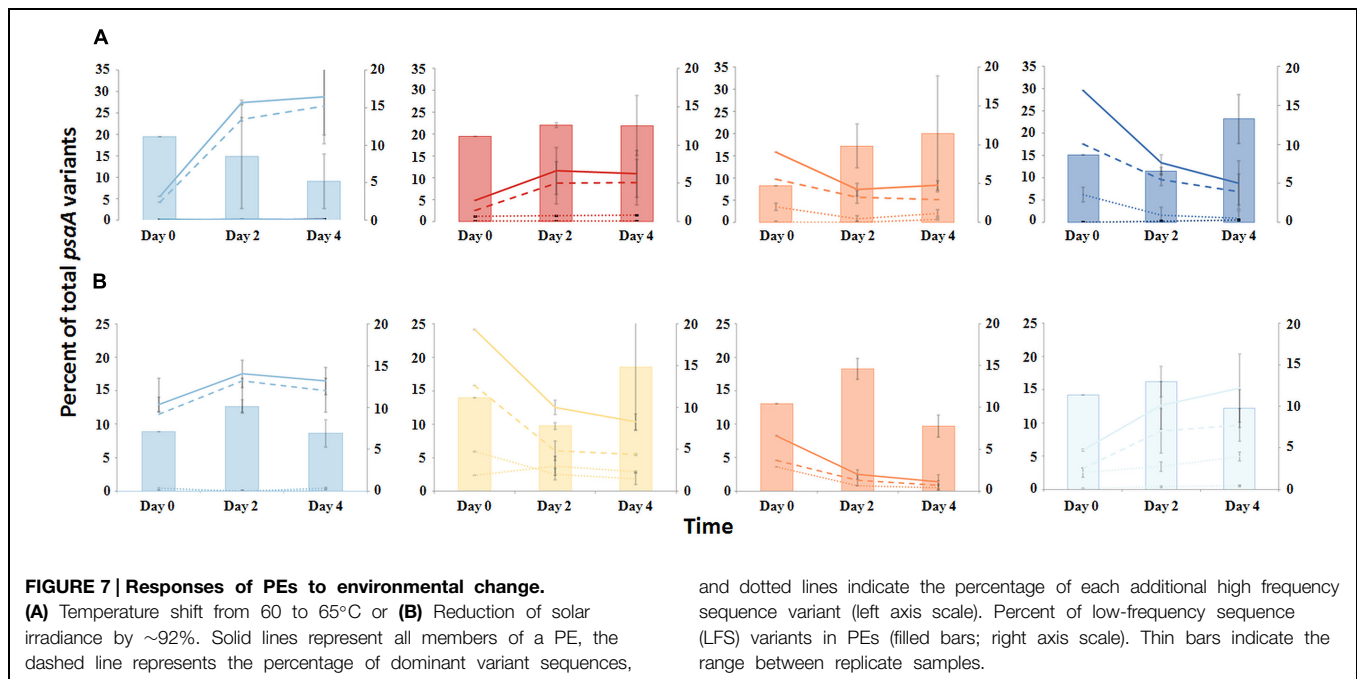
near the surface of warmer samples, with the exception that two variants were found in deeper samples. Fine-scale demarcation suggested that conservatively demarcated PE B′9 contains two ecotypes (PE B′9- and PE B′9 -2; **Figures 1B** and **6E**). PE B′9-2 represented 90% of the population and was surface-associated. PE B′9-1 represented ∼10% of the population and contained the two variants that were more deeply distributed and other surface-associated variants, possibly explaining the subsurface bulge in the PE B′9 profile shown in **Figure 3A**.

Putative ecotypes B′8 and B′12 exhibited broader distributions relative to temperature and mat depth. Fine-scale demarcation

FIGURE 6 | Canonical correspondence analyses of high-frequency *psaA* **sequence variants predicted by Ecotype Simulation to belong to the same PE. (A,B)** Clustering of predominant variants (squares) to the exclusion of the least abundant variant (diamonds) within A-like PEs demarcated using the fine-scale approach. **(C–F)** Clustering of variants within PEs demarcated by the fine-scale approach (different symbols) that were demarcated as a single PE by the conservative approach (all points). Legends indicate how symbols correlate with PEs shown in **Figure 1** and include the *p*-value associated with the cluster being different from random.

FIGURE 7 | Responses of PEs to environmental change.
(A) Temperature shift from 60 to 65°C or **(B)** Reduction of solar irradiance by ~92%. Solid lines represent all members of a PE, the dashed line represents the percentage of dominant variant sequences, and dotted lines indicate the percentage of each additional high frequency sequence variant (left axis scale). Percent of low-frequency sequence (LFS) variants in PEs (filled bars; right axis scale). Thin bars indicate the range between replicate samples.

characteristic of the 60°C mat to that roughly characteristic of the 65°C mat, as expected from temperature distributions (Supplementary Table S4; Becraft et al., 2011). While PEs A6 and A14 had similar vertical distributions at 60–63°C sites (**Figure 3**), all variants within PE A6 decreased, while all variants within PE A14 increased after shifting from 60°C to 65°C, suggesting that these PEs have different optimal growth temperatures.

Light Alteration

Light alteration experiments were conducted at 63°C in order to focus on very closely related A-like PEs (**Figure 7B**; Supplementary Table S5). The most abundant PEs were significantly different in their responses to removal of 92% of the ambient light (ANCOVA test of covariance with time, $P = 0.0262$; see Supplementary Table S2A). After shading, PEs A4 and A6 declined, whereas PEs A1 and A12 increased in relative abundance (Supplementary Table S14). Thus, while PEs A1 and A4 had similar distributions along the vertical aspect of the mat at 60–63°C sites (**Figure 3**), they had different responses to light reduction.

Homogeneity of Responses within PEs

Most high-frequency sequences within a given PE appeared to respond similarly to the perturbations (**Figure 7** and Supplementary Table S2). We tested for homogeneity of response within an abundant PE by addressing whether the pool of all low-frequency sequences changed its relative frequency within the PE after an environmental perturbation. The proportion of low-frequency sequence variants among all sequences within a PE changed little over the environmental perturbation (by 0.3–7.6%) (generalized linear model analysis, $P > 0.34$; see Supplementary Table S3), despite 1.5- to 5-fold increases or decreases in PE

relative abundances in response to environmental perturbations (**Figures 7A,B**).

Transcription Patterns

Sequencing of *psaA* transcripts from 60°C samples collected over a diel cycle demonstrated differential expression of transcripts contributed by the dominant *Synechococcus* PEs B′9 and A1 (**Figure 8A**; G-test, $P < 0.001$; Supplementary Figure S3 and Table S1). The expression of results as relative percentages limited our understanding of the basis for these differences. However, we were able to use the genome sequences of isolates representative of PEs *Synechococcus* B′ and A to recruit separately A-like and B′-like transcripts of photosynthesis and N$_2$ fixation genes from a metatranscriptome of this mat produced over the same diel cycle (Liu et al., 2012). In the case of photosynthesis genes, B′-like transcripts were expressed before A-like transcripts (**Figure 8B**). Because both B′-like and A-like photosynthesis genes were transcribed diurnally, and because PE B′9 and PE A1 transcripts dominated the pools of B′-like and A-like *psaA* transcripts, the decreased relative abundance of PE B′9 transcripts in mid-day (**Figure 8A**) was apparently due to the increase in relative abundance of PE A1 transcripts. Dominance of PE B′9 transcripts at night was observed in the data on relative abundance, but *psaA* transcripts in general were low at night (**Figure 8B**). A similar offset was observed in B′-like and A-like nitrogen fixation gene transcription.

The later transcription of PE A1 photosynthesis genes may result from the limited penetration of light into the mat until between 1000 and 1100 h, when irradiance reached 5–100 μmol photons m^{-2} s^{-1} between 320 and 720 μm below the mat surface (**Figure 8C**), the depth where this population was predominantly located (**Figure 3**). Oxygen concentrations at these mat depths increased between 1000 and 1100 h reaching 3- to 6-times air
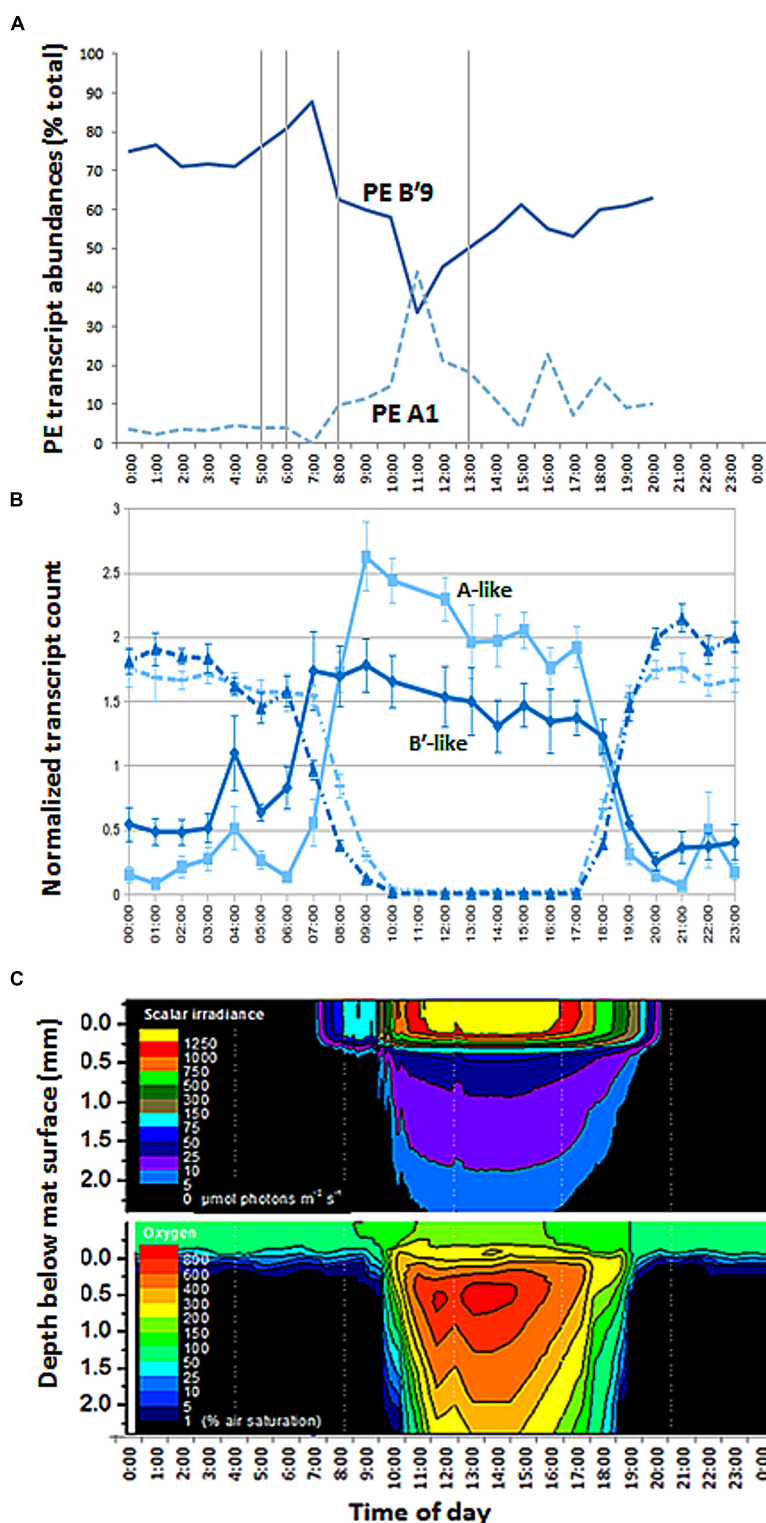
FIGURE 8 | Diel transcription and change in irradiance and oxygen concentration with depth in 60°C Mushroom Spring mat. (A) Relative abundances of transcripts of abundant PEs B'9 and A1 measured from *psaA* cDNA over the diel cycle. Samples corresponding to time points at 0500, 0600, 0800, 1300 (vertical lines), 2100, 2200, and 2300 are missing due to failed sequencing reactions. **(B)** Normalized count of B'-like (dark blue) and A-like (light blue) combined photosystem (solid line) and nitrogen fixation (dashed line) transcripts over the diel cycle. **(C)** Isopleth diagrams showing depth distribution of scalar irradiance (μmol photons m^{-2} s^{-1}) and O$_2$ concentration (% air saturation) over the diel cycle. Samples at time points 1700, 1800, 1900, and 2000 h in part A were from the previous evening, 11 September 2009.

saturation, and further accumulated to >8 times air saturation by 1300 h, which reflects increased photosynthetic activity at greater depth. In contrast, photon irradiance increased to between ∼500 and 1250 μmol photons m^{-2} s^{-1} at the mat surface between 1000 and 1100 h in parallel with a decline in PE B′9 transcript relative abundance.

Discussion

Next-generation sequencing technology allowed deep coverage of a broad array of environmental samples, and this has allowed us to demonstrate the ecological distinctions of most of the predominant sequence clusters predicted to be PEs by Ecotype Simulation. Importantly, the deep coverage provided access to sequence variation within PEs, allowing us to test whether the variants predicted by Ecotype Simulation to comprise a given PE are equivalently distributed along ecological gradients and behave equivalently in response to environmental change (i.e., they are ecologically interchangeable).

Pyrosequencing showed that PEs collected along the thermal gradient in the Mushroom Spring effluent channel in 2008 differed in their temperature ranges, as suggested by our previous studies of samples collected in 2006 and analyzed by cloning and sequencing (**Tables 1** and **2**; Supplementary Table S4). However, sequencing and CCA analyses in the present study further demonstrated that the variants predicted to belong to the same PE were distributed similarly, indicating ecological homogeneity within most PEs. PEs also responded differently to shifts in temperature, providing further evidence of divergence in their temperature preferences, and these shifts were consistent with temperature adaptations of cultivated *Synechococcus* strains (Miller and Castenholz, 2000; Allewalt et al., 2006). Furthermore, these experiments provided additional evidence that all of the members predicted to belong to the same PE respond to temperature in concert.

Pyrosequencing analyses also allowed us to resolve PEs that were distributed uniquely in the vertical aspect of the 60–63°C mat. CCA analyses demonstrated that all variants predicted to belong to the same PE were distributed similarly. These findings significantly extend earlier observations, in which vertical positioning of PEs was obscured either by the use of slowly evolving molecular markers, such as 16S rRNA and the 16S–23S rRNA internal transcribed spacer region (Ramsing et al., 2000; Ferris et al., 2003) or by the inability of DGGE to discern closely related sequence variants (Becraft et al., 2011). The light environment changes significantly over the top 1 mm of the microbial mat both with depth and over time (**Figure 8C**). The uppermost 0.1–0.2 mm experience high photon irradiance (>200–2000 μmol photons m^{-2} s^{-1}) for most of the sun-exposed part of the day, while deeper mat layers (>0.5 mm depths) experience much lower light levels (maximally ∼50 μmol photons m^{-2} s^{-1} during mid-day), and an altered light spectrum, due to the strong attenuation and scattering of visible wavelengths, and oxygenic photosynthesis in these mat layers thus remains light limited throughout most of the day (**Figure 8C**). PEs responded differently to reduced

irradiance, suggesting that adaptation to light may, at least in part, underlie these distribution patterns, and similar responses of all members of each PE again suggested ecological homogeneity of PE populations. Indeed, in other studies reported in the second paper of this series (Nowack et al., 2015, this issue), we have observed that *Synechococcus* strains representative of some of the PEs residing in deeper parts of the mat green layer have adaptations to low irradiance that confer the ability to acclimate to low irradiance values that are consistent with the light levels present at these depths.

Our observation that PE B′9 transcripts were expressed earlier than PE A1 transcripts matches the relatively longer and shorter light exposure periods experienced by these surface- and subsurface-associated populations, respectively (**Figure 8C**). Compared to deeper, subsurface populations (e.g. PE A1), the surface localization of PE B′9 might suggest that genes encoding its photosynthetic apparatus need to be transcribed earlier (i.e., before sunrise) to be available to harvest light energy in the morning. Interestingly, previous *in situ* microsensor studies showed that peak oxygenic photosynthesis moved downward to deeper portions of the 1 mm-thick upper green *Synechococcus* layer as light intensity increased during the morning (Ramsing et al., 2000). This might reflect the increased activity of deeper *Synechococcus* PEs, possibly combined with inhibition of photosynthesis in PEs nearer the mat surface when irradiance is highest.

While pyrosequencing provided an opportunity to examine the ecological interchangeability of individual variants within PEs, these results must be interpreted with caution. At issue is whether the resolution provided by a single gene, in this case *psaA*, can identify the most newly divergent ecotypes; also at issue is the appropriate coarseness of demarcation by Ecotype Simulation. A conservative demarcation of a PE might yield an ecologically heterogeneous set of variants, which could comprise multiple ecotypes instead of just one. It appears that conservative demarcation of several B′-like PEs yielded an ecologically heterogeneous set of organisms. For example, fine-scale demarcation of PE B′8 yielded PEs B′8-3 and B′8-4, which are significantly different in their temperature and depth distributions. In other cases, such as PE B′8-5, PEs cannot be resolved from others by their temperature and depth distributions, and so unmeasured parameters may explain their distinctions and their ability to co-exist. Comparative genomic analyses reported in the third paper of this series (see Olsen et al., 2015, this issue) are beginning to provide insights into such unsuspected adaptations.

In principle, complete ecological interchangeability is expected only in the populations that have descended from an ancestor that survived the most recent periodic selection event and have not subsequently split to form a new ecologically distinct population. The need to find the smallest clades with ecologically interchangeable membership emphasizes the importance of molecular resolution. For instance, even fine-scale PE demarcation yielded extremely closely related populations within PEs A6 and A12, which might prove to be the ecologically homogeneous ecotypes.

The deeper coverage of genetic variants and habitats provided by pyrosequencing analyses resulted in the demarcation by Ecotype Simulation of a greater number of non-singleton PEs than before. Whereas Becraft et al. (2011) used the conservative approach to detect 7 A-like and 12 B′-like PEs, here we detected 11 A-like 18 B′-like PEs. Although ∼500 times more sequences were included in the present study, all the high-frequency sequences found in the present study were detected previously in our smaller samples. This implies that the greater number of PEs demarcated resulted from the deeper coverage provided by pyrosequencing rather than changes in PE abundances in the mat between 2006 and 2008. Some PEs that were newly predicted in the present analyses corresponded to subclades previously embedded within PEs that had not been identified as PEs from more limited clone sequence data (e.g., PE A12; as shown in **Tables 1** and **2**), or to low-abundance populations not detected in clone libraries (e.g., PE A10). Many of these newly discovered PEs appear to be comprised of ecologically homogeneous members (see Supplementary Figure S4). Although cloning and sequencing studies analyzed only 0.2% of the amount of sequence data obtained in pyrosequencing study of samples collected in 2008, this approach still detected 66% of the currently demarcated PEs. In addition, newly demarcated PEs generally exhibited low relative abundance; this implies that the predominant *Synechococcus* populations contributing to the environment in these habitats have confidently been identified.

In summary, our analysis of sequence variation of an essential protein-encoding locus identified populations that meet the expectations of ecological species: (i) they are ecologically distinct populations and (ii) as shown by the similarity of their environmental distributions and their responses to environmental perturbations, the memberships of many predicted ecological species are ecologically interchangeable. These and other results (Melendrez et al., 2011) indicate that molecular divergence among ecological species is much less than among the species recognized by microbial systematists, including definitions of 1–3% 16S rRNA divergence (Wayne et al., 1987; Stackebrandt and Ebers, 2006) or >95% (Konstantinidis and Tiedje, 2005; Konstantinidis, 2011) or 95–96% (Chun and Rainey, 2014; Kim et al., 2014) average nucleotide identity (ANI). As will be demonstrated in the accompanying papers (Nowack

et al., 2015, this issue; Olsen et al., 2015, this issue), *psaA* was able to detect species that were identical or nearly identical (0.06% divergent) in their full-length 16S rRNA gene sequences and >98.35% identical in their average nucleotide indeces. In these and other habitats, where microbial diversity is organized by physico-chemical parameters, ecological species populations are likely to be the fundamental units of community composition and structure. If the level of molecular resolution needed to detect species in the communities we have studied is typical, the widespread use of high-throughput 16S rRNA sequencing methods may be masking the true species diversity and dynamics across microbial communities in a wide variety of habitats.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.00590

References

Allewalt, J. P., Bateson, M. M., Revsbech, N. P., Slack, K., and Ward, D. M. (2006). Effect of temperature and light on growth of and photosynthesis by Synechococcus isolates typical of those predominating in the octopus spring microbial mat community of Yellowstone National Park. *Appl. Environ. Microbiol.* 72, 544–550. doi: 10.1128/AEM.72.1.544-550.2006

Becraft, E. D., Cohan, F. M., Kuhl, M., Jensen, S. I., and Ward, D. M. (2011). Fine-scale distribution patterns of Synechococcus ecological diversity in microbial mats of Mushroom Spring, Yellowstone National Park. *Appl. Environ. Microbiol.* 77, 7689–7697. doi: 10.1128/AEM.05927-5911

Béjà, O., Spudich, E. N., Spudich, J. L., Leclerc, M., and Delong, E. F. (2001). Proteorhodopsin phototrophy in the ocean. *Nature* 411, 786–789. doi: 10.1038/35081051

Bhaya, D., Grossman, A. R., Steunou, A.-S., Khuri, N., Cohan, F. M., Hamamura, N., et al. (2007). Population level functional diversity in a microbial

community revealed by comparative genomic and metagenomic analyses. *ISME J.* 1, 703–713. doi: 10.1038/ismej.2007.46

Chun, J., and Rainey, F. A. (2014). Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *Int. J. Syst. Evol. Microbiol.* 64, 316–324. doi: 10.1099/ijs.0.054171-54170

Cohan, F. M. (1994). The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. *Am. Nat.* 143, 956–986. doi: 10.1086/285644

Cohan, F. M., and Perry, E. B. (2007). A systematics for discovering the fundamental units of bacterial diversity. *Curr. Biol.* 17, 373–386. doi: 10.1016/j.cub.2007.03.032

Connor, N., Sikorski, J., Rooney, A. P., Kopac, S., Koeppel, A. F., Burger, A., et al. (2010). Ecology of speciation in the genus Bacillus. *Appl. Environ. Microbiol.* 76, 1349–1358. doi: 10.1128/AEM.01988-09-1989

Denef, V. J., Kalnejais, L. H., Mueller, R. S., Wilmes, P., Baker, B. J., Thomas, B. C., et al. (2010). Proteogenomic basis for ecological divergence of closely related

bacteria in natural acidophilic microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2383–2390. doi: 10.1073/pnas.0907041107

Felsenstein, J. (1989). Mathematics vs. Evolution: Mathematical Evolutionary Theory. *Science* 246, 941–942. doi: 10.1126/science.246.4932.941

Ferris, M. J., Kuhl, M., Wieland, A., and Ward, D. M. (2003). Cyanobacterial ecotypes in different optical microenvironments of a 68OC hot spring mat community revealed by 16S-23S rRNA internal transcribed spacer region variation. *Appl. Environ. Microbiol.* 69, 2893–2898. doi: 10.1128/AEM.69.5.2893-2898.2003

Ferris, M. J., and Ward, D. M. (1997). Seasonal distributions of dominant 16S rRNA-defined populations in a hot spring microbial mat examined by denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* 63, 1375–1381.

Francisco, J. C., Cohan, F. M., and Krizanc, D. (2012). "Demarcation of bacterial ecotypes from DNA sequence data: a comparative analysis of four algorithms," in *Proceedings of the 2nd IEEE International Conference on Computational Advances in Bio and Medical Sciences (ICCABS)*. Washington, DC. doi: 10.1109/iccabs.2012.6182633

Hunt, D. E., David, L. A., Gevers, D., Preheim, S. P., Alm, E. J., and Polz, M. F. (2008). Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320, 1081–1085. doi: 10.1126/science.1157890

Jensen, S. I., Steunou, A.-S., Bhaya, D., Kuhl, M., and Grossman, A. R. (2011). In situ dynamics of O2, pH and cyanobacterial transcripts associated with CCM, photosynthesis and detoxification of ROS. *ISME J.* 5, 317–328. doi: 10.1038/ismej.2010.131

Jing, H., Lacap, D. C., Lau, C. Y., and Pointing, S. B. (2006). Community phylogenetic diversity of cyanobacterial mats associated with geothermal springs along a tropical intertidal gradient. *Extremophiles* 10, 159–163. doi: 10.1007/s00792-005-0477-9

Johnson, Z. I., Zinser, E. R., Coe, A., Mcnulty, N. P., Woodward, E. M. S., and Chisholm, S. W. (2006). Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* 311, 1737–1740. doi: 10.1126/science.1118052

Jordan, P., Fromme, P., Witt, H. T., Klukas, O., Saenger, W., and Krauss, N. (2001). Three-dimensional structure of cyanobacterial photosystem I at 2.5 A resolution. *Nature* 411, 909–917. doi: 10.1038/35082000

Kashtan, N., Roggensack, S. E., Rodrigue, S., Thompson, J. W., Biller, S. J., Coe, A., et al. (2014). Single-cell genomics reveals hundreds of coexisting subpopulations in wild *Prochlorococcus*. *Science* 344, 416–420. doi: 10.1126/science.1248575

Kim, M., Oh, H. S., Park, S. C., and Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 64, 346–351. doi: 10.1099/ijs.0.059774-0

Klatt, C. G., Wood, J. M., Rusch, D. B., Bateson, M. M., Hamamura, N., Heidelberg, J. F., et al. (2011). Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. *ISME J.* 5, 1262–1278. doi: 10.1038/ismej.2011.73

Koch, A. L. (1974). The pertinence of the periodic selection phenomenon to prokaryote evolution. *Genetics* 77, 127–142.

Koeppel, A., Perry, E. B., Sikorski, J., Krizanc, D., Warner, A., Ward, D. M., et al. (2008). Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2504–2509. doi: 10.1073/pnas.0712205105

Koeppel, A. F., Wertheim, J. O., Barone, L., Gentile, N., Krizanc, D., and Cohan, F. M. (2013). Speedy speciation in a bacterial microcosm: new species can arise as frequently as adaptations within a species. *ISME J.* 7, 1080–1091. doi: 10.1038/ismej.2013.3

Konstantinidis, K. T. (2011). "Metagenomic insights into bacterial species," in *Handbook of Molecular and Microbial Ecology*, Vol. 1, *Metagenomics and Complementary Approaches*, ed. F. J. D. Bruijn (Hoboken, NJ: John Wiley & Sons, Inc.), 89–98. doi: 10.1002/9781118010518.ch11

Konstantinidis, K. T., and Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2567–2572. doi: 10.1073/pnas.0409727102

Kopac, S., Wang, Z., Wiedenbeck, J., Sherry, J., Wu, M., and Cohan, F. M. (2014). Genomic heterogeneity and ecological speciation within one

subspecies of *Bacillus subtilis*. *Appl. Environ. Microbiol.* 80, 4842–4853. doi: 10.1128/AEM.00576-4

Kühl, M. (2005). Optical microsensors for analysis of microbial communities. *Methods Enzymol.* 397, 166–199. doi: 10.1016/S0076-6879(05)97010-9

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., Mcwilliam, H., et al. (2007). Clustal W and clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Lau, M. C., Aitchison, J. C., and Pointing, S. B. (2009). Bacterial community composition in thermophilic microbial mats from five hot springs in central Tibet. *Extremophiles* 13, 139–149. doi: 10.1007/s00792-008-0205-3

Legendre, P., and Legendre, L. (1998). *Numerical Ecology*. Amsterdam: Elsevier.

Liu, Z., Klatt, C. G., Ludwig, M., Rusch, D. B., Jensen, S. I., Kuhl, M., et al. (2012). Candidatus Thermochlorobacter aerophilum: an aerobic chlorophotoheterotrophic member of the phylum Chlorobi defined by metagenomics and metatranscriptomics. *ISME J.* 6, 1869–1882. doi: 10.1038/ismej.2012.24

Liu, Z., Klatt, C. G., Wood, J. M., Rusch, D. B., Ludwig, M., Wittekindt, N., et al. (2011). Metatranscriptomic analyses of chlorophototrophs of a hot-spring microbial mat. *ISME J.* 5, 1279–1290. doi: 10.1038/ismej.2011.37

Manning, S. D., Motiwala, A. S., Springman, A. C., Qi, W., Lacher, D. W., Ouellette, L. M., et al. (2008). Variation in virulence among clades of *Escherichia coli* O157:H7 associated with disease outbreaks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4868–4873. doi: 10.1073/pnas.0710834105

Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D., and Lefeuvre, P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 26, 2462–2463. doi: 10.1093/bioinformatics/btq467

Martiny, A. C., Tai, A. P. K., Veneziano, D., Primeau, F., and Chisholm, S. W. (2009). Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. *Environ. Microbiol.* 11, 823–832. doi: 10.1111/j.1462-2920.2008.01803.x

Melendrez, M. C., Lange, R. K., Cohan, F. M., and Ward, D. M. (2011). Influence of molecular resolution on sequence-based discovery of ecological diversity among *Synechococcus* populations in an alkaline siliceous hot spring microbial mat. *Appl. Environ. Microbiol.* 77, 1359–1367. doi: 10.1128/AEM.02032-0

Miller, S. R., and Castenholz, R. W. (2000). Evolution of thermotolerance in hot spring cyanobacteria of the genus *Synechococcus*. *Appl. Environ. Microbiol.* 66, 4222–4229. doi: 10.1128/AEM.66.10.4222-4229.2000

Miller, S. R., Strong, A. L., Jones, K. L., and Ungerer, M. C. (2009). Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Appl. Environ. Microbiol.* 75, 4565–4572. doi: 10.1128/AEM.02792-08

Nowack, S., Olsen, M. T., Schaible, G. A., Becraft, E. D., Shen, G., Klapper, I., et al. (2015). The molecular dimension of microbial species: 2. *Synechococcus* strains representative of putative ecotypes inhabiting different depths in the Mushroom Spring microbial mat exhibit different adaptive and acclimative responses to light. *Front. Microbiol.* 6:626. doi: 10.3389/fmicb.2015.00626

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., et al. (2013). *Vegan: Community Ecology Package, R Package Version 2.1-27/r2451*. Vienna: R Core Team.

Olsen, M. T., Nowack, S., Wood, J. M., Becraft, E. D., Labutti, K., Lipzen, A., et al. (2015). The molecular dimension of microbial species: 3. Comparative genomics of *Synechococcus* strains with different light responses and in situ diel transcription patterns of associated ecotypes in the Mushroom Spring microbial mat. *Front. Microbiol.* 6:604. doi: 10.3389/fmicb.2015.00604

Quince, C., Lanzen, A., Curtis, T. P., Davenport, R. J., Hall, N., Head, I. M., et al. (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat. Methods* 6, 639–641. doi: 10.1038/nmeth.1361

Quince, C., Lanzen, A., Davenport, R. J., and Turnbaugh, P. J. (2011). Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* 12:38. doi: 10.1186/1471-2105-12-38

Ramsing, N. B., Ferris, M. J., and Ward, D. M. (2000). Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter-thick photic zone of a hot spring cyanobacterial mat. *Appl. Environ. Microbiol.* 66, 1038–1049. doi: 10.1128/AEM.66.3.1038-1049.2000

Reeder, J., and Knight, R. (2009). The 'rare biosphere': a reality check. *Nat. Methods* 6, 636–637. doi: 10.1038/nmeth0909-636

Roberts, D. W. (2013). Labdsv: ordination and multivariate analysis for ecology. *R. Package Version* 1, 6–11.

Ruff-Roberts, A. L., Kuenen, J. G., and Ward, D. M. (1994). Distribution of cultivated and uncultivated cyanobacteria and Chloroflexus-like bacteria in hot spring microbial mats. *Appl. Environ. Microbiol.* 60, 697–704.

Shapiro, B. J., Friedman, J., Cordero, O. X., Preheim, S. P., Timberlake, S. C., Szabo, G., et al. (2012). Population genomics of early events in the ecological differentiation of bacteria. *Science* 336, 48–51. doi: 10.1126/science.1218198

Sikorski, J. (2008). Populations under microevolutionary scrutiny: what will we gain? *Arch. Microbiol.* 189, 1–5. doi: 10.1007/s00203-007-0294-x

Smith, N. H., Gordon, S. V., De La Rua-Domenech, R., Clifton-Hadley, R. S., and Hewinson, R. G. (2006). Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*. *Nat. Rev. Microbiol.* 4, 670–681. doi: 10.1038/nrmicro1472

Stackebrandt, E., and Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiology* 33, 152–155.

Ter Braak, C. J. F. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67, 1167–1179. doi: 10.2307/1938672

Vos, M. (2011). A species concept for bacteria based on adaptive divergence. *Trends Microbiol.* 19, 1–7. doi: 10.1016/j.tim.2010.10.003

Walk, S. T., Alm, E. W., Calhoun, L. M., Mladonicky, J. M., and Whittam, T. S. (2007). Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. *Environ. Microbiol.* 9, 2274–2288. doi: 10.1111/j.1462-2920.2007.01341.x

Ward, D. M. (1998). A natural species concept for prokaryotes. *Curr. Opin. Microbiol.* 1, 271–277. doi: 10.1016/S1369-5274(98)80029-5

Ward, D. M., Bateson, M. M., Ferris, M. J., Kuhl, M., Wieland, A., Koeppel, A., et al. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos. Trans. R. Soc. B Biol. Sci.* 361, 1997–2008. doi: 10.1098/rstb.2006.1919

Ward, D. M., and Cohan, F. M. (2005). "Microbial diversity in hot spring cyanobacterial mats: pattern and prediction," in *Geothermal Biology and Geochemistry in Yellowstone National Park*, eds W. B. Inskeep and T. R. Mcdermott (Bozeman, MT: Thermal Biology Institute), 185–201.

Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., et al. (1987). Report of the Ad-Hoc-committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol.* 37, 463–464. doi: 10.1099/00207713-37-4-3

West, N. J., and Scanlan, D. J. (1999). Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic Ocean. *Appl. Environ. Microbiol.* 65, 2585–2591.

The molecular dimension of microbial species: 2. *Synechococcus* strains representative of putative ecotypes inhabiting different depths in the Mushroom Spring microbial mat exhibit different adaptive and acclimative responses to light

Shane Nowack[1,2]*, Millie T. Olsen[3], George A. Schaible[3], Eric D. Becraft[3], Gaozhong Shen[4], Isaac Klapper[1,5], Donald A. Bryant[4,6] and David M. Ward[3]

[1] Department of Mathematical Sciences, Montana State University, Bozeman, MT, USA, [2] School of Environmental Sciences, University of Guelph, Guelph, ON, Canada, [3] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA, [4] Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA, [5] Department of Mathematics, Temple University, Philadelphia, PA, USA, [6] Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA

Closely related strains of thermophilic *Synechococcus* were cultivated from the microbial mats found in the effluent channels of Mushroom Spring, Yellowstone National Park (YNP). These strains have identical or nearly identical 16S rRNA sequences but are representative of separate, predicted putative ecotype (PE) populations, which were identified by using the more highly resolving *psaA* locus and which predominate at different vertical positions within the 1-mm-thick upper-green layer of the mat. Pyrosequencing confirmed that each strain contained a single, predominant *psaA* genotype. Strains differed in growth rate as a function of irradiance. A strain with a *psaA* genotype corresponding to a predicted PE that predominates near the mat surface grew fastest at high irradiances, whereas strains with *psaA* genotypes representative of predominant subsurface populations grew faster at low irradiance and exhibited greater sensitivity to abrupt shifts to high light. The high-light-adapted and low-light-adapted strains also exhibited differences in pigment content and the composition of the photosynthetic apparatus (photosystem ratio) when grown under different light intensities. Cells representative of the different strains had similar morphologies under low-light conditions, but under high-light conditions, cells of low-light-adapted strains became elongated and formed short chains of cells. Collectively, the results presented here are consistent with the hypothesis that closely related, but distinct, ecological species of *Synechococcus* occupy different light niches in the Mushroom Spring microbial mat and acclimate differently to changing light environments.

Keywords: microbial species, light adaptation, light acclimation, cyanobacteria, photosynthesis

Introduction

As reviewed by the first paper in this three-paper series on the molecular dimension of microbial species (Becraft et al., 2015), in the course of a long-term effort to understand the composition, structure, and function of microbial mat communities inhabiting alkaline, siliceous hot springs in Yellowstone National Park (YNP), we have focused on sequence variation in increasingly more divergent genes, progressing from 16S rRNA (Ferris et al., 1997) to the 16S–23S rRNA internal transcribed spacer region (Ferris et al., 2003) to protein-encoding genes (Becraft et al., 2011; Melendrez et al., 2011). Due to different cyanobacterial gene sequence variants being predominant at different locations along thermal gradients, we hypothesized the existence of temperature-adapted *Synechococcus* ecotypes, which was later demonstrated by obtaining representative strains and studying their temperature preferences (Allewalt et al., 2006). Similar temperature adaptations were reported for *Synechococcus* strains cultivated from Oregon hot springs by Peary and Castenholz (1964) and Miller and Castenholz (2000).

Differences in the distribution of 16S rRNA and 16S–23S internal-transcribed-spacer sequence variants along vertical profiles in the upper 1 mm-thick photic zone of these mats (Ramsing et al., 2000; Ferris et al., 2003) led us to hypothesize the existence of different light-adapted *Synechococcus* ecotypes. Microsensor studies have shown that the dense populations of mat inhabitants alter light quantity and wavelength distribution dramatically with depth in the upper 1–2 mm of the mat (see Figure 4 in Becraft et al., 2015), providing selection conditions for evolutionary adaptations along these light gradients and to other environmental parameters that vary with depth. Additionally, microsensor analyses revealed that oxygenic photosynthesis in a mat recovering from physical disturbance exhibited two maxima, one nearer, and another farther from the mat surface (Ferris et al., 1997), providing further evidence in support of the existence of *Synechococcus* ecotypes adapted to different light microenvironments. Becraft et al. (2015) combined pyrosequencing analysis of the gene encoding *psaA*, a core subunit of the photosystem I reaction center, with analysis of the *psaA* sequences by Ecotype Simulation, an algorithm based on the Stable Ecotype Model of species and speciation that predicts ecological species populations from sequence variation (Koeppel et al., 2008). These hypothetical species are called putative ecotypes (PEs) until they are shown to exhibit properties expected of ecological species (Becraft et al., 2015). This analysis permitted prediction of *Synechococcus* PEs and provided a conceptual basis for ensuing studies of their vertical distributions in the microbial mats. By examining 80 μm-thick vertical sections of mat samples collected at 60–63°C, which were obtained by cryotome sectioning, a progression from the mat surface downward of *Synechococcus* PEs B′9, A1, A4, A14, and A6 was observed (**Table 1**; see also Figures 3 and 4 in Becraft et al., 2015). The predicted A-like PEs exhibit identical or nearly identical 16S rRNA sequences (see Olsen et al., 2015).

The existence of ecotypes of another hot spring cyanobacterium, *Plectonema notatum*, that are evolutionarily adapted to different irradiance levels, was demonstrated by Sheridan (1976, 1978). These findings, however, contrast with previous studies of light responses of native *Synechococcus* populations in such mats (Brock and Brock, 1969; Madigan and Brock, 1977), which had been interpreted as acclimative changes of a single *Synechococcus* population that was physiologically adjusting to a change in the environment (M.T. Madigan, personal communication). [Note: We will use the term acclimation to mean the physiological response of an organism to an environmental change; we will use adaptation to mean an alteration in the structure or function of an organism or any of its parts that results *from natural selection* and by which the organism becomes better fitted to survive and multiply in its environment.] If evolutionarily adapted *Synechococcus* ecotypes exist, changes in the relative abundances of differently adapted *Synechococcus* ecotypes, such as those observed in light alteration experiments by Becraft et al. (2015), would provide an alternate explanation of the responses observed in earlier studies.

Previous studies of evolutionary adaptation to light were performed on *Synechococcus* strains obtained from low-dilution enrichments with 16S rRNA sequences representative of predominant natural populations (Allewalt et al., 2006) or substrains therefrom (Kilian et al., 2007). As such, these strains might contain multiple ecotypes with the same 16S rRNA sequence, which were shown to exist by higher-molecular resolution, theory-based analyses (Becraft et al., 2011, 2015). Here we report on *Synechococcus* strains obtained from more highly diluted mat samples, which are representative of PEs designated from *psaA* sequence variation. These *Synechococcus* cultures are not axenic because they presently contain heterotrophic contaminants, but we refer to them as strains since we will show that they represent individual genotypes of distinct *Synechococcus* PEs (i.e., with respect to the cultivated *Synechococcus*, they are strains of a known species). We will demonstrate that strains representative of PEs that are known to predominate at different positions along the vertical light gradient in the mat have different evolutionary adaptations and acclimative responses to light, even though they have identical or nearly identical 16S rRNA sequences. The genomic sequences of these strains are described in the third paper of this series (Olsen et al., 2015).

Materials and Methods

Sample Collection

Samples for cultivation were collected from Mushroom Spring, YNP, at sites with temperatures of 60, 63, and 65°C on September 7, 2010 using a #4 cork borer (8 mm diameter). The top green layer, ∼1 mm in thickness, of each mat sample was removed with a razor blade, placed in a 1.5-mL microcentrifuge tube, and then returned to the lab in a thermos containing Mushroom Spring source water (cooled to the temperature at which the sample was collected). The time from sampling to the lab was ∼2 h, and the temperature in the thermos was ∼7 to 10°C degrees

TABLE 1 | Summary of *psaA* Ti454-barcode sequencing analyses of *Synechococcus* strains.

Strain name	Temperature[a]	Depth[b]	Total sequences	Dominant variant sequences[c]			Other sequences[e]	
				Putative ecotype (PE)[d]	DVs	%	Total	%
JA-3-3Ab	58–65°C	400–600	1116	All	931	83.4		
				A1	930	83.3	185	16.6
				A14	1	0.1		
JA-2-3B′a (2–13)	51–61°C	N/A	2054	All	1690	82.3		
				B′19	1385	67.4	364	17.8
				B′24	300	14.6		
				B′11	5	0.2		
CIW-10	51–61°C	N/A	1528	All	1077	70.5		
				B′24	438	28.7	413	27
				€	639	41.8	38	2.5
65AY6Li	65°C	400–600	980	**A1**	780	79.6	200	20.4
65AY6A5	65°C	640–720	1325	All	1106	83.5	219	16.5
				A4	1105	83.4		
				A1	1	0.1		
60AY4M2	60°C	800	920	**A14**	756	82.2	164	17.8

[a]*Temperature of the sample from which the strain was obtained.*

[b]*Depth from surface where PE is most abundant at 60°C in Mushroom Spring microbial mat. Units are μm. N/A means PE was not a dominant population in the 60°C mat at any depth. Data reported in Becraft et al. (2015).*

[c]*Number of sequences that are identical to the dominant variant (DV; identical sequences that made up the plurality of a putative ecotype PE clade) of the PE listed in the same row.*

[d]*PEs in bold are the dominant PEs in the culture.*

[e]*Other sequences that are closely related to the DV sequence.*

[€]*A dominant sequence variant in the culture that is distinct from PE B′24 by 17 single-nucleotide polymorphisms.*

cooler upon arrival at the lab than that measured at the field site.

Microscope Counts

To obtain an estimate of the number of *Synechococcus* cells found in the top green layer of a #4 mat core, the green layer of a 63°C mat sample was homogenized in 10 mL of autoclaved Mushroom Spring water. The cells in a 10-μL subsample were counted using a Bright-line hemacytometer (Hausser Scientific, Horsham, PA, USA) and microscope (Zeiss Axioskop 2 plus with an HBO 100 UV lamp) at 40× magnification. *Synechococcus* cells were identified by their orange–red autofluorescence.

Synechococcus Isolation

Using Castenholz's (1969) medium D supplemented with HEPES buffer (DH), Allewalt et al. (2006) were only successful in cultivating *Synechococcus* strains out to 10^5-fold dilution of mat samples that contained ~10^8 to 10^9 *Synechococcus* cells/mL. For this study, we developed a modified medium containing sodium acetate and yeast extract (medium DHAY), each added at 0.01% (w/v), which resulted in the recovery of strains of *Synechococcus* from more highly diluted mat samples and faster growth rates in liquid cultures. The medium was solidified with 1.5% (w/v) Gelrite™ (Sigma–Aldrich, St. Louis, MO, USA) throughout the isolation process.

Mat samples containing *Synechococcus* cells were diluted 10-fold to extinction prior to inoculation into autoclaved media that had been cooled to 50°C, which was then poured into sterile 60 mm-diameter Petri dishes. After allowing the Gelrite

to solidify, the plates were placed in sealed Ziploc bags with wetted paper towels and were then incubated at 52°C under 50 μmol photons $m^{-2}sec^{-1}$ of white fluorescent light. Plates were inspected daily over a one-month period to monitor the growth of *Synechococcus* colonies. Isolated colonies formed by *Synechococcus* that grew on plates inoculated with mat samples that were diluted at least 10^5-fold from the original mat sample were picked with a sterile toothpick. In order to focus efforts on strains representative of different PEs, we screened strains by *psaA* sequence analysis before attempting further purification. Thus, to obtain sufficient material for sequencing, colonies were suspended in 2 mL of liquid medium DHAY in glass test tubes with caps that were loosely fastened to allow gas exchange. The strains were grown to stationary phase and were then scaled up in a step-wise manner to 20 mL and then to 100 mL by the addition of DHAY medium. Incubation conditions were as described above and remained constant throughout this process.

Molecular Assessment of *psaA* Genotype, Strain Purification, and Purity
DNA Extraction and PCR Amplification

Cells from a 1.5-mL aliquot of liquid culture were pelleted by centrifugation for 3 min at 4800 × *g* and resuspended in 200 μL of lysis buffer. Cells were lysed using the FastPrep Cell Disrupter (Bio101 Savant Instruments, Holbrook, NY, USA) and DNA was extracted and purified using the FastDNA Spin kit (Molecular Biosciences, Boulder, CO, USA) by following the manufacturer's instructions. Amplicons of *psaA* for DNA sequencing were

produced by polymerase chain reaction (PCR) amplification as described by Becraft et al. (2011).

Sanger Sequencing and Ecotype Assignment

Sanger sequencing of *psaA* amplicons was performed at the Idaho State University Molecular Biosciences Core Facility. *Synechococcus* strain sequences were aligned to sequences obtained from the mat, which were used to demarcate PEs (Becraft et al., 2015). Batch cultures producing amplicons with no ambiguous base calls and low background signals, and with sequences that were identical to a dominant variant (DV) representative of a predominant PE, were further purified.

Purification of Strains

Strains representative of different *psaA*-based PEs were diluted to extinction again in liquid medium DHAY to increase the probability that strains were representative of a single PE. The resulting highest-dilution subculture was prepared for pyrosequencing (see methods below). Strains with sequences containing ambiguous base calls and/or high background signal were re-diluted to extinction on plates, and the cultivation process was repeated by picking well-isolated colonies from high-dilution plates.

Pyrosequencing for Assessing Culture Purity

Strains meeting the purity criteria described above were grown to late exponential growth phase (between 2 and 5 $\times$ 10^7 cells/mL), and the cells from 60-mL liquid cultures were pelleted by centrifugation for 30 min at 1000 $\times$ *g*. Cell pellets were flash-frozen in liquid nitrogen and stored at $-80°C$. In order to have DNA of sufficient quality for both assessing culture purity and genome sequencing (see Olsen et al., 2015), an enzymatic cell lysis protocol was followed. Each frozen pellet was thawed and resuspended in medium DHAY to produce a 1-mL total volume; DNA was extracted by a phenol/chloroform/isoamyl alcohol method following cell lysis by lysozyme/proteinase K (see complete DNA extraction protocol in Supplementary information). RNA and other impurities were removed using the RNase I treatment protocol described by the Department of Energy Joint Genome Institute[1]. The DNA concentration was quantified using a NanoDrop Spectrophotometer ND-1000 (NanoDrop Technologies, Wilmington, DE, USA), and PCR was performed as described above to ensure that the DNA could be amplified.

As was the case in Allewalt et al. (2006), cyanobacterial strains contained heterotrophic contaminants, which we could not eliminate with repeated streaking for isolation or using spent media from the mixed strains or pure-cultures of heterotrophic isolates. To characterize the composition of the strains, pyrosequencing of 16S rRNA amplicons produced with primers 28F and 519R was performed according to the methodology posted on The Research and Testing Laboratory website[2]. Pyrosequencing of *psaA* gene amplicons was performed as described in Becraft et al. (2015) to assess the complexity of the cultivated *Synechococcus* population. Systematic errors,

---

[1]http://my.jgi.doe.gov/general/protocols.html
[2]www.researchandtesting.com

defined as single-nucleotide polymorphisms that were common to different strains, were removed. A culture was considered to contain a single predominant cyanobacterial ecotype if it contained a single dominant *psaA* sequence, as well as closely related, less-abundant genetic variants with 1 or 2 randomly distributed nucleotide substitutions compared to the dominant sequence. The genetic variants might have arisen during cultivation or represent sequencing error (Gilles et al., 2011), which might have occurred because a non-high-fidelity polymerase (Qiagen HotStar Taq polymerase) was used. These sequences are available upon request.

Final Strain Purification and Purity Check

To achieve an extra level of confidence regarding purity with respect to cyanobacterial PEs, all strains were again diluted to extinction in liquid medium DHAY and the subculture from the highest dilution was subjected to a second round of *psaA* amplicon sequencing using pyrosequencing. This sequence-based criterion was also used to test the purity of the two *Synechococcus* strains (JA-3-3Ab and JA-2-3B′a (2–13)) that were originally cultivated by Allewalt et al. (2006), and whose genomes were sequenced (Bhaya et al., 2007), and a third *Synechococcus* strain (CIW-10), which was derived from strain JA-2-3B′a (2–13) by Kilian et al. (2007; see Results).

Growth at Different Irradiances

All growth experiments were performed with medium DHAY using an illuminated growth chamber that consisted of an aquarium of dimensions 152.4 cm $\times$ 15.2 cm $\times$ 38.1 cm that served as a water bath. The temperature (52 or 60°C) was controlled with a PolyScience circulator, Series 7000 (Niles, IL, USA). The chamber was illuminated with two identical ATI 6 $\times$ 80 W SunPower T5 high-output fluorescent fixtures (Denver, CO, USA), one on each side of the aquarium, using a total of 12 150-cm-long fluorescent tubes of type F80W-T5-841-ECO (General Electric, Fairfield, CN, USA). Different light conditions, spanning the range of irradiances observed in nature, were achieved by applying various layers of neutral-density filter covering (GAM products, Los Angeles, CA, USA) around the culture tubes. Before sterilizing the culture tubes, irradiance was measured using a scalar irradiance probe, model QSL2100 from Biospherical Instruments (San Diego, CA, USA), by submersing the probe in 100 mL of water. To reduce the chance of CO_2 limitation, a sterilized, cotton-plugged 7 mm glass tube was connected to a gas cylinder containing 6% CO_2 in air (GENDCO, Bozeman, MT, USA) and was used to sparge the cultures at approximately one bubble per second. To prevent possible acidification of the medium, 26 mM $NaHCO_3$ was added to the growth medium (before autoclaving). The optimal concentration of $NaHCO_3$ was determined in a preliminary experiment by sparging ($\sim$1 bubble/second) liquid medium DHAY with 6% v/v CO_2 and identifying the minimum concentration of $NaHCO_3$ required to stabilize the pH over a period of 3 days.

Prior to inoculation, liquid cultures were pre-grown to late exponential phase (to minimize lag phase), or to a density of $\sim$2 $\times$ 10^7 cells/mL, at either (i) 52°C and at a scalar irradiance of

50 μmol photons m^{-2}sec^{-1}, or (ii) 60°C and 500 μmol photons m^{-2}sec^{-1} of white fluorescent light under a supply of 6% CO_2 in air, as described above. Duplicate cultures for each light intensity were inoculated by adding either 5.5×10^5 cells/mL or 1×10^6 cells/mL to 100 mL of medium DHAY in 175-mL glass P/T culture tubes (Bellco, Vineland, NJ, USA). The tubes were capped with silicon sponge closures (Sigma–Aldrich, St. Louis, MO, USA) to allow for gas exchange. Light intensity and pH were measured before and after each experiment to ensure that these parameters had remained stable during the experiment. Samples (1.0 mL) were taken every 12 h, fixed with glutaraldehyde (0.125% final concentration), and frozen at −80°C. Cell counts were obtained using flow cytometry (BD-FACSCanto or BD-FACSAria II flow cytometer and BD counting beads (BD Biosciences, San Jose, CA, USA) or CountBright Absolute counting beads (Invitrogen, Life Technologies, Waltham, MA, USA)). Each sample was filtered through a 70-μm screen-cap filter (Fisher Scientific) before analysis on the cell counter to avoid clogging of the flow cell. Growth rates were determined by estimating log-linear slopes during exponential growth phase.

Synechococcus cells are typically 8–10 μm in length, and microscopic examination revealed that they were noticeably longer than any of the heterotrophic cells that were detected in the culture and that were also able to pass through the screen-cap filter (see **Figure 1**). *Synechococcus* cells but not heterotrophic contaminants contain chlorophyll, Chl *a*, a pigment that is excited by the SYTO 17-A laser on the FACSAria II flow cytometer. Therefore, plots of forward scatter (a measure of cell size) versus autofluorescence excited by the SYTO 17-A laser equipped with a 650–670 nm emission filter, and plots of forward scatter versus side-scatter (cell complexity) were analyzed to distinguish heterotrophic cells from the *Synechococcus* cells (see Supplementary Figure S1).

Post-Experiment Validation of *psaA* Ecotype

When cultures reached exponential growth phase (cell density between 2 and 5×10^7 cells/mL), cells were harvested by centrifugation for 30 min at 1000 × *g*. The cell pellets were frozen in liquid nitrogen and stored in a −80°C freezer. DNA was extracted and a segment of *psaA* was amplified by PCR and sequenced, as described above, to verify that the *psaA* genotype had not changed during the experiment due to the selection of a rare *Synechococcus* variant with a different light adaptation. No evidence of changes in genotypes were found in this study.

Cell Morphology

To observe the cell morphologies of the heterotrophic contaminants and to identify any possible morphological differences among the *Synechococcus* strains, differential interference contrast and fluorescence microscopy were performed using a Nikon Eclipse 80i microscope with a Nikon Intensilight C-HGFl UV lamp and a Nikon DS-Ri1 camera. Near the end of exponential growth phase (cell density of 2 to 5×10^7 cells/mL) for the 52°C pre-growth experiment, photomicrographs of cells grown under low light (25 μmol photons m^{-2}sec^{-1}) or high light (600 μmol photons m^{-2}sec^{-1}) were obtained (see Supplementary Figure S2). Composite images were then produced by combining cells (at an equal proportion) that were grown under the two light conditions into a single sample. The purpose of obtaining these images was to seek evidence of morphological acclimation to light. These same samples were also analyzed on the BD-FACSCanto flow cytometer. Cell size (forward scatter) and autofluorescence intensity were measured and compared to the microscopic observations.

Measurement of Pigments

Total pigments were extracted from cells grown under different light conditions with 100% methanol to determine the Chl *a* and carotenoid contents. Spectroscopic measurements were carried out with a Genesys^{TM} 10 UV/Vis scanning spectrophotometer (Thermo Scientific, Rochester, NY, USA). The concentrations of Chl *a* and carotenoids were determined on the basis of equivalent cell concentrations as determined by equal OD$_{730\ nm}$ values as described in Shen et al. (2002). Phycobiliprotein (PBP) contents were determined by comparing the absorbance difference between untreated cells and cells that were incubated at
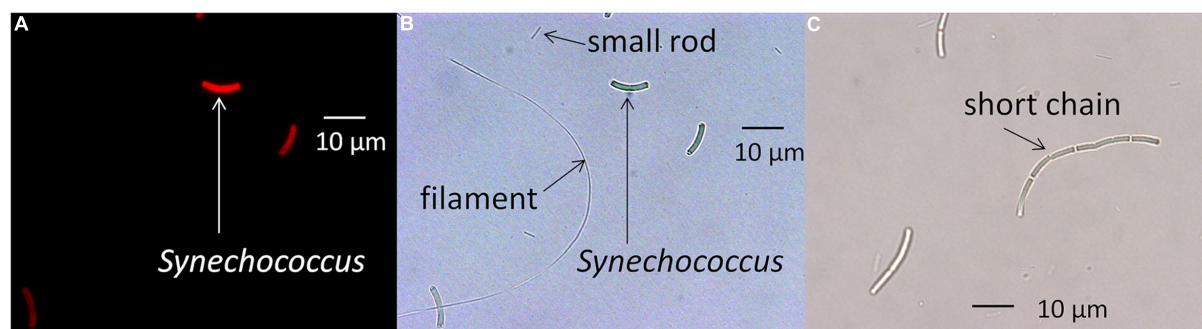
FIGURE 1 | Microscopic images of a *Synechococcus* strain and heterotrophic contaminants. (A) Fluorescence microscopy photomicrograph of the PE A1 strain (65AY6Li) grown at a scalar irradiance of 50 μmol photons m^{-2}sec^{-1}. **(B)** Same image using differential interference contrast microscopy, showing rod-shaped and filamentous heterotrophic contaminants. **(C)** Differential interference contrast photomicrograph showing elongated cells of the PE A14 strain grown at a scalar irradiance of 600 μmol photons m^{-2}sec^{-1} after pre-growth at 50 μmol photons m^{-2}sec^{-1}. Scale bars are 10 μm.

75°C for 5 min. The PBP content was estimated by the previously described procedure (Zhao et al., 2001).

Low-Temperature Fluorescence Emission Spectroscopy

Low-temperature (77 K) fluorescence emission spectra were measured for cells grown at different light conditions using an SLM8000-based spectrofluorometer modified for computerized, solid-state operation by On-Line Instrument Systems, Inc. (Bogart, GA, USA) as described previously (Shen et al., 2008). After determination of the cell density at 730 nm, cells were adjusted to $OD_{730 nm} = 0.5$ in 50 mM HEPES, pH = 7 containing 60% v/v glycerol. After loading into measuring tubes, samples were incubated for 5 min in the dark and then quickly frozen in liquid nitrogen. The excitation wavelength was set at 440 nm to excite proteins containing Chl *a* selectively or to 590 nm to excite PBP selectively.

Results

Synechococcus Isolation from Mat Samples

Direct microscopic counts of mat samples from 63°C revealed that there were ~1.8×10^8 *Synechococcus* cells within the 1 mm-thick green layer sampled with a #4 cork borer, which corresponded to an approximate volume of 0.05 mL (hence ~3.6×10^9 cells/mL). After homogenization and plating on DHAY medium, growth of *Synechococcus* colonies was observed as early as day three on low-dilution plates, and as late as day 21 on high-dilution plates. The plated dilutions resulted in growth of *Synechococcus* colonies out to the 10^7-fold dilution for the 60°C sample, and out to the 10^8-fold dilution for the 63 and 65°C samples. The cell counts reported here were nearly identical to those reported by Brock (1978), who found the cell densities to be constant at several temperatures between 57.2 and 68°C in the Mushroom Spring mat. Thus, if this is generally true, we estimate that less than 1% of the inoculated cells from 60°C, ~5 to 20% from 63°C, and ~50 to 100% from 65°C formed colonies on the plated dilution series. The estimated numbers of cells in the inocula of the final dilutions demonstrating growth on plates were ~300 from the 60°C sample, ~30 from 63°C sample, and ~3 from the 65°C sample. Additional extincting dilutions resulted in improvements in the efficiency of clonal isolation for all strains; specifically, ~10 to 100% of the estimated cells in the inoculum formed colonies on plates.

Hundreds of well-isolated colonies containing cells with the typical unicellular morphology of *Synechococcus* were picked from high-dilution plates, suspended in 2 mL of liquid medium DHAY, and then gradually scaled up to larger volumes in preparation for PE classification and physiological characterization.

Molecular and Morphological Descriptions of Strains

As shown in **Figure 1A**, microscopic observation revealed that strains contained rod-shaped, red-autofluorescing *Synechococcus* cells that were ~8–10 μm in length. The strains also contained smaller rod-shaped cells approximately 2–5 μm in length and filamentous cells greater than 50 μm in length (**Figure 1B**). 16S rRNA pyrosequencing analyses (Supplementary Table S1) suggested that *Meiothermus* sp. and *Caldilinea aerophila*-like sequences were present in these strains, and the presence of rod-shaped and filamentous cells is consistent with morphologies of cultivated strains of *Meiothermus* sp. (Nobre and Da Costa, 2001) and *C. aerophila* (Sekiguchi et al., 2003), respectively.

Sanger and pyrosequencing revealed strains that were dominated by *psaA* sequences representative of predominant A-like PEs that are differently distributed in the vertical profile at 60 to 63°C (PEs A1, A4, and A14). These strains were assigned the names 65AY6Li, 65AY6A5, and 60AY4M2, respectively, but for simplicity they will be referred to by the *psaA*-based PEs that they represent, i.e., PEs A1, A4, and A14 (**Table 1**). Pyrosequencing analyses also revealed that *Synechococcus* strain JA-3-3Ab previously cultivated by Allewalt et al. (2006) was heavily dominated by the expected DV of PE A1 and associated random singleton variants; a single sequence corresponding to PE A14 was detected (representing less than 0.1% of the variants in the culture). In contrast, previously cultivated *Synechococcus* strain JA-2-3B′a (2–13) contained multiple high-frequency sequences representative of three different *Synechococcus* PEs (B′19, B′24, and B′11), two of which were present in substantial amounts (see **Table 1**). Similar observations were made regarding the presence of multiple *Synechococcus* PEs in strain CIW-10 (Kilian et al., 2007). In fact, the pyrosequencing data revealed that strain CIW-10 was dominated by PE B′24, which was detected in the parent culture (JA-2-3B′a (2–13)) but was not the dominant PE in that culture. A previously unidentified genotype (€ in **Table 1**) apparently arose during cultivation as the predominant variant detected in this culture.

Adaptive and Acclimative Light Behavior
Effect of Heterotrophic Growth on *Synechococcus* Light Responses

A preliminary experiment was conducted to investigate the possibility that the measured light responses of the *Synechococcus* strains might be secondary to effects of light on the heterotrophic contaminants. By taking advantage of the unique fluorescence and light-scattering characteristics of the cells in these populations, the growth of both was tracked simultaneously. The results for the PE A1 strain are shown in **Figure 2**. During the first 12 h, the growth rate of the heterotrophic population (black) was nearly identical at all light intensities (~12–13 doublings/day). During this time the *Synechococcus* population (blue) remained in lag phase. After 24 h the heterotrophic population began to decrease and the *Synechococcus* population began to increase, eventually overtaking the culture. As expected for phototrophic organisms, the growth rate of the *Synechococcus* cells varied with light intensity, whereas the growth rate of the heterotrophic cells did not. Thus, we hypothesize that the heterotrophs pre-condition the medium and that any possible impact that the heterotroph metabolic activity has on cyanobacterial response to light would be constant over the different experiments (see **Figure 2**).
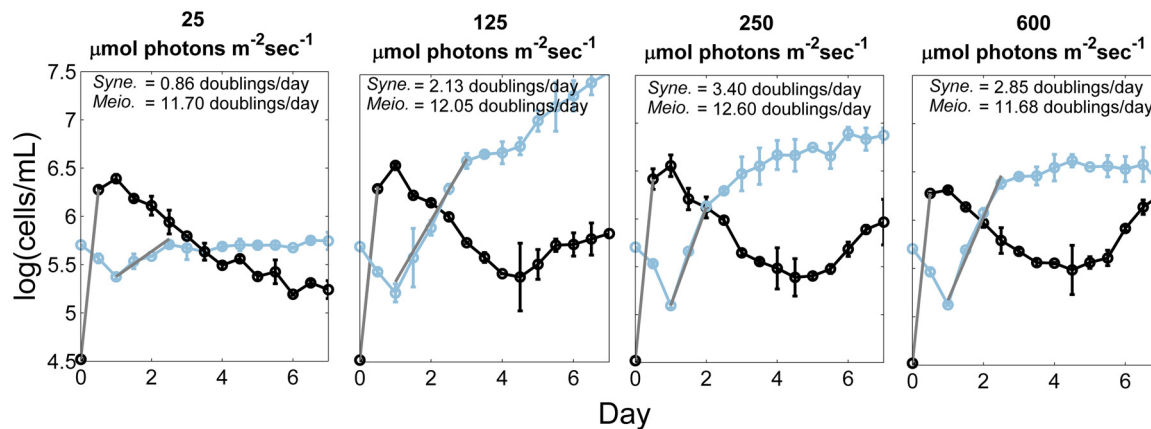
FIGURE 2 | Growth of heterotrophic and *Synechococcus* cells at different irradiances. Number of cells of the heterotrophic population (black) and the PE A1 *Synechococcus* strain (65AY6Li; blue) at four irradiance values (25, 125, 250, and 600 μmol photons m^{-2}sec^{-1}), over a 7-day period at 52°C, and without any additional dissolved inorganic carbon provided to medium DHAY. The culture used as inoculum was grown at 50 μmol photons m^{-2}sec^{-1}. The gray lines are the least-square best fit lines that were used to calculate the growth rate during exponential growth phase. Each point is the average of two measurements and the range is shown by the bars.

Growth Behavior of *Synechococcus* Strains in Response to Irradiance

The adaptations to irradiance of the three *Synechococcus* strains grown at 60°C, when cells were pre-grown at 50 μmol photons m^{-2}sec^{-1} without CO_2 sparging, are shown in **Figure 3**. At the lowest irradiance (25 μmol photons m^{-2}sec^{-1}), the PE A4 and PE A14 strains grew faster than the PE A1 strain ($p < 0.05$ in a two-factor ANOVA analysis (**Figure 3A**)). The PE A14 strain grew faster than the other strains at irradiances of up to ~200 μmol photons m^{-2}sec^{-1}, and the PE A1 strain grew faster than the other strains at irradiances greater than ~250 μmol photons m^{-2}sec^{-1}. The highest scalar irradiance supporting growth for the PE A1 strain (**Figure 3B**, blue) was at least 2900 μmol photons m^{-2}sec^{-1} at 60°C, the maximum

intensity that could be achieved in the lab. The PE A4 strain (**Figure 3B**, orange) was able to grow at 2200 μmol photons m^{-2}sec^{-1}, but not at 2500 μmol photons m^{-2}sec^{-1}. The PE A14 strain (**Figure 3B**, red) had the lowest light tolerance of the three strains tested and was able to grow at 1050 μmol photons m^{-2}sec^{-1}, but growth was not observed at 1550 μmol photons m^{-2}sec^{-1}.

The differences in the observed upper-light limits of the three strains might be due to cells of different PEs having a higher or lower survival probability when shifted from very low light, 52°C and CO_2-limiting conditions to very high light, 60°C, and CO_2-replete conditions. When temperature- and CO_2-shifts were eliminated and the light-shift was reduced by pre-growing at a light intensity of 500 μmol photons m^{-2}sec^{-1} the PE A14 strain

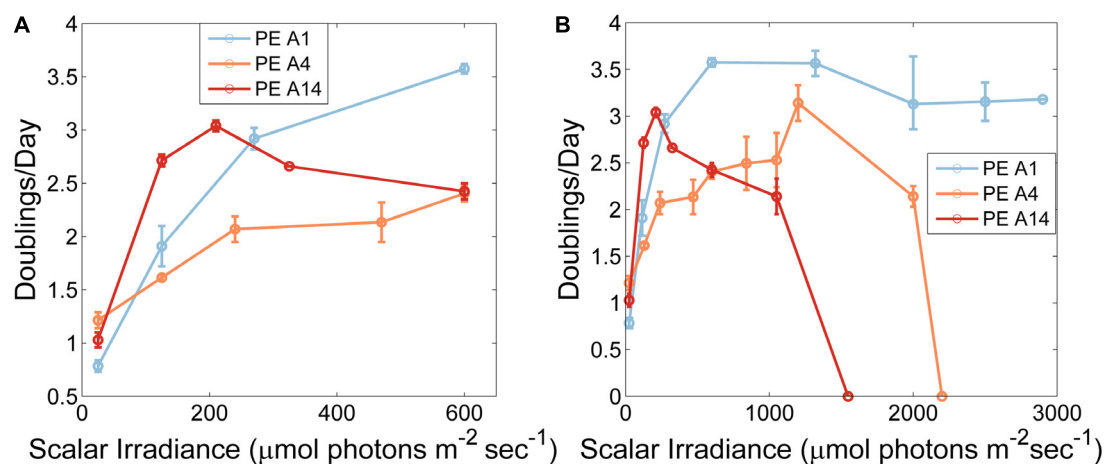**FIGURE 3 | Growth rates of *Synechococcus* strains representative of predominant PEs as a function of irradiance when grown at 60°C and sparged with 6% (v/v) CO_2 in air.** Strains used in this experiment were pre-grown at 50 μmol photons m^{-2}sec^{-1}, 52°C and without CO_2 sparging **(A)** Low irradiances only. **(B)** All tested irradiances. Each point is the average of two measurements and the range is shown by the bars.

was able to grow at light intensities up to at least 2250 μmol photons m^{-2}sec^{-1} (**Figure 4**). At light intensities $\leq$325 μmol photons m^{-2}sec^{-1}, cells pre-grown at the higher irradiance value grew after a longer lag phase than cells pre-grown at the lower light intensity, whereas at light intensities of 600 and 1050 μmol photons m^{-2}sec^{-1}, the opposite growth pattern was observed (Supplementary Figure S3).

Pigment Content of Cells Grown Under Different Irradiances

As shown in **Table 2**, the cellular contents of Chl a, carotenoids, and PBP were compared in cells grown at low (25 μmol photons m^{-2}sec^{-1}) and high (600 μmol photons m^{-2}sec^{-1}) irradiance. In general, Chl a and PBP levels were lower and carotenoid contents were higher when cells were grown at higher irradiance. However, the reductions of Chl a (about 45–50%) and increases in carotenoid content (about 45%) were more pronounced in the PE A4 and A14 strains compared to the PE A1 strain (about 35 and 10%, respectively). The reduction in PBP contents when cells were grown at high irradiance were of similar magnitude, but the levels were slightly lower for the strains representative of PEs A4 and A14 than for the PE A1 strain. These results indicate that all three strains regulate their pigment content as a function of

irradiance in a manner similar to other cyanobacteria (Bernstein et al., 2014), but that the PE A4 and A14 strains differ in their specific responses from that of the PE A1 strain.

Low-Temperature Fluorescence Emission Spectra

Figure 5A shows low-temperature fluorescence emission spectra for whole cells of the three strains when the excitation wavelength was 440 nm, which selectively excites Chl a. When cells were grown at high irradiance (**Figure 5A**, dashed lines), the overall fluorescence amplitudes for the three strains were similar but generally reflected the differences in Chl a content shown in **Table 2**. As judged by the similar fluorescence emission amplitudes at 683 and 695 nm, the PS II contents of the three strains were very similar, and the emission maximum (722 nm) and amplitudes for PS I emission were also similar. These data indicate that the high-irradiance adapted strain of PE A1 synthesizes slightly more Chl a and produces the most photosynthetic apparatus when grown under high irradiance, but that the PS I to PS II ratio of the three strains is similar when cells are grown at 600 μmol photons m^{-2}sec^{-1}.
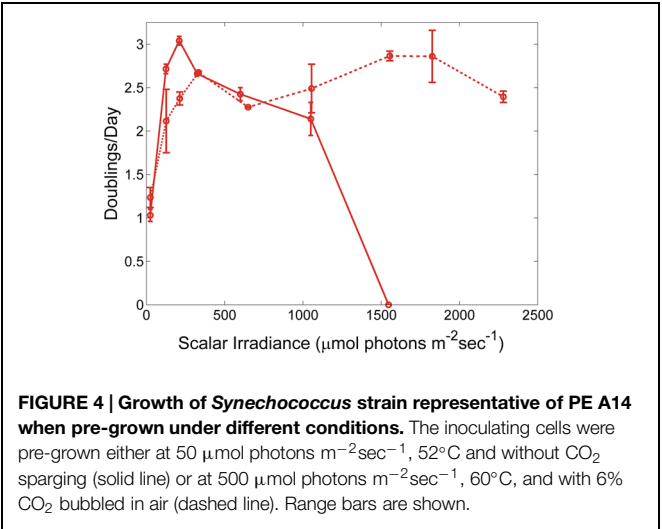
FIGURE 4 | Growth of *Synechococcus* strain representative of PE A14 when pre-grown under different conditions. The inoculating cells were pre-grown either at 50 μmol photons m^{-2}sec^{-1}, 52°C and without CO_2 sparging (solid line) or at 500 μmol photons m^{-2}sec^{-1}, 60°C, and with 6% CO_2 bubbled in air (dashed line). Range bars are shown.

TABLE 2 | Content of chlorophyll (Chl) *a*, carotenoids, and phycobiliproteins (PBP) in cells grown under two irradiance levels.

Strain	Irradiance[a]	Chl a[b]	Carotenoids[b]	Relative PBP
65AY6Li	25	6.88 ± 0.31	2.29 ± 0.14	135%
(PE A1)	600	4.21 ± 0.26	2.56 ± 0.11	100%
65AY6A5	25	7.08 ± 0.37	2.31 ± 0.08	147%
(PE A4)	600	3.23 ± 0.19	3.46 ± 0.32	100%
60AY4M2	25	7.61 ± 0.22	2.25 ± 0.17	145%
(PE A14)	600	3.84 ± 0.17	3.47 ± 0.28	100%

[a]*Irradiance under which strain was grown. Units are μmol photons m^{-2}sec^{-1}.*
[b]*Units are μg/OD_{730}/mL. OD_{730} is an abbreviation for optical density at 730 nm.*
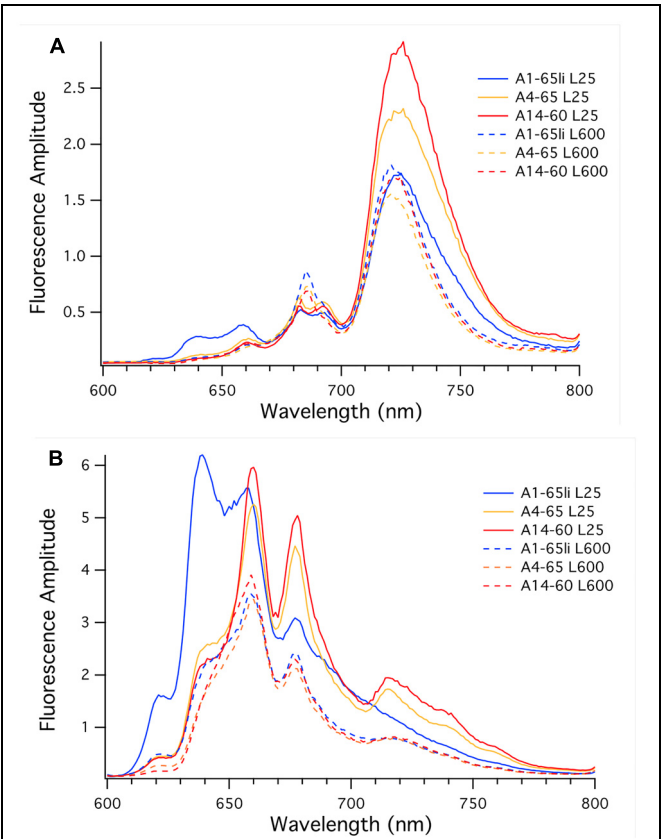
FIGURE 5 | Low-temperature fluorescence emission spectra of whole cells of *Synechococcus* strains representative of PEs A1 (blue), A4 (orange), and A14 (red). (A) The excitation wavelength was set at 440 nm to excite chlorophyll a selectively. **(B)** The excitation wavelength was set at 590 nm to excite phycobiliproteins selectively. Cells were grown at 25 μmol photons m^{-2}sec^{-1} (solid lines) or 600 μmol photons m^{-2}sec^{-1} (dashed lines).

Figure 5A (solid lines) also shows results for the three strains when cells were grown at 25 μmol photons m^{-2}sec^{-1}. As reflected by the similar fluorescence emission amplitudes at 683 and 695 nm, low-light-grown cells of the three strains had very similar PS II contents, but the fluorescence emission from PS I varied strikingly. The emission maximum for the strain of PE A1 had the lowest emission amplitude for PS I; the PE A4 strain had much more PS I than the PE A1 strain, and the PE A14 strain had the most PS I when cells were grown at low light. These values correlate well with the Chl *a* contents of the cells, which increased in the order PE A1 < PE A4 < PE A14. The larger PS I to PS II ratio in strains of PEs A4 and A14 and the increased total Chl *a* content are expected because PS I contains 96 Chl *a* molecules per monomer while PS II only contains 35 Chl *a* molecules per monomer (Jordan et al., 2001; Umena et al., 2011). The observation that the two low-irradiance strains have substantially broader emission bands and greater emission amplitudes between 700 and 800 nm when cells are grown under low irradiance reflects acclimative differences from the PE A1 strain. These differences could originate from Chl-protein complexes that are missing or minimally expressed in the PE A1 strain.

Figure 5B shows the low-temperature fluorescence emission spectra for whole cells of the three strains when the excitation wavelength is 590 nm, which selectively excites PBP. Based upon the genome sequences of *Synechococcus* strains representative of 16S rRNA genotypes A and B′ (Bhaya et al., 2007), these cells should synthesize only two major PBP, phycocyanin, and allophycocyanin. Consistent with the lower relative PBP contents of cells grown at high irradiance (**Table 2**), the emission amplitudes from phycocyanin (∼640 nm) and allophycocyanin (∼660 nm) and the terminal emitters of phycobilisomes (∼678 nm) were much lower in cells grown at high irradiance (**Figure 5B**, dashed lines), but the emission amplitudes were nearly identical for the three strains. These data indicate that the PBP contents of the three strains are very similar when the cells are grown under high irradiance, and that high-light grown cells have lower PBP contents than the cells that are grown under low irradiance.

Compared to the fluorescence emission spectra for strains of PEs A4 and A14, the spectrum for the PE A1 strain was quite different when cells were grown at low irradiance. The PE A1 strain had large emission peaks at 639 and 660 nm, which indicates the presence of substantial pools of uncoupled phycocyanin and allophycocyanin, respectively. The emission amplitude for phycobilisomes was much lower and there was minimal and smoothly declining fluorescence emission beyond 700 nm. The fluorescence emission spectra of cells of strains of PEs A4 and A14 grown at low light were similar and showed strong emission from allophycocyanin (661 nm) and phycobilisomes (678 nm), and intriguingly showed enhanced emission peaks at approximately 720, 740, and 760 nm. The enhanced emission at ∼720 nm might reflect greater coupling of PBP/phycobilisomes with Photosystem I in the cells of strains of PEs A4 and A14 grown at low-irradiance, which is known to occur under state transition conditions (Liu et al., 2013).

However, the latter two emission bands are not observed in cells of these strains grown under high irradiance and are not observed in the PE A1 strains under either condition, and thus these must arise from an acclimative response that specifically occurs in the two low-irradiance strains. Because these emission bands are much stronger when the excitation wavelength was 590 nm, it is likely that these emission bands are associated with far-red-absorbing PBP that are present in cells of PEs A4 and A14 grown under low irradiance but that do not occur in the PE A1 strain. Thus, the fluorescence emission and pigment composition data strongly indicate that the three strains differ in both adaptive and acclimative responses. In particular, the fluorescence data clearly show that strains of PEs A4 and A14 can perform an acclimative response to growth under low-irradiance conditions that the strain of PE A1 cannot perform, which reflects an adaptive difference between PE A1 and PEs A4 and A14. This correlates very well with gene content differences in these three strains, which will be described and discussed in the third paper of this series (Olsen et al., 2015).

Cell Differences Observed with Microscopy and Flow Cytometry

Evidence of *Synechococcus* cells acclimated to different light conditions was also provided by comparing photomicrographs and flow cytometry output of cells grown under low-irradiance and high-irradiance conditions (**Figure 6**). Cells pre-grown at 50 μmol photons m^{-2}sec^{-1} and 52°C without sparging with CO_2 were subsequently grown at 25 or 600 μmol photons m^{-2}sec^{-1}, 60°C and continuous sparging with 6% (v/v) CO_2 in air. As shown in **Figures 6A,B**, cells were smaller, greener, and more brightly autofluorescent when grown at low irradiance compared to when they were grown under high irradiance. This was also observed in the flow cytometer output (**Figure 6C**), in which the cells had a higher autofluorescence signal (PerCP-Cy5-5-A) when grown under low irradiance (green data points) than when grown under high irradiance (red data points). Cells of strains representative of PEs A4 and A14 also showed a rightward shift in forward scatter and an upward shift in side scatter when grown at high irradiance (**Figure 6D**), which suggests that these cells were larger than those grown at low light. Microscopy also revealed that a large percentage of PE A4 and A14 cells (roughly 10 and 25%, respectively), but not the PE A1 cells (<1%), appeared to be unable to immediately separate from the parent cell after division (see **Figure 1C**) when grown at high light. These observations are all consistent with the data on pigment content (**Table 2**) and the fluorescence spectra (**Figure 5**) presented above.

Discussion

The results shown here establish that *Synechococcus* strains corresponding to PEs predicted from pyrosequencing distribution analyses of *psaA* sequences exhibit different light adaptation and acclimation patterns as hypothesized by Becraft et al. (2015) in the first paper of this series. Strains of PEs A4 and A14, which predominate in the deepest part of the
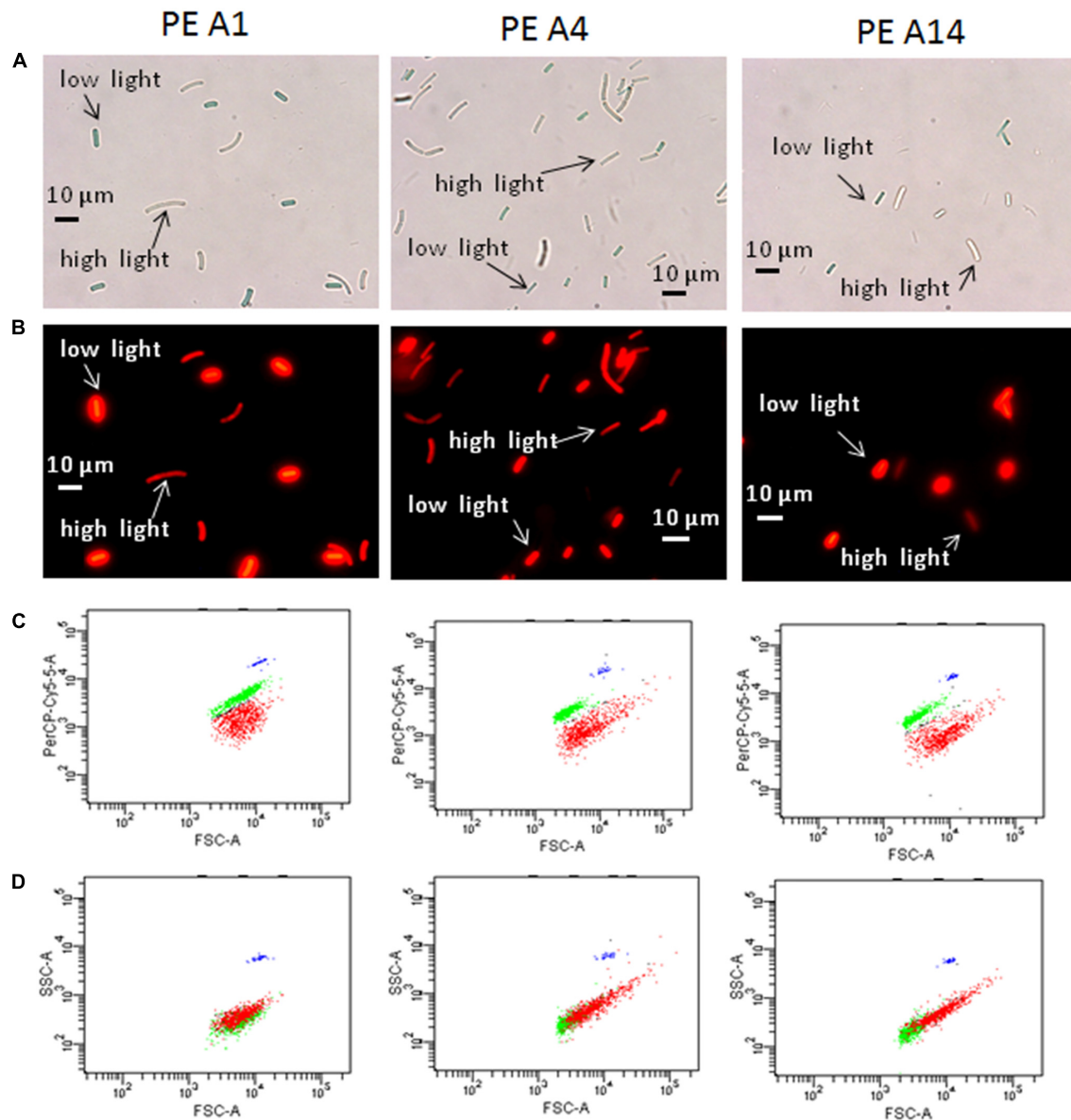
FIGURE 6 | Microscopic and flow cytometric analyses of the strains representative of PEs A1, A4, and A14 grown at 60°C and bubbled with 6% CO₂ in air, under a low-light (25 μmol photons m^{-2}sec^{-1}) and a high-light condition (600 μmol photons m^{-2}sec^{-1}). Pre-growth light intensity was 50 μmol photons m^{-2}sec^{-1}. **(A)** Differential interference contrast microscopy of low-light and high-light grown cells that were collected at the end of exponential growth phase and then mixed for comparative analyses. Scale bar is 10 μm. **(B)** Same images using fluorescence microscopy. **(C)** Scatter plots from BD-FACSCanto flow cytometer of forward scatter (cell size, horizontal axis) versus fluorescence signal (PerCP-Cy5-5-A, vertical axis) of mixed low-light and high-light samples shown in **(A)** and **(B)**. Cells grown under low irradiance are represented by the green data points and cells grown under high irradiance are represented by the red data points. **(D)** Plots of forward scatter (cell size, horizontal axis) versus side scatter (complexity, vertical axis) of mixed low-light and high-light samples shown in **(A)** and **(B)**. The blue data points represent the fluorescent counting beads that serve as a fluorescence control.

1 mm-thick upper green layer of the mat, are able to grow faster at low irradiance levels typically found at this depth than a strain of PE A1, which resides above them and receives more light (see Becraft et al., 2015, Figures 3 and 4). Compared to the PE A1 strain, cells of strains of PEs A4 and A14 had higher Chl *a*, PS I, and PBP contents when grown under low irradiance.

These differences in pigment content, as well as changes in the photosynthetic apparatus, are consistent with shifts in absorbance with depth that have been measured using optical microsensors (see Becraft et al., 2015, Figure 4). The low-light adapted strains also appear to be more sensitive to large changes in irradiance (combined with changes in temperature and CO₂

supply). When light was the only parameter changed between pre-growth and experiment and the degree of increase in light was reduced, the PE A14 strain still grew faster than a strain representative of PE A1 at low irradiances, but was also able to grow at higher irradiances. This might suggest there is an evolutionary trade-off associated with the ability to grow more rapidly under low-light conditions. The two low-light adapted strains also differed from the PE A1 strain in their tendency to increase cell size and decrease post-division cell-separation rate when shifted abruptly to high irradiance conditions.

Previous studies on irradiances supporting optimal photosynthesis were conducted by observing the effects of shading native *Synechococcus* populations in Yellowstone mats with neutral-density filters (Brock and Brock, 1969; Madigan and Brock, 1977). After 10–16 days of exposure to natural light levels, and light levels reduced by 73 and 93%, samples were exposed to a range of irradiances for 5–10 min before addition of $^{14}CO_2$ and incubation for 1 h. Results showed that optimal incorporation of the radiolabel occurred at lower irradiance levels in samples collected from shaded regions of the mat. The upper-light limit (or activity at the upper limit) was also reduced in samples from the shaded mat region. These results were attributed to the acclimation of a single *Synechococcus* population in the mat. Our findings and those of Becraft et al. (2015) suggest that the previously reported results were more likely due to changes in the abundances of *Synechococcus* PEs that are differently adapted to light. The decreased upper-light limit for photosynthetic fixation of $^{14}CO_2$ may have occurred because of sudden exposure of the populations to light intensities much higher than those before the short-term labeling experiments were conducted.

Unfortunately, we did not recover strains representative of PE B′9, which predominates in the uppermost region of the 60–63°C mat. This may have been due to initial cultivation under low-irradiance conditions ($\sim$50 μmol photons m^{-2}sec^{-1}). In a study of *Prochlorococcus* isolates, high-light adapted isolates were unable to grow when very low-irradiance conditions were provided in culture (Moore et al., 1998). Hence, we hypothesize that cells representative of the PE B′9 population may have a higher low-light threshold than the A-like strains characterized here. Interestingly, unlike *Prochlorococcus* sp. isolates, the low-light adapted *Synechococcus* PE A14 strain did have a higher upper-light tolerance when cells were not abruptly shifted to different light, temperature, and CO_2 supply conditions. Preferential selection of A-like *Synechococcus* strains at low light might also explain differences in recovery efficiency at different temperatures, as PE B′9 is present at high relative abundance at 60–63°C, but not at higher temperatures, conditions in which A-like *Synechococcus* predominate (Becraft et al., 2015).

The results clearly demonstrate that strains previously obtained from low-dilution samples contained more than one *Synechococcus* PE. In particular, JA-2-3B′a (2–13) whose genome has been sequenced (Bhaya et al., 2007), might be a consensus genome from three or more different *Synechococcus* PEs. By screening using pyrosequencing, we were able to ensure that the strains studied here were heavily dominated by a single PE. While this approach allowed us to obtain a large number of sequences,

the possibility that very rare members of other PEs are present in these strains still exists (Shrestha et al., 2013). Post-experiment sequencing confirmed that there had not been a shift in PEs under different growth conditions and that measured light relations were for a strain of a single PE. In cases not reported here, we have observed such shifts (Nowack, 2014).

The protocol we applied regularly resulted in growth of colonies from inocula that were three orders of magnitude more dilute than those used by Allewalt et al. (2006). The basis for improved recovery of *Synechococcus* from more highly diluted samples appears to be stimulation of heterotrophic contaminants by the addition of acetate and/or yeast extract. These organisms, in particular *Meiothermus*, develop rapidly in freshly inoculated strains, whereas *Synechococcus* growth initiates as these populations decline. Tank and Bryant (2015) described the purification of another strain from this mat, *Chloracidobacterium thermophilum*, from similar contaminants and the approaches they used may help us to obtain axenic *Synechococcus* cultures as well. At this time we do not know the exact cause of the sudden decline in the *Meiothermus* population. This may possibly be due to (i) nutrient depletion; the relatively small amount of organic substrate added to the medium may be depleted after 24 h of heterotrophic growth at a high rate (12–13 doublings/day), or (ii) a preference for low oxygen levels; at about the time the *Meiothermus* population reaches stationary phase, the *Synechococcus* population starts to increase exponentially, and the increased oxygen produced through photosynthesis may have an adverse effect on the heterotroph (Tank and Bryant, 2015). We are also currently investigating whether *Meiothermus* provides a nutrient(s) needed by *Synechococcus*, or alternatively, may remove toxic substances that limit the development of cyanobacteria (Sakamoto et al., 1998). Model organism experiments suggest that heterotrophs can protect cyanobacteria against reactive oxygen species as one specific beneficial interaction, but sharing of metabolites also clearly occurs (Beliaev et al., 2014).

Although we have demonstrated the existence of *Synechococcus* strains with different adaptations and acclimation responses to light, which are representative of PEs with correspondingly different vertical positioning in the upper mat green layer, it is important to keep in mind that many other physical and chemical parameters vary along the vertical aspect of the mat. For instance, when the rate of photosynthesis is high, oxygen levels are extremely high and CO_2 is reduced to low concentrations, which can result in a sudden and significant increase in pH (Revsbech and Ward, 1984). Differences in the availability of other nutrients (e.g., nitrogen, iron, sulfate, phosphate) may also exist and change dielly. Indeed, genomic analyses of these strains by Olsen et al. (2015), as reported in the third paper in this series, are beginning to reveal such differences.

Acknowledgments

Supplementary Material

References

Allewalt, J. P., Bateson, M. M., Revsbech, N. P., Slack, K., and Ward, D. M. (2006). Effect of temperature and light on growth of and photosynthesis by *Synechococcus* isolates typical of those predominating in the Octopus Spring microbial mat community of Yellowstone National Park. *Appl. Environ. Microbiol.* 72, 544–550. doi: 10.1128/AEM.72.1.544-550.2006

Becraft, E. D., Cohan, F. M., Kühl, M., Jensen, S. I., and Ward, D. M. (2011). Fine-scale distribution patterns of *Synechococcus* ecological diversity in microbial mats of Mushroom Spring, Yellowstone National Park. *Appl. Environ. Microbiol.* 77, 7689–7697. doi: 10.1128/AEM.05927-11

Becraft, E. D., Wood, J. M., Rusch, D. B., Kühl, M., Jensen, S. I., Bryant., D. A., et al. (2015). The molecular dimension of microbial species: 1. Ecological distinctions among, and homogeneity within, putative ecotypes of *Synechococcus* inhabiting the cyanobacterial mat of Mushroom Spring, Yellowstone National Park. *Front. Microbiol.* 6:590. doi: 10.3389/fmicb.2015.00590

Beliaev, A. S., Romine, M. F., Serres, M., Bernstein, H. C., Linggi, B. E., Markille, L. M., et al. (2014). Inference of interactions in cyanobacterial-heterotrophic co-cultures via transcriptome sequencing. *ISME J.* 8, 2243–2255. doi: 10.1038/ismej.2014.69

Bernstein, H. C., Konopka, A., Melnicki, M. R., Hill, E. A., Kucek, L. A., Zhang, S., et al. (2014). Effect of mono- and dichromatic light quality on growth rates and photosynthetic performance of *Synechococcus* sp. PCC 7002. *Front. Microbiol.* 5:488. doi: 10.3389/fmicb.2014.00488

Bhaya, D., Grossman, A. R., Steunou, A. S., Khuri, N., Cohan, F. M., Hamamura, N., et al. (2007). Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analysis. *ISME J.* 1, 703–713. doi: 10.1038/ismej.2007.46

Brock, T. D. (1978). *Thermophilic Microorganisms and Life at High Temperatures*. New York: Springer. doi: 10.1007/978-1-4612-6284-8

Brock, T. D., and Brock, M. L. (1969). Effect of light intensity on photosynthesis by thermal algae adapted to natural and reduced sunlight. *ASLO* 14, 334–341. doi: 10.4319/lo.1969.14.3.0334

Castenholz, R. W. (1969). Thermophilic blue-green algae and the thermal environment. *Bacteriol. Rev.* 33, 476–504.

Ferris, M. J., Kühl, M., Wieland, A., and Ward, D. M. (2003). Cyanobacterial ecotypes in different optical microenvironments of a 68°C hot spring mat community revealed by 16S-23S rRNA internal transcribed spacer region variation. *Appl. Environ. Microbiol.* 69, 2893–2898. doi: 10.1128/AEM.69.5.2893-2898.2003

Ferris, M. J., Nold, S. C., Revsbech, N. P., and Ward, D. M. (1997). Population structure and physiological changes within a hot spring microbial mat community following disturbance. *Appl. Environ. Microbiol.* 63, 1367–1374.

Gilles, A., Meglecz, E., Pech, N., Ferreira, S., Malausa, T., and Martin, J. F. (2011). Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Gen.* 12:245. doi: 10.1186/1471-2164-12-245

Jordan, P., Fromme, P., Witt, H. T., Klukas, O., Saenger, W., and Krauss, N. (2001). Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature* 411, 909–917. doi: 10.1038/35082000

Kilian, O., Steunou, A. S., Fazeli, F., Bailey, S., Bhaya, D., and Grossman, A. R. (2007). Responses of a thermophilic *Synechococcus* isolate from the microbial mat of Octopus Spring to light. *Appl. Environ. Microbiol.* 73, 4268–4278. doi: 10.1128/AEM.00201-07

Koeppel, A., Perry, E. B., Sikorski, J., Krizanc, D., Warner, A., Ward, D. M., et al. (2008). Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2504–2509. doi: 10.1073/pnas.0712205105

Liu, H., Zhang, H., Niedzwiedzki, D. M., Prado, M., He, G., Gross, M. L., et al. (2013). Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria. *Science* 342, 1104–1107. doi: 10.1126/science.1242321

Madigan, M. T., and Brock, T. D. (1977). Adaptation by hot spring phototrophs to reduced light intensities. *Arch. Microbiol.* 113, 111–120. doi: 10.1007/BF00428590

Melendrez, M. C., Lange, R. K., Cohan, F. M., and Ward, D. M. (2011). Influence of molecular resolution on sequence-based discovery of ecological diversity among *Synechococcus* populations in an alkaline siliceous hot spring microbial mat. *Appl. Environ. Microbiol.* 77, 1359–1367. doi: 10.1128/AEM.02032-10

Miller, S. R., and Castenholz, R. W. (2000). Evolution of thermotolerance in hot spring cyanobacteria of the genus *Synechococcus*. *Appl. Environ. Microbiol.* 66, 4222–4229. doi: 10.1046/j.1529-8817.1999.00001-143.x

Moore, L. R., Rocap, G., and Chisholm, S. W. (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* 393, 464–467. doi: 10.1038/30965

Nobre, M. F., and Da Costa, M. S. (2001). "The genus *Meiothermus*," in *Bergey's Manual of Systematic Bacteriology*, eds D. R. Boone, R. W. Castenholz, and G. M. Garrity (New York: Springer).

Nowack, S. (2014). *Niche Character in a Temporally Varying Environment*. Ph.D. dissertation, Montana State University, Bozeman, MT.

Olsen, M. T., Nowack, S., Wood, J. M., Becraft, E. D., LaButti, K., Lipzen, A., et al. (2015). The molecular dimension of microbial species: 3. Comparative genomics of *Synechococcus* strains with different light responses and *in situ* diel transcription patterns of associated ecotypes in the Mushroom Spring microbial mat. *Front. Microbiol.* 6:604. doi: 10.3389/fmicb.2015.00604

Peary, J. A., and Castenholz, R. W. (1964). Temperature strains of a thermophilic blue-green alga. *Nature* 202, 720–721. doi: 10.1038/202720a0

Ramsing, N. B., Ferris, M. J., and Ward, D. M. (2000). Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter thick photic zone of a hot spring cyanobacterial mat. *Appl. Environ. Microbiol.* 66, 1038–1049. doi: 10.1128/AEM.66.3.1038-1049.2000

Revsbech, N. P., and Ward, D. M. (1984). Microelectrode studies of interstitial water chemistry and photosynthetic activity in a hot spring microbial mat. *Appl. Environ. Microbiol.* 48, 270–275.

Sakamoto, T., Delgaizo, V. B., and Bryant, D. A. (1998). Growth on urea can trigger death and peroxidation of the cyanobacterium *Synechococcus* sp. strain PCC 7002. *Appl. Environ. Microbiol.* 64, 2361–2366.

Sekiguchi, Y., Yamada, T., Hanada, S., Ohashi, A., Harada, H., and Kamagata, Y. (2003). *Anaerolinea thermophila* gen. nov., sp. nov. and *Caldilinea aerophila* gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain Bacteria at the subphylum level. *Int. J. Syst. Evol. Micrbiol.* 53, 1843–1851. doi: 10.1099/ijs.0.02699-0

Shen, G., Schluchter, W. M., and Bryant, D. A. (2008). Biogenesis of phycobiliproteins I. cpcS and cpcU mutants of the cyanobacterium *Synechococcus* sp. PCC 7002 define a heterodimeric phycobiliprotein lyase specific for Cys-82/84 sites of the β-phycocyanin and allophycocyanin subunits. *J. Biol. Chem.* 283, 7503–7512. doi: 10.1074/jbc.M7081 64200

Shen, G., Zhao, J., Reimer, S. K., Antonkine, M. L., Cai, Q., Weiland, S. M., et al. (2002). Assembly of Photosystem I. I. Inactivation of the *rubA* gene encoding a membrane-associated rubredoxin in the cyanobacterium *Synechococcus* sp. PCC 7002 causes a loss of Photosystem I activity. *J. Biol. Chem*. 277, 20343–20354. doi: 10.1074/jbc.M201103200

Sheridan, R. P. (1976). Sun-shade ecotypes of a bluegreen alga in a hot spring. *J. Phycol.* 12, 279–285. doi: 10.1111/j.1529-8817.1976.tb02845.x

Sheridan, R. P. (1978). Seasonal variation in sun-shade ecotypes of *Plectonema notatum* Cyanophyta. *J. Phycol.* 15, 223–226. doi: 10.1111/j.1529-8817.1979.tb02989.x

Shrestha, P. M., Nevin, K. P., Shrestha, M., and Lovley, D. R. (2013). When is a microbial culture "pure"? Persistent cryptic contaminant escapes detection even with deep genome sequencing. *mBio* 4:e00591-12. doi: 10.1128/mBio.00591-12

Tank, M., and Bryant, D. A. (2015). Nutrient requirements and growth physiology of the photoheterotrophic Acidobacterium, *Chloracidobacterium thermophilum. Front. Microbiol.* 6:226. doi: 10.3389/fmicb.2015.00226

Umena, Y., Kawakami, K., Shen, J. R., and Kamiya, N. (2011). Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* 473, 55–60. doi: 10.1038/nature09913

Zhao, J., Shen, G., and Bryant, D. A. (2001). Photosystem stoichiometry and state transition in a mutant of the cyanobacterium *Synechococcus* sp. PCC 7002 lacking phycocyanin. *Biochim. Biophys. Acta.* 1505, 248–257. doi: 10.1016/S0005-2728(01)00175-X

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The molecular dimension of microbial species: 3. Comparative genomics of *Synechococcus* strains with different light responses and *in situ* diel transcription patterns of associated putative ecotypes in the Mushroom Spring microbial mat

***Correspondence:**
Millie T. Olsen,
Department of Land Resources and
Environmental Sciences, Montana
State University, 960 Technology Blvd,
Molecular Biosciences Building,
Bozeman, MT 59718, USA
millie.thornton@msu.montana.edu

Millie T. Olsen[1]*, Shane Nowack[2], Jason M. Wood[1], Eric D. Becraft[1], Kurt LaButti[3], Anna Lipzen[3], Joel Martin[3], Wendy S. Schackwitz[3], Douglas B. Rusch[4], Frederick M. Cohan[5], Donald A. Bryant[6,7] and David M. Ward[1]

[1] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA, [2] Department of Mathematical Sciences, Montana State University, Bozeman, MT, USA, [3] Department of Energy, Joint Genome Institute, Walnut Creek, CA, USA, [4] J. Craig Venter Institute, Rockville, MD, USA, [5] Department of Biology, Wesleyan University, Middletown, CT, USA, [6] Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA, [7] Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA

Genomes were obtained for three closely related strains of *Synechococcus* that are representative of putative ecotypes (PEs) that predominate at different depths in the 1 mm-thick, upper-green layer in the $60°C$ mat of Mushroom Spring, Yellowstone National Park, and exhibit different light adaptation and acclimation responses. The genomes were compared to the published genome of a previously obtained, closely related strain from a neighboring spring, and differences in both gene content and orthologous gene alleles between high-light-adapted and low-light-adapted strains were identified. Evidence of genetic differences that relate to adaptation to light intensity and/or quality, CO_2 uptake, nitrogen metabolism, organic carbon metabolism, and uptake of other nutrients were found between strains of the different putative ecotypes. *In situ* diel transcription patterns of genes, including genes unique to either low-light-adapted or high-light-adapted strains and different alleles of an orthologous photosystem gene, revealed that expression is fine-tuned to the different light environments experienced by ecotypes prevalent at various depths in the mat. This study suggests that strains of closely related PEs have different genomic adaptations that enable them to inhabit distinct ecological niches while living in close proximity within a microbial community.

Keywords: comparative genomics, microbial species, thermophilic *Synechococcus*, microbial mats, adaptation

Introduction

Thermophilic cyanobacteria of the genus *Synechococcus* predominate in microbial mat communities inhabiting the effluent channels of alkaline, siliceous hot springs and have been extensively studied and characterized for over 50 years (Peary and Castenholz, 1964; Brock, 1978; Ward et al., 2012). Strains of *Synechococcus* from mats in Hunters Hot Springs, OR, which were found to exhibit different temperature adaptations, were first observed by Peary and Castenholz (1964). Molecular analyses of Octopus Spring, Yellowstone National Park (YNP), based on 16S rRNA sequences provided further evidence of *Synechococcus* ecotypes—five closely related genotypes were found to be distributed differently from high to low temperatures along the thermal gradient of the effluent channel (Ferris and Ward, 1997). Genotype A″ inhabited the highest temperatures, followed by A′, A, B′, and B as temperatures decreased with distance from the source, though some overlap of adjacent genotypes was observed. Strains with these 16S rRNA genotypes were shown to have different temperature adaptations that correlated with their distribution along the effluent channel (Allewalt et al., 2006). 16S rRNA genotypes were also found to differ along the vertical aspect of the mat. For example, differential vertical distributions of genotype B′, which occurred above genotype A in the 60°C mat, were observed (Ramsing et al., 2000). However, fluorescence microscopy, combined with estimates of oxygenic photosynthetic rates calculated using oxygen microsensor measurements in a 68°C mat sample, revealed physiologically distinct populations in the lower and upper parts of the top green mat layer. These populations were identical at the 16S rRNA locus and might have been interpreted as one genotype acclimated differently in response to lower light intensity. However, surface and subsurface populations were genetically distinct at the more rapidly evolving internal transcribed spacer locus that separates the 16S and the 23S rRNA genes (Ferris et al., 2003), suggesting the possible existence of yet more closely related *Synechococcus* populations with different adaptations to light.

Based on these observations, Ward and Cohan (2005) and Ward et al. (2006) foresaw the need for theory-based models to predict putative ecological species, or ecotypes, from natural variation in sequence data, and for studying genes with even higher molecular resolution than 16S rRNA and the 16S–23S rRNA internal transcribed spacer region. Most recently, *Synechococcus* putative ecotypes (PEs) have been demarcated using highly resolving, protein-encoding loci, from which the evolutionary simulation algorithm Ecotype Simulation (Koeppel et al., 2008) has predicted an even greater number of PEs than demarcated by the internal transcribed spacer region (Becraft et al., 2011; Melendrez et al., 2011). As an example, in the first paper of this three-paper series on the molecular dimensions of microbial species, Becraft et al. (2015) used *psaA* sequence variation and Ecotype Simulation to predict many PEs, including seven PEs in the A′ lineage, 15 PEs in the A lineage, and 24 PEs in the B′ lineage, several of which were shown to have different vertical distributions in the mat. At 60–63°C, PEs B′9, A1, A4, A14, and A6 were found to be progressively

predominant from the mat surface to the bottom of the upper 1 mm-thick green layer (summarized in **Table 1**; also see Figure 3 in Becraft et al., 2015). This led to the hypothesis that these PEs are adapted to different irradiances corresponding the light levels they experience *in situ*. In the second paper of the series, Nowack et al. (2015) reported the successful cultivation of strains representative of PEs A1, A4, and A14, which they used to test the hypothesis. Strains representative of these PEs were shown to have distinctive growth patterns, pigment contents, and low-temperature fluorescence emission spectra when grown at either high or low irradiance. These differences indicated adaptive and/or acclimative responses to irradiance levels and light qualities that are characteristic of the depth at which each PE predominates *in situ* (see Figures 4, 8C in Becraft et al., 2015).

While individual genes can be used to predict ecotypes whose unique niches have been inferred from their microhabitat distributions (Becraft et al., 2015), and whose existence can be confirmed by the phenotypes of strains (Nowack et al., 2015), whole genome comparative analysis can reveal genetic differences among strains that may be responsible for the adaptive and acclimative mechanisms. For example, the genomes of closely related strains of *Prochlorococcus* spp., which are prevalent phototrophs in marine environments, have been sequenced and compared. Strains that have different light adaptations maintain differences in gene content related to adaptation to the specific light and nutrient environment of the surface-associated high-light or deep-water-associated low-light layers in the ocean (Rocap et al., 2003). *Prochlorococcus* spp. strains with similar light adaptations also maintained "genomic islands" that may aid in niche differentiation of ecotypes that co-exist within the high-light or the low-light portions of the water column (Coleman et al., 2006; Kettler et al., 2007). Furthermore, genomic analyses are open-ended and unconstrained by the limits of our intuition in that they may reveal unsuspected differences for physiological or metabolic functions that have not yet been tested experimentally. For instance, in previous work the genome of a *Synechococcus* genotype found in downstream regions of these hot spring mats (i.e., a B′-like strain) was shown to have genes for nitrogen storage and metabolism and for phosphonate utilization that were lacking in an upstream genotype (i.e., A-like), indicating that populations along the flow path differ in adaptations for nutrient metabolism as well as temperature (Bhaya et al., 2007).

In this study, we compared the genomes of four *Synechococcus* strains within the A lineage that are representative of PEs known to be predominant at different depths in the 60–63°C Mushroom Spring mat (Becraft et al., 2015). Strains representative of PE A1, which is found closer to the mat surface, and PEs A4 and A14, which are found deeper in the mat upper green layer, were shown to have different adaptations and acclimative behaviors to low and high irradiance (Nowack et al., 2015). A second PE A1 strain, which had been previously cultivated from Octopus Spring (Allewalt et al., 2006) was shown to have light responses that were indistinguishable from that of the other PE A1 strain (Nowack, 2014). We compared the genomes of these strains of high-light and low-light adapted organisms to identify differences in gene content and specific alleles that might underlie these and

TABLE 1 | Genomic, phenotypic, and environmental information for *Synechococcus* strains of putative ecotypes with different depth distributions[a] or from different hot spring mats.

Strain	A4	A14	A1-MS	A1-OS
Strain name	65AY6A5	60AY4M2	65AY6Li	JA-3-3Ab
Putative ecotype (PE)	A4	A14	A1	A1
Collection source	Mushroom Spring	Mushroom Spring	Mushroom Spring	Octopus Spring
Isolation date	September 2010	September 2010	September 2010	July 2002
Isolation temperature (°C)	65	60	65	58–65[b]
PE relative abundance $\geq$5% at 60°C Mushroom Spring	400–800 μm	400–960 μm	0–960 μm	0–960 μm
PE relative abundance $\geq$10% at 60°C Mushroom Spring	640–720 μm	560–960 μm	0–800 μm	0–800 μm
Light adaptation	Low-light	Low-light	High-light	High-light
Sequencing method	Illumina	Illumina	Illumina	Sanger
Depth of coverage	22×	35×	16×	2×
Length (Mbp)	2.98	3.16	2.93	2.93
Number of contigs	9	6	2	1
Largest contig (bp)	2,508,234	3,142,301	2,795,989	2,932,766
%GC content	60.4	60.4	60.3	60.2
CDS	2622	2597	2623	2760
tRNAs	47	47	49[c]	47
rRNA operons	2	2	2	2
Number of genes unique to the strain with respect to the reference A1-OS	173	204	131	N/A

[a] *Distributions are expressed as a range of depths where the PE has relative abundance of either $\geq$5 or $\geq$10% abundance, rather than emphasizing the peak population abundance of a PE (e.g., Table 1 in Nowack et al., 2015).*

[b] *Temperatures in Octopus Spring fluctuate continuously over a 4.5 min cycle (Miller et al., 1998), therefore the 7°C range of the isolation site is given.*

[c] *A1-MS contains a duplication of the 23S rRNA locus in one operon and two adjacent tRNA loci in one operon.*

other adaptations and acclimative responses. We also used these genomes to probe transcript abundances for specific ecotypes using a diel metatranscriptome dataset we had previously obtained (Liu et al., 2012). We sought evidence of differences in transcription patterns for homologous genes shared among species, which were divergent enough to differentiate PEs, as well as strain-specific genes, which may be representative of each PE, including genes involved in light harvesting and nutrient uptake. These differences may be indicative of mechanisms underlying adaptive and acclimative responses to light intensity, light quality, and nutrient use that reflect the distinct, ecological niches of these PEs.

Materials and Methods

Synechococcus Strains

Strains representative of *Synechococcus* PE A1 (65AY6Li), PE A4 (65AY6A5), and PE A14 (60AY4M2), were selected for comparative genomic analysis. For simplicity we will refer to these strains by their PE affiliations or by their known adaptive and/or acclimative responses to low light (strains of PEs A4 and A14) or high light (strains of PE A1). The DNA samples used to demonstrate strain purity by Ti-454 barcode sequencing (Nowack et al., 2015) were also used for genome sequencing. These genomes were compared to the genome of a second strain of PE A1 (JA-3-3Ab), previously obtained from the microbial mat of Octopus Spring (Bhaya et al., 2007), which was also shown to be adapted to high light (Nowack, 2014). To distinguish between

these two PE A1 strains, we will refer to them here as strain PE A1-MS (from Mushroom Spring) and PE A1-OS (from Octopus Spring). Mushroom Spring is located ~0.5 km from Octopus Spring, and the two alkaline siliceous springs have been shown to have similar major ion chemistry over decades (see Brock, 1978; Inskeep et al., 2013; the YNP Research Coordination Network website[1]) and inhabitants (Ramsing et al., 2000; Becraft, 2014).

Genome Sequencing, Assembly, and Annotation

Purified total genomic DNA from each strain was submitted to the Department of Energy Joint Genome Institute for sequencing and assembly. The DNA from each strain was randomly sheared into ~270 bp fragments and the resulting fragments were used to create fragment libraries. These libraries were sequenced on Illumina sequencers generating 150-bp paired-end reads. All general aspects of library construction and sequencing are described on the JGI website[2]. Because the variant detection pipeline requires some non-overlapping paired reads prior to variant detection, the reads were then trimmed to 125 bp. These trimmed reads were then aligned to the reference genome *Synechococcus* sp. JA-3-3Ab using the Burrows-Wheeler Aligner (BWA) (Li and Durbin, 2009), and putative single-nucleotide polymorphisms (SNPs) and small indels were identified using samtools and mpileup (Li et al., 2009). Putative structural variants were identified using BreakDancer (Chen

---

[1]http://www.rcn.montana.edu/Features/DataDetail.aspx?id=124 for Mushroom Spring data and http://www.rcn.montana.edu/Features/DataDetail.aspx?id=37 for Octopus Spring data

[2]http://www.jgi.doe.gov/

et al., 2009), filtering for a confidence score of >90. Genomes were also assembled *de novo* for each strain. Each FASTQ file was QC-filtered for artifact/process contamination and subsequently assembled with AllPathsLG (Gnerre et al., 2011). The resulting contigs included DNA sequences of heterotrophic contaminants that occur in these strains (Nowack et al., 2015). These sequences were separated from *Synechococcus* sequences by binning the sequences using NCBI BLASTN and a database of the *Synechococcus* spp. JA-3-3Ab and JA 2-3B′a(2-13) reference genomes (Bhaya et al., 2007), and those of the possible contaminants (*Meiothermus* spp., CP005385; *Anyoxybacillus* spp., CP000922; *Rubrobacter* spp., CP000386) using a method similar to that described in Klatt et al. (2011). The *Synechococcus* DNA assemblies were submitted to the automatic annotation pipelines NCBI PGAAP [NCBI Handbook (Internet) 2nd edition (Tatusova et al., 2013)] and RAST (Aziz et al., 2008) for annotation using the default parameters. The three draft genome sequences have been submitted to Genbank under the following accession numbers: PRJNA209725 (*Synechococcus* sp. 65AY6Li, PE A1), PRJNA210217 (*Synechococcus* sp. 65AY6A5, PE A4), and PRJNA210214 (*Synechococcus* sp. 60AY4M2, PE A14).

Comparative Analyses

The strain genome phylogeny was obtained using a concatenation of 460 marker proteins identified by Phyla-AMPHORA (Wang and Wu, 2013), which uses phylum-specific conserved proteins for metagenomic phylotyping. Only conserved cyanobacterial proteins present in all five genomes (the strains studied here and *Synechococcus* sp. JA-2-3B′a(2-13), which was used as an outgroup) were selected (highlighted in Supplementary Table S1). Proteins were aligned with ClustalO (Sievers et al., 2011) and a Newick tree was computed using FastTree, with local support values calculated using the Shimodaira_Hasegawa test (Price et al., 2009). Comparative analyses of genomes were conducted using the RAST SEEDViewer for gene content analyses (Overbeek et al., 2014) and the best-hit and reciprocal best-hit average nucleotide identity (ANI) calculator (Goris et al., 2007).

Metatranscriptomic Analyses

Diel metatranscriptomic datasets described by Liu et al. (2012), which were based on analysis of pooled triplicate samples from the Mushroom Spring 60°C mat, collected at hourly intervals throughout a complete diel cycle, were reanalyzed by BWA (Li and Durbin, 2009) to locate transcripts for specific genes. Genes targeted in these analyses were chosen either because they are specific to low-light- or high-light-adapted strains or, in the case of orthologous genes, because they exhibited at least 3% nucleotide sequence difference between/among homologous genes in different strains. We used the methods described in Liu et al. (2011, 2012), except that (i) we used genomes instead of metagenomic assemblies to recruit transcripts, and (ii) transcripts associated with the unique alleles of different ecotypes were recruited without allowing any mismatches (i.e., an exact sequence match was required). Recruitment of transcripts associated with B′-like *Synechococcus* was done using the

published *Synechococcus* B′ genome (Bhaya et al., 2007), but, as in Liu et al. (2012), up to 5 mismatches were allowed. This was done because this genome is not representative of the predominant B′ PE (B′9) in the 60°C mat (Becraft et al., 2015), which we do not yet have in culture (Nowack et al., 2015). Raw transcript counts were normalized by the total number of mRNA-specific transcripts at each time point and then by the geometric mean of normalized transcript counts across all time points (Liu et al., 2011, 2012).

Results

Genomic Properties

Basic characteristics of the genomes of strains representative of PEs A1-OS, A1-MS, A4, and A14 are presented in **Table 1**. The strains are 99.93–100% identical at the 16S rRNA locus and share 2201 orthologous genes as their core genome, including most RAST annotated subsystem genes found in the A1-OS reference genome. The ANI among orthologous genes in the different strains ranged from 98.35 to 99.32%, so many of the shared genes predict proteins with 100% sequence identity and are likely to be functionally identical (**Table 2**).

The GC contents of the strains ranged from 60.2 to 60.4%, while genome sizes varied from 2.93 Mbp for strains of PEs A1-OS and A1-MS to 3.16 Mbp for the PE A14 strain (**Table 1**). The strain genome phylogeny reflects the *psaA* phylogeny (see Figure 1 in Becraft et al., 2015), with the high-light-adapted strains and the low-light-adapted strains each forming a distinct clade (**Figure 1**). Although all sequenced A-lineage strains are very closely related, there are several genomic differences among the strains that may underlie the niche differentiation among the PEs. The three strains were selected because they are representative of PEs that differ in vertical position and exhibit different adaptations to irradiance, but the low-light-adapted strains of PEs A4 and A14 were similar to each other phenotypically (see Nowack et al., 2015) and were very different from the two high-light-adapted strains A1-MS and A1-OS. Hence, this discussion will focus on differences between the two low-light-adapted strains and the two high-light-adapted strains. Differences discussed in the main text, including any subsystem genes missing in the newly sequenced strains, are presented in **Table 3** and a full ortholog table, which also presents differences in the percentage amino acid identity of homologous genes,

TABLE 2 | Average nucleotide index (ANI) and genes with identical sequence in *Synechococcus* strain genomes.

	A4	A14	A1-MS	A1-OS
A4		42.65%	37.65%	36.63%
A14	99.05%		37.08%	43.77%
A1-MS	98.41%	98.72%		55.07%
A1-OS	98.42%	98.35%	99.32%	

The ANI between each pair of genomes, using whole-genome reciprocal best hits is shown below the diagonal. The percentage of genes that encode for proteins with identical amino acid sequences (and may be functionally identical) between each pair of genomes is shown above the diagonal.

FIGURE 1 | Phylogeny based on a concatenation of 460 conserved cyanobacterial proteins found in the A-lineage strains. The tree was rooted by *Synechococcus* sp. JA-2-3B′a(2-13) and the scale bar is equal to 0.001 amino acid substitutions/site.
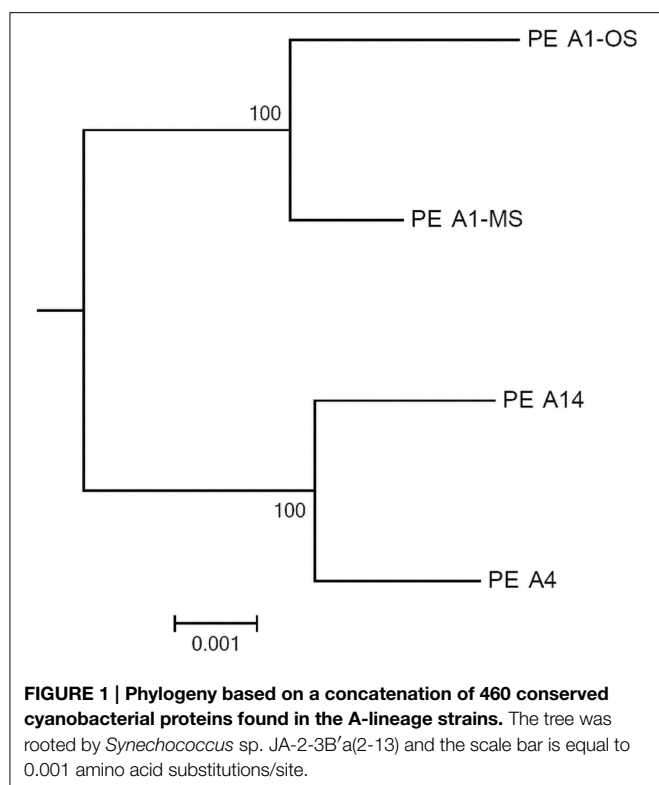
is presented in Supplementary Table S1. Specific differences between the genomes of PE A4 and A14 strains will also be considered below.

Genes Found Only in Low-light-adapted Strains and their Diel Transcription Patterns

The low-light-adapted strains representative of PEs A4 and A14 have a unique, possibly horizontally acquired gene cluster that includes: a potential photoreceptor predicted to have four PAS domains, two GAF domains, and a histidine kinase domain, which could act as a light-activated response regulator; *apcD4* and *apcB3* genes, which are predicted to encode a variant allophycocyanin that probably has enhanced far-red absorption (Gan et al., 2014b); and a gene for an IsiA-like protein, which we have tentatively named IsiX. Allophycocyanins are phycobiliproteins that absorb red and far-red light and form light-harvesting antenna complexes for Photosystems I and II in cyanobacteria (Gan et al., 2014a,b; Sidler, 2004). ApcD4 is approximately 40% identical and 62% similar to ApcD1 (CYA_2790) in these *Synechococcus* strains and is ~63% similar to ApcA (CYA_2227) of the PE A1-OS strain; ApcB3 is ~80% similar to the ApcB (CYA_2226) of the PE A1-OS strain. In contrast, the products of the *apcA* and *apcB* genes, which are not located in this cluster, are highly conserved in all four strain genomes (100% identical). ApcD4 and ApcB3 are only found in a few cyanobacteria and probably form a variant type of allophycocyanin with enhanced far-red light absorption (670–710 nm absorption maximum) (Gan et al., 2014b).

IsiA and its paralogs are chlorophyll-binding proteins that form specialized light-harvesting antenna complexes in

cyanobacteria (Kouril et al., 2005; Murray et al., 2006). IsiX belongs to the PsbC/IsiA superfamily (Kouril et al., 2005; Murray et al., 2006) of chlorophyll (Chl) *a*-binding antenna complex proteins but is quite distinct from IsiA, which typically produces a specialized light-harvesting complex under iron-starvation conditions (**Table 3**). IsiX, which has a C-terminal extension of nearly 100 amino acids and probably has one additional transmembrane helix relative to other PsbC/IsiA proteins, is only 48% similar to the paralogous Photosystem II core subunit, PsbC, which is >99% similar among all four strains. Moreover, IsiX is only 46% similar to IsiA (CYA_2606), which is likely to be iron-regulated because of its co-localization with *isiB*, encoding flavodoxin (IsiB; CYA_2605), as observed in most other cyanobacteria (Straus, 1994). These observations are consistent with the idea that ApcD4-ApcB3 and IsiX are specialized antenna proteins that function in PE A4 and A14 strains under low irradiance or possibly far-red light (or both) conditions. Consistent with this idea, we have noted that strains corresponding to PEs A4 and A14 have enhanced absorption above 700 nm compared to the PE A1-MS strain (data not shown, but see Nowack et al., 2015). However, at this point it is not yet clearly established whether these proteins are linked to these differences.

Transcript abundances for genes encoding components of the photosynthetic apparatus in the Mushroom Spring mat *Synechococcus* generally rise sharply at sunrise, are maximal during the mid-day, and decline in the late afternoon (see Figure 4D in Liu et al., 2012). This pattern is observed for *psbC*, which encodes a core subunit of the Photosystem II reaction center (**Figure 2A**). However, the transcript abundance for *isiX* has a different pattern and is most abundant during the low-light periods in the early morning (07:00–10:00) and late afternoon (15:00–19:00) (**Figure 2A**). Transcripts for the *apcD4* and *apcB3* genes have a similar overall abundance pattern to *isiX* (**Figure 2B**). Although transcripts for *apcA*, *apcB*, *apcD4*, and *apcB3* were all maximal at the same time in the morning (09:00), transcripts for *apcD4* and *apcB3* were maximal about an hour later than transcripts for *apcA* and *apcB* in the late afternoon period. Our confidence in the observed patterns is based on trends established by adjacent, closely spaced time points, highly similar co-expression patterns for different genes in this cassette, and correspondence with previously published transcription patterns (Steunou et al., 2006, 2008; Liu et al., 2012).

Another gene cassette unique to the low-light-adapted strains representative of PEs A4 and A14 contains the *feoAB* genes, which encode subunits of a ferrous iron transporter. The transcript abundances for these genes are maximal in the late afternoon, when the mat is becoming anoxic (**Figure 2C**; and see Figure 8C in Becraft et al., 2015). The most closely related FeoAB protein sequences are found in other cyanobacteria, but it is not clear if the genes were acquired by horizontal gene transfer, or lost in the high-light-adapted strains, which are representative of a PE that predominates in the more oxic portions of the mat. Strains of PEs A14 and A4 also share an ABC transporter cassette for sugar (possibly maltose/maltodextrin) transport and a paralogous methyl-accepting chemotaxis protein, one of many copies found in all four genomes. The latter gene has two identical

TABLE 3 | Ortholog table showing discussed gene content differences among strains with different light adaptations and representative of putative ecotypes with different vertical positioning in the 60–63°C Mushroom Spring mat.

A4	A14	A1-MS	A1-OS
Chlorophyll a(b) binding protein, photosystem II CP43 protein homolog *isiX*/ Allophycocyanin subunit *apcB3*/ Allophycocyanin subunit *apcD4*/Hypothetical protein (putative photoreceptor)	Chlorophyll a(b) binding protein, photosystem II CP43 protein homolog *isiX*/ Allophycocyanin subunit *apcB3*/ Allophycocyanin subunit *apcD4*/ Hypothetical protein (putative photoreceptor)		
Ferrous iron transport cassette *feoAB* (2 genes)	Ferrous iron transport cassette *feoAB* (2 genes)		
Maltose/maltodextrin transport cluster *malK* (4 genes)	Maltose/maltodextrin transport cluster *malK* (4 genes)		
Ammonium transporter *amtB2*	Ammonium transporter *amtB2*		
Ammonium transporter *amtB1*	Ammonium transporter *amtB1*	Ammonium transporter *amtB1*	Ammonium transporter *amtB1*
Assimilatory nitrate reductase	Assimilatory nitrate reductase	Assimilatory nitrate reductase (inactivated)	Assimilatory nitrate reductase (inactivated)
Methyl-accepting chemotaxis protein	Methyl-accepting chemotaxis protein	Methyl-accepting chemotaxis protein (inactivated)	Methyl-accepting chemotaxis protein (inactivated)
		Carbonic Anhydrase	Carbonic Anhydrase
		Peptide/opine/ nickel PepT ABC Transport Cassette (5 genes)	Peptide/opine/ nickel PepT ABC Transport Cassette (5 genes)
		Urease Cluster 1 (5 genes)	Urease Cluster 1 (5 genes)
		Cystine ABC Transporter (2 genes)	Cystine ABC Transporter (2 genes)
		Succinate dehydrogenase flavoprotein subunit *sdhA*/ Omega-amino acid-pyruvate amino-transferase	Succinate dehydrogenase flavoprotein subunit *sdhA*/ Omega-amino acid-pyruvate amino-transferase
		Type III CRISPR/*cas* array	Type III CRISPR/*cas* array
Urea Carboxylase Cassette (7 genes)	Urea Carboxylase Cassette (7 genes)	Urea Carboxylase Cassette (7 genes)	
Polar amino acid ABC Transport (PAAT) (3 genes)		Polar amino acid ABC Transport (PAAT) (3 genes)	Polar amino acid ABC Transport (PAAT) (3 genes)
	Sugar ABC Transport Cassette (3 genes)	Sugar ABC Transport Cassette (3 genes)	Sugar ABC Transport Cassette (3 genes)
Spermidine Putrescine ABC Transporter PotABCD (6 genes)			
Beta-carotene ketolase			
Formamidase *amiF*/Amidase Gene Cluster (8 genes)			
		Type I restriction-modification system/SS exonuclease associated with Rad50/Mre11 complex/Bipolar DNA Helicase	

in-frame stop codons in genomes of the high-light-adapted PE A1-OS and A1-MS strains.

In addition to horizontal gene transfer, gene duplication and subsequent nucleotide divergence can provide novel functionality to an organism, even though the resulting variant protein retains homology to the original product. All four strains contain a gene (*amtB1*) encoding a putative ammonium transporter, and the predicted AmtB1 proteins are ∼90% identical. However, the strains of PEs A4 and A14 additionally contain a paralogous gene that has apparently arisen by duplication and divergence: AmtB2 is ∼70% identical to AmtB1. The transcription patterns of the *amtB1* gene in the PE A1 strains and in the PE A4 and A14 strains are comparable, but the transcription pattern of the *amtB2* gene in the PE A4 and

A14 strains differ from the *amtB1* pattern. This gene has a transcription pattern similar to many genes for components of the photosynthetic apparatus (Liu et al., 2012) and largely reflects the light period except for a late-afternoon decline (**Figure 2D**). Functional studies have shown that AmtB can transport both NH_3 and CO_2 (Musa-Aziz et al., 2009), and it is possible that these variants are functionally differentiated with respect to substrate. Because the transcript abundance pattern mirrors photosynthetic activity in the mat, this pattern is consistent with the possibility that AmtB2 could be a CO_2 transporter. Alternatively, AmtB2 could transport ammonium but have a high affinity for the substrate. In addition to the duplicated *amtB* genes, strains of PEs A4 and A14 contain a second copy of *narB*, encoding assimilatory nitrate reductase, which is transcribed
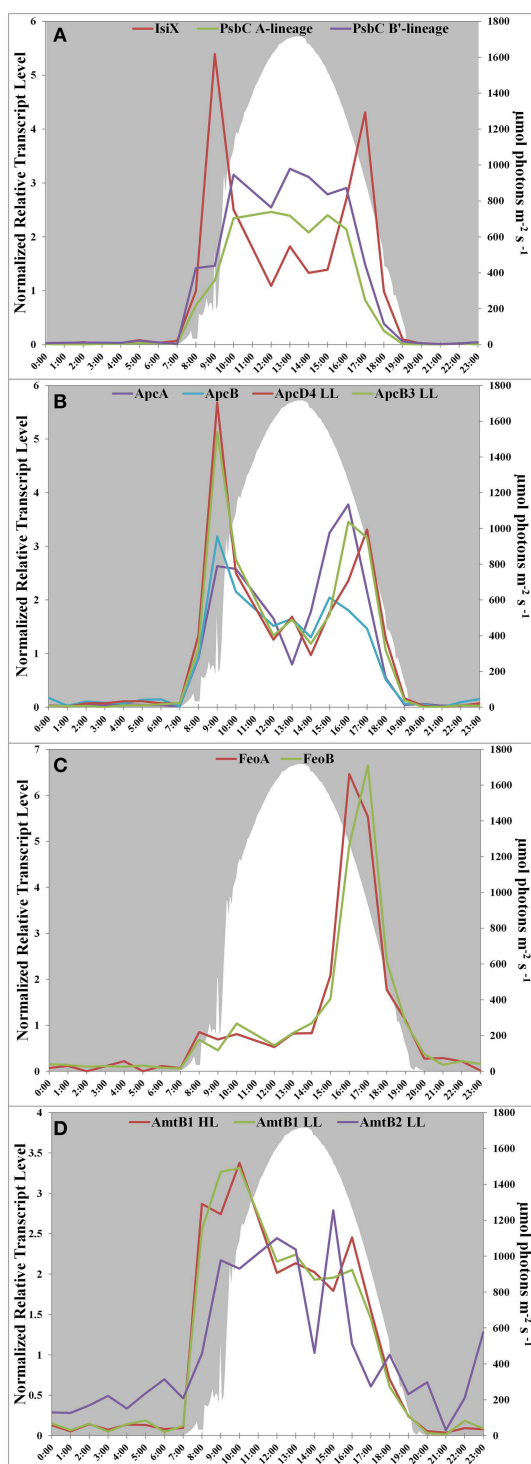
FIGURE 2 | Transcription patterns of transcripts encoding (A) PsbC in all A-lineage strains and IsiX, present in PE A4 and A14 strains, (B) ApcA/ApcB, present in all A-lineage strains, and ApcD4 LL/ApcB3 LL, present in PE A4 and A14 strains, (C) FeoA and FeoB, present in PE A4 and A14 strains, and (D) ammonium transporter genes in PE A1 (*amtB1*), PE A4/A14 (*amtB1* LL), and duplication (*amtB2* LL) in PE A4/A14. All panels show downwelling irradiance (μmol photons m^{-2}s^{-1}) measured at Mushroom Spring from September 11–12, 2009.

diurnally (Supplementary Figure S1). This copy is unlinked and divergent (∼68% amino acid identity) from the nitrate reductase of the *nirA-narB* gene cluster found in all of the strains. Although this gene is also found in the genomes of the PE A1-MS and A1-OS strains, it is disrupted by a mobile element gene and thus is not likely to be active.

Genes Found Only in High-light-adapted Strains and their Diel Transcription Patterns

The genomes of high-light-adapted strains PE A1-MS and A1-OS possess a copy of a carbonic anhydrase gene that has 68% amino acid identity to the zinc-dependent, gamma-class carbonic anhydrase found in *Thermosynechococcus* sp. NK55a. Carbonic anhydrase catalyzes the interconversion of carbon dioxide and bicarbonate by a reversible hydration reaction, and while the carbon-concentrating mechanism (CCM, present in all four strains) also includes a different carbonic anhydrase, cyanobacteria (Cannon et al., 2010), and other prokaryotes (Smith and Ferry, 2000) can have multiple copies of the genes and multiple classes of the enzyme that may play different functional roles in photosynthesis. The expression of this gene was too low to ascertain its transcription pattern confidently.

The urease cassette (Cluster 1 urease in Bhaya et al., 2007) found in the genome of the PE A1-OS strain is also found in the PE A1-MS strain genome, but these genes are not present in the genomes of the low-light adapted PE A4 and A14 strains. This urease cassette includes the genes that encode the larger alpha subunit UreC, smaller beta and gamma subunits UreB and UreA, which form the heterotrimeric urease enzyme, and UreDEFG accessory proteins that aid in assembly of the nickel metallocenter of the enzyme (Farrugia et al., 2013). All of the genes in the urease cassette have >90% identity to the urease genes found in *Thermus islandicus*, which indicates a possible horizontal gene transfer of this cassette to an ancestor common to the high-light-adapted strains but not the low-light-adapted strains.

The PE A1-OS and A1-MS strain genomes have a five-gene cluster annotated as a peptide/opine/nickel ABC transporter (PepT family), which includes a periplasmic substrate-binding protein, two permease subunits, and two ATP-binding protein genes. Additionally, the PE A1-OS and A1-MS strain genomes possess two components of a cystine ABC transporter, genes encoding the periplasmic cystine binding protein and the permease protein, as well as two genes, flanked by genes for mobile element proteins, that are annotated as succinate dehydrogenase flavoprotein subunit *sdhA* and omega-amino acid-pyruvate aminotransferase. Transcript abundances for all of these genes are higher during the light period and lower at night, similar to other genes that are expressed during the day (Supplementary Figure S2). Finally, along with the Type I and Type II CRISPR/*cas* arrays that are conserved among all four strains, the PE A1-MS genome contains a Type III CRISPR/*cas* array previously found in the PE A1-OS strain genome by Heidelberg et al. (2009). This is a unique CRISPR/*cas* array that is shared by *Roseiflexus* sp. RS-1, an anoxygenic photosynthetic organism that is also abundant in these microbial mat communities (Klatt et al., 2011). Although the amino acid similarities of the homologous genes are only 40–66% between

the two organisms, there are transposons flanking the array in the PE A1 strain genomes, which suggests a possible, if not recent, lateral gene transfer event in the mat (Heidelberg et al., 2009).

PsbA Allele and Diel Transcription Differences between High-light- and Low-light-adapted Strains

PsbA, also known as D1, is one of the core subunits of Photosystem II reaction center (Umena et al., 2011; Murray, 2012). The genome of the high-light-adapted PE A1-OS strain encodes four *psbA* genes, CYA_1274, CYA_1748, CYA_1811, and CYA_1894, while the B′ genome [JA-2-3B′a (2-13)] of *Synechococcus* has three *psbA* genes, designated CYB_0216, CYB_0371, and CYB_0433 (Bhaya et al., 2007). CYA_1274, CYA_1811, CYA_1849, CYB_0371, and CYB_0433 are nearly identical and differ by only 1 or 2 conserved amino acids. CYA_1748 and CYB_0216 are very similar to one another (94% identity, 96% similarity) but are only about 73% identical and 85% similar to the other PsbA sequences. These latter sequences have been called "rogue PsbA" sequences (rPsbA) by Murray (2012). Rogue PsbA sequences lack key functional residues and thus are not expected to support oxygen evolution by Photosystem II complexes that might contain them. The new strain genomes also possess multiple copies of the *psbA* gene: the PE A1-MS genome has four copies, which appear to be orthologous to those in the PE A1-OS genome, while the PE A4 and A14 genomes each have three *psbA* genes, two of which are identical to CYA_1274 and CYA_1849, as well as a copy of the *rpsbA* gene. The gene encoding rPsbA, CYA_1748, is sufficiently divergent to differentiate between the high-light- and low-light-adapted strains (the PE A4 and A14 alleles are 90 and 91% identical to the *rpsbA* gene in the PE A1-MS strain, respectively). Interestingly, the low-light- and high-light-adapted strains exhibited similar, but clearly temporally offset transcript abundance patterns (**Figure 3**), with the transcripts of low-light-adapted strains declining later in the morning and increasing earlier in the late afternoon. The transcript abundance for *rpsbA* (CYB_0216) of the B′-lineage *Synechococcus* also declines earlier in the morning than those of the low-light adapted strains, but they increase even later than transcripts of PE A1 strains.

Other Gene Content Differences among Strains

The genomes of strains PEs A1-MS, A4, and A14 encode the genes necessary to synthesize urea carboxylase, including urea carboxylase/allophanate hydrolase, two urea carboxylase-related aminomethyltransferases, three genes for a urea carboxylase-related ABC transporter, and a biotin-protein ligase gene (**Table 3**). The urea carboxylase cluster proteins have between 60 and 85% identity to proteins found in other cyanobacteria, so it is not clear if the cluster was acquired horizontally or vertically. Strains of PEs A1-OS, A1-MS, and A4 share a three-gene cassette for an ABC transporter for polar amino acids that is not found in the PE A14 strain (**Table 3**). This is a potentially interesting difference because aspartate and glutamate are the only two amino acids that are not taken up and metabolized by *Chloracidobacterium thermophilum*, which is co-localized with the low-light-adapted *Synechococcus* that occur deeper in the mat

(see Tank and Bryant, 2015). The PE A1-MS strain also has a gene cluster that is not found in the PE A1-OS strain and that consists of genes predicted to encode a bipolar DNA helicase, a Type I restriction-modification system DNA-methyltransferase subunit M, and the single-stranded exonuclease associated with Rad50/Mre11 complex (**Table 3**). The PE A14 strain possesses a PotABCD cassette for spermidine/putrescine transport that is not found in any of the other strains, which could provide an alternative nitrogen source for this strain. The proteins are most similar to PotABCD proteins in alpha and gamma proteobacteria, which may be indicative of a lateral gene transfer event. Additionally, the genome of the PE A14 strain encodes beta-carotene ketolase (CrtO), which is also encoded in the genome of the B′ strain JA-2-3B′a (2-13) (Bhaya et al., 2007). Keto-carotenoids provide better protection from reactive oxygen species than hydroxylated xanthophyll derivatives and are differently localized in membranes than other xanthophyll derivatives (Zhu et al., 2010). Further suggesting functional differences among these ecological species, transcripts were found for all of the strain-specific genes *in situ*.

Discussion

In this study we compared the genomes of strains from extremely closely related yet ecologically distinct PEs, each of which has a unique distribution along the vertical gradient at 60–63°C (Becraft et al., 2015) and differing light adaptations and acclimation responses corresponding to their vertical distributions (Nowack et al., 2015). Our aim was to discover the genetic bases for the physiological differences that cause these organisms to occupy different niches along the vertical gradient. Some of the most conspicuous differences are found between the high-light-adapted strains of PE A1 and the low-light-adapted strains of PEs A4 and A14. Becraft et al. (2015), the first paper of this series, showed that PE A1 predominates in the upper to middle part (0–760 μm deep) of the upper green layer of the mat, while PEs A4 and A14 are most abundant in deeper layers of the mat (640–720 μm and 640–960 μm, respectively). The difference in scalar irradiance received by the different populations is striking (see Figure 8C in Becraft et al., 2015); while members of PE A1 may experience up to 1250 μmol photons $m^{-2}s^{-1}$ scalar irradiance, PE A4 and A14 populations may only experience 50–75 μmol photons $m^{-2}s^{-1}$ at the peak irradiance level during a diel cycle. These ecophysiological differences are reflected in the gene contents of these organisms.

Gene content differences suggest different adaptations for the high-light- and low-light-adapted organisms. The low-light-adapted strains of PEs A4 and A14 possess a gene cluster with xenologous copies of *apcD4*, *apcB3*, and *isiX*, all of which are highly expressed *in situ*. Genes in this cassette are most likely responsible for the long-wave absorption and fluorescence emission features observed in those strains when grown at low irradiance, but are missing in high-light-adapted organisms, as reported in the second paper of this series (Nowack et al., 2015). This would be consistent with selection pressure to improve and expand light harvesting when the ambient light is strongly filtered by Chl *a* and phycobiliproteins by organisms in the upper regions
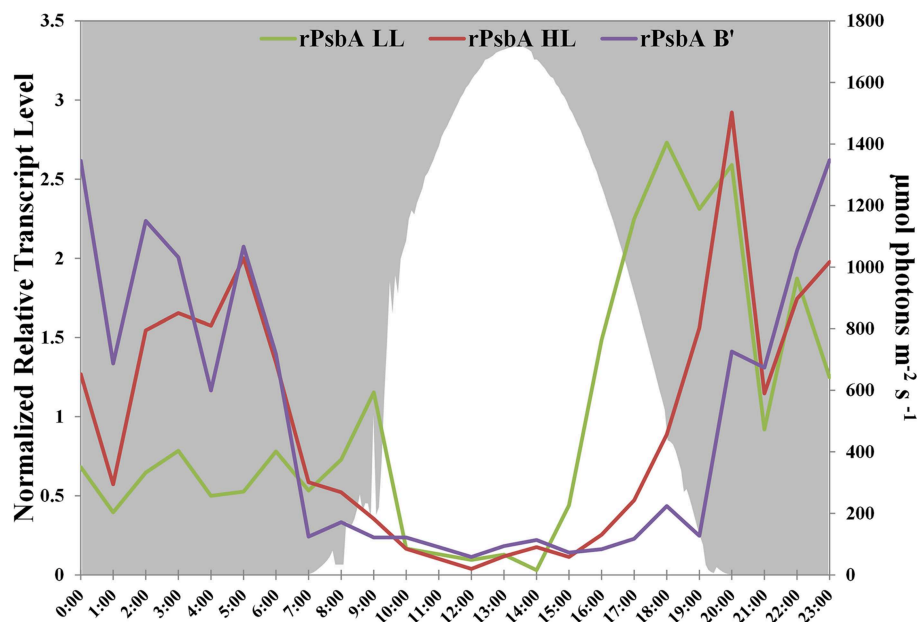
FIGURE 3 | rPsbA diel transcription patterns of alleles from the (i) B′ lineage, (ii) populations represented by high-light adapted strains (HL) of PE A1-OS and A1-MS, and (iii) populations represented by low-light adapted strains (LL) of PEs A4 and A14. Downwelling irradiance (μmol photons m^{-2}s^{-1}) measured at Mushroom Spring on September 11–12, 2009, is also shown.

of the mat and by the greater relative abundance of far-red light at increasing depth in the mat (see Figure 4 in Becraft et al., 2015). Some of this shift to the far-red would simply be due to greater penetration by light of longer wavelengths, which is less readily scattered. This gene cassette is also found in several other cyanobacteria (see Shih et al., 2013 Figure S5 CP43 phylogeny, members of clade CBPVI; and Gan et al., 2014b), including *Chlorogloeopsis* spp. PCC 6912 and PCC 9212, *Fischerella* sp. PCC 9605, *Chroococcidiopsis thermalis* PCC 7203, *Gloeocapsa* sp. PCC 7428, *Xenococcus* sp. PCC 7305, and *Leptolyngbya* sp. PCC 6406. These genes may encode a common adaptive mechanism among low-light-adapted cyanobacteria that are primarily found in benthic or terrestrial environments, by enabling them to acclimate to low irradiance conditions and/or to far-red light. Only three genes, *apcD4*, *apcB3*, and *isiX*, are required, which is far simpler than the FaRLiP response recently described by Gan et al. (2014a) that involves 17 genes and leads to changes to all three major photosynthetic complexes. Interestingly, several organisms that can perform FaRLiP (Gan et al., 2014a,b) also have this simpler system, which strongly suggests that the systems, at least in those cyanobacteria that have both capabilities, respond to different light cues.

The PE A4 and A14 strains additionally contain the *feoAB* genes for the ferrous iron transport system, which were initially described to be present in the metagenomes of the Mushroom Spring and Octopus Spring mats. The transcription pattern of the genes in the metatranscriptome (Liu et al., 2012 and **Figure 2C**) match the transcription pattern of *feoB* measured with q-RT-PCR over a diel cycle by Bhaya et al. (2007). Under alkaline conditions, ferrous iron is only present in the absence of oxygen, which

may occur more often in the deeper parts of the mat, away from the higher levels of oxygen in the upper part of the mat that are produced by *Synechococcus* populations experiencing higher irradiance levels and longer periods of exposure to light (Jensen et al., 2010). Interestingly, the *feoAB* genes discovered by Bhaya et al. (2007) were found on metagenomic clones that were most closely related to the B′-lineage, which may indicate the existence of low-light-adapted B′-lineage ecotypes as well as low-light-adapted A-lineage PEs. This might explain the inability of the B′-like strain studied by Kilian et al. (2007) to grow at high irradiance, if it contained only low-light-adapted, B′-lineage ecotypes.

In contrast to the low-light-adapted strains, the high-light-adapted strains PE A1-OS and A1-MS both contain an extra carbonic anhydrase gene, which may enhance growth under CO_2-limiting conditions when bicarbonate is present. The extra carbonic anhydrase may enhance conversion of bicarbonate to CO_2. CO_2 limitation caused by high rates of photosynthesis during peak irradiance has been indicated by an increase in pH when rates of oxygenic photosynthesis are high (Jensen et al., 2010). This observation led us to demonstrate that the growth rate of the PE A1-MS strain, but not a strain without the extra carbonic anhydrase gene, was increased by the addition of bicarbonate under carbon-limiting conditions (Supplementary Figure S3 and Supplementary Methods), which implies that this gene may provide increased fitness under such conditions. The high-light-adapted PE A1-OS and A1-MS strains also have unique genes involved in the TCA cycle (*sdhA*) and virus infection (Type III CRISPR/*cas* array), which may be indicative of uncharacterized environmental realities of

the high-light-adapted strains compared to the low-light-adapted strains.

Transcription patterns differ for genes associated with strains representative of different PEs. We were able to exploit the relatively high sequence divergence of the *rpsbA* gene to show that the transcription timing of this gene by low-light-adapted PE A4 and A14 populations found deepest in the mat green layer differed from that of the high-light-adapted PE A1 population residing above them. Specifically, transcription in PEs found deeper in the mat started earlier in the afternoon and ended later in the morning. Jensen et al. (2010) reported a similar transcription pattern for this gene in B′-like *Synechococcus*. Furthermore, by recruiting B′-like transcripts from the metatranscriptome, we were able to show that B′-like populations in the 60°C mat, which have been shown to predominate in the uppermost part of the mat green layer (see Figure 3 of Becraft et al., 2015 and Ramsing et al., 2000), express *rpsbA* genes even later in the afternoon and have declining transcript abundances for these genes even earlier in the morning. Similarly, Becraft et al. (2015) reported offsets in the timing of B′-like and A-like expression of photosynthesis and nitrogen fixation genes.

The function of rogue-PsbA in Photosystem II has not yet been established, but because this subunit is missing essential amino acid residues for the Mn_4CaO_5 cluster of the water oxidation center and has key differences in the binding pocket for quinone Q_B, it seems unlikely that Photosystem II complexes containing this protein can oxidize water (Murray, 2012). Considering that transcript abundance pattern for this gene is similar to those for nitrogen fixation genes (Figure 8B in Becraft et al., 2015), and that transcripts for "typical" *psbA* alleles increase rapidly as nitrogen fixation wanes and photosynthesis increases, we hypothesize that rPsbA subunits are involved in the oxidation of sulfide, which is present in the mats due to sulfate reduction during periods of anoxia (van der Meer et al., 2005; Dillon et al., 2007). Although *Synechococcus* lacks sulfide quinone reductase, which occurs in some cyanobacteria that oxidize sulfide to polysulfide (e.g., *Oscillatoria limnetica*; Arieli et al., 1991, 1994), most cyanobacteria that oxidize sulfide actually produce thiosulfate as the sole product in a reaction that has never been fully characterized biochemically (De Wit and van Gemerden, 1987; Rabenstein et al., 1995). We hypothesize that rPsbA is involved in the oxidation of sulfide to thiosulfate, and that this process could provide electrons for nitrogen fixation by nitrogenase, which would otherwise be inactivated by oxygen production if Photosystem II contained "typical" PsbA subunits. This scenario is further supported and is completely consistent with previous results suggesting that sulfide stimulated early morning incorporation of CO_2 into cyanobacterial lipids (van der Meer et al., 2005). Such a process would be expected to occur under anoxic conditions, which occur earlier in the afternoon in deeper portions of the mat (see Figure 8C in Becraft et al., 2015).

Additionally, we observed gene content differences among strains that might reflect alternative strategies for nitrogen metabolism. For instance, both PE A1 strains are capable of urea degradation with urease, while strains of PEs A1-MS, A4, and A14 have urea carboxylase. Urea degradation with urea carboxylase involves two separate reactions and is ATP-dependent, while urease involves only one reaction and is not ATP-dependent, but requires nickel for the enzyme metallocenter (Sakamoto and Bryant, 2001; Solomon et al., 2010; Farrugia et al., 2013). Rates of urea uptake are usually higher than for nitrate or nitrite, even when the concentration of these oxidized nitrogen sources is higher, and urea is preferable in CO_2-limited environments because CO_2 is a useful by-product of urea assimilation (Solomon et al., 2010). The peptide/opine/nickel transport cassette in PE A1-OS and A1-MS strains may provide the nickel for the urease enzyme when available, or it might be involved in scavenging of environmental peptides or opines as a source of both nitrogen and organic carbon. Similarly, the cystine transport genes in PE A1-OS and A1-MS strains, the polar amino acid transport cassette in PE A1-OS, A1-MS, and A4 strains, and the PotABCD spermidine/putrescine transporter in the PE A14 strain are all transcribed *in situ*, and all transport possible sources of nitrogen into the cells. Other gene content differences among strains may indicate differences in organic carbon use (the putative MalK transport cassette in strains of PEs A4 and A14) and DNA protection and repair (bipolar DNA helicase and single-strand exonuclease in the PE A1-MS strain and beta-carotene ketolase in the PE A14 strain). Some of these gene content differences may help to explain the niche differentiation between the two low-light-adapted strains of PEs A4 and A14. Although both grow faster at lower irradiances than the PE A1 strains and are thus characterized as low-light-adapted, they do have different patterns of growth relative to light intensity (Nowack et al., 2015) and different vertical distributions in the mat (Becraft et al., 2015). The PE A4 distribution is maximal in the lower-middle part of the mat upper green layer, while PE A14 is maximal at the greatest depths where irradiance is most attenuated.

This three-paper series was designed to address the issue of the molecular dimension of microbial species. Since Woese and Fox (1977) used the highly conserved 16S ribosomal RNA sequence to estimate phylogenetic relatedness among organisms to reveal inaccuracies of traditional classification methods [e.g., the complete oversight of the domain Archaea (Balch et al., 1977)], the extensive use of this approach has led to a somewhat arbitrary molecular demarcation of microbial species that is widely accepted and used by many microbiologists. Molecular cutoffs were created by observing the sequence divergence among strains of classically named species (e.g., Seki et al., 1978 within the genus *Bacillus*): that a >3% divergence of the 16S rRNA locus between two organisms (Stackebrandt and Goebel, 1994) or, more recently, >1% divergence at the 16S rRNA locus (Stackebrandt and Ebers, 2006) is required to consider that the two strains belong to different species. Using the highly resolving locus *psaA*, we have (i) predicted the existence of different putative ecological species within traditional 16S rRNA-defined species using a theory-based model (Becraft et al., 2011), (ii) shown that they are ecologically distinct through differences in distribution, (iii) shown that many contain ecologically homogeneous members (Becraft et al., 2015), (iv) shown that strains representative of different PEs have different adaptations and acclimation responses to light (Nowack et al., 2015), and (v)

through comparative genomic analysis of these strains, shown that strains of the different PEs contain differences in gene content and gene alleles which appear to underlie the adaptations and acclimation responses of each PE to their distinct ecological niches. Comparative analyses of multiple strains of different PEs, including analyses of genes under selection within and between ecotypes, will provide further evidence for the ecological differentiation among PEs and will demonstrate whether the adaptations and acclimative responses of these strains are typical of members of a PE. These results, along with the differences in the timing of gene expression by different PEs located in distinct niches, demonstrate that extremely closely related ecologically adapted populations, which may in fact be true ecological species, matter in microbial communities.

Author Contributions

MO cultivated strains, performed molecular procedures, performed genomic analyses, and wrote the manuscript. SN cultivated strains and performed growth experiments. JW and EB assisted in analyzing sequence data. EB assisted with molecular procedures. KL, AL, JM, and WS assisted in genome sequencing and assembly. DR, FC, and DB were acting Co-PIs for the Joint Genome Institute Community sequencing project. FC and DB discussed results and edited the manuscript. DW was acting PI for the Joint Genome Institute Community sequencing project, assisted with experimental design, discussed the results, and edited the manuscript at all stages.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.00604

References

Allewalt, J. P., Bateson, M. M., Revsbech, N. P., Slack, K., and Ward, D. M. (2006). Effects of temperature and light on growth of and photosynthesis by *Synechococcus* isolates typical of those predominating in the Octopus Spring microbial mat community of Yellowstone National Park. *Appl. Environ. Microbiol.* 72, 544–550. doi: 10.1128/AEM.72.1.544-550.2006

Arieli, B., Padan, E., and Shahak, Y. (1991). Sulfide-induced sulfide-quinone reductase activity inthylakoids of *Oscillatoria limnetica*. *J. Biol. Chem.* 266, 104–111.

Arieli, B., Shahak, Y., Taglicht, D., Hauska, G., and Padan, E. (1994). Purification and characterization of sulfide-quinone reductase, a novel enzyme driving anoxygenic photosynthesis in *Oscillatoria limnetica*. *J. Biol. Chem.* 269, 5705–5711.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75

Balch, W. E., Magrum, L. J., Fox, G. E., Wolfe, R. S., and Woese, C. R. (1977). An ancient divergence among the Bacteria. *J. Mol. Evol.* 9, 305–311. doi: 10.1007/BF01796092

Becraft, E. D. (2014). *The Ecology and Evolution of Thermophilic* Synechococcus *Species*. Ph.D. dissertation, Montana State University, Bozeman, MT.

Becraft, E. D., Cohan, F. M., Kühl, M., Jensen, S. I., and Ward, D. M. (2011). Fine-scale distribution patterns of *Synechococcus* ecological diversity in microbial mats of Mushroom Spring, Yellowstone National Park. *Appl. Environ. Microbiol.* 77, 7689–7697. doi: 10.1128/AEM.05927-11

Becraft, E. D., Wood, J. M., Rusch, D. B., Kühl, M., Jensen, S. I., Bryant, D. A., et al. (2015). The molecular dimension of microbial species: 1. Ecological distinctions among, and homogeneity within, putative ecotypes of *Synechococcus* inhabiting the cyanobacterial mat of Mushroom Spring, Yellowstone National Park. *Front. Microbiol.* 6:590. doi: 10.3389/fmicb.2015.00590

Bhaya, D., Grossman, A. R., Steunou, A. S., Khuri, N., Cohan, F. M., Hamamura, N., et al. (2007). Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. *ISME J.* 1, 703–713. doi: 10.1038/ismej.2007.46

Brock, T. D. (1978). *Thermophilic Microorganisms and Life at High Temperatures*. New York, NY: Springer.

Cannon, G. C., Heinhorst, S., and Kerfeld, C. A. (2010). Carboxysomal carbonic anhydrases: structure and role in microbial CO_2 fixation. *Biochim. Biophys. Acta* 1804, 382–392. doi: 10.1016/j.bbapap.2009.09.026

Chen, K., Wallis, J. W., McLellan, M. D., Larson, D. E., Kalicki, J. M., Pohl, C. S., et al. (2009). BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat. Methods* 6, 677–681. doi: 10.1038/nmeth.1363

Coleman, M. L., Sullivan, M. B., Martiny, A. C., Steglich, C., Barry, K., DeLong, E. F., et al. (2006). Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* 311, 1768–1770. doi: 10.1126/science.1122050

De Wit, R., and van Gemerden, H. (1987). Oxidation of sulfide to thiosulfate by *Microcoleus chtonoplastes*. *FEMS Microbiol. Lett.* 45, 7–13. doi: 10.1111/j.1574-6968.1987.tb02332.x

Dillon, J. G., Fishbain, S., Miller, S. R., Bebout, B. M., Habicht, K. S., Webb, S. M., et al. (2007). High rates of sulfate reduction in a low-sulfate hot spring microbial mat are driven by a low level of diversity of sulfate-respiring microorganisms. *Appl. Environ. Microbiol.* 73, 5218–5226. doi: 10.1128/AEM.00357-07

Farrugia, M. A., Macomber, L., and Hausinger, R. P. (2013). Biosynthesis of the urease metallocenter. *J. Biol. Chem.* 288, 13178–13185. doi: 10.1074/jbc.R112.446526

Ferris, M. J., Kühl, M., Wieland, A., and Ward, D. M. (2003). Cyanobacterial ecotypes in different optical microenvironments of a 68°C hot spring mat community revealed by 16S-23S rRNA internal transcribed spacer region variation. *Appl. Environ. Microbiol.* 69, 2893–2898. doi: 10.1128/AEM.69.5.2893-2898.2003

Ferris, M. J., and Ward, D. M. (1997). Seasonal distributions of dominant 16S rRNA-defined populations in a hot spring microbial mat examined

by denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* 63, 1375–1381.

Gan, F., Shen, G., and Bryant, D. A. (2014b). Occurrence of far-red light photoacclimation (FaRLiP) in diverse cyanobacteria. *Life* 5, 4–24. doi: 10.3390/life5010004

Gan, F., Zhang, S., Rockwell, N. C., Martin, S. S., Lagarias, J. C., and Bryant, D. A. (2014a). Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light. *Science* 345, 1312–1317. doi: 10.1126/science.1256963

Gnerre, S., MacCallum, I., Przybylski, D., Ribeiro, F. J., Burton, J. N., Walker, B. J., et al. (2011). High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1513–1518. doi: 10.1073/pnas.1017351108

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijs.0.64483-0

Heidelberg, J. F., Nelson, W. C., Schoenfeld, T., and Bhaya, D. (2009). Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLoS ONE* 4:e4169. doi: 10.1371/journal.pone.0004169

Inskeep, W. P., Jay, Z. J., Tringe, S. G., Herrgård, M. J., Rusch, D. B., Committee, Y. M. P. S., et al. (2013). The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem. *Front. Microbiol.* 4:67. doi: 10.3389/fmicb.2013.00067

Jensen, S. I., Steunou, A. S., Bhaya, D., Kühl, M., and Grossman, A. R. (2010). *In situ* dynamics of O_2, pH and cyanobacterial transcripts associated with CCM, photosynthesis and detoxification of ROS. *ISME J.* 5, 317–328. doi: 10.1038/ismej.2010.131

Kettler, G. C., Martiny, A. C., Huang, K., Zucker, J., Coleman, M. L., Rodrigue, S., et al. (2007). Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. *PLoS Genet.* 3:e231. doi: 10.1371/journal.pgen.0030231

Kilian, O., Steunou, A. S., Fazeli, F., Bailey, S., Bhaya, D., and Grossman, A. R. (2007). Responses of a thermophilic *Synechococcus* isolate from the microbial mat of Octopus Spring to light. *Appl. Environ. Microbiol.* 73, 4268–4278. doi: 10.1128/AEM.00201-07

Klatt, C. G., Wood, J. M., Rusch, D. B., Bateson, M. M., Hamamura, N., Heidelberg, J. F., et al. (2011). Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. *ISME J.* 5, 1262–1278. doi: 10.1038/ismej.2011.73

Koeppel, A., Perry, E. B., Sikorski, J., Krizanc, D., Warner, A., Ward, D. M., et al. (2008). Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2504–2509. doi: 10.1073/pnas.0712205105

Kouril, R., Arteni, A. A., Lax, J., Yeremenko, N., D'Haene, S., Rögner, M., et al. (2005). Structure and functional role of supercomplexes of IsiA and Photosystem I in cyanobacterial photosynthesis. *FEBS Lett.* 579, 3253–3257. doi: 10.1016/j.febslet.2005.03.051

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Liu, Z., Klatt, C. G., Ludwig, M., Rusch, D. B., Jensen, S. I., Kühl, M., et al. (2012). 'Candidatus Thermochlorobacter aerophilum:'an aerobic chlorophotoheterotrophic member of the phylum Chlorobi defined by metagenomics and metatranscriptomics. *ISME J.* 6, 1869–1882. doi: 10.1038/ismej.2012.24

Liu, Z., Klatt, C. G., Wood, J. M., Rusch, D. B., Ludwig, M., Wittekindt, N., et al. (2011). Metatranscriptomic analyses of chlorophototrophs of a hot-spring microbial mat. *ISME J.* 5, 1279–1290. doi: 10.1038/ismej.2011.37

Melendrez, M. C., Lange, R. K., Cohan, F. M., and Ward, D. M. (2011). Influence of molecular resolution on sequence-based discovery of ecological diversity among *Synechococcus* populations in an alkaline siliceous hot spring microbial mat. *Appl. Environ. Microbiol.* 77, 1359–1367. doi: 10.1128/AEM.02032-10

Miller, S. R., Wingard, C. E., and Castenholz, R. W. (1998). Effects of visible light and UV radiation on photosynthesis in a population of a hot spring cyanobacterium, a *Synechococcus* sp., subjected to high-temperature stress. *Appl. Environ. Microbiol.* 64, 3893–3899.

Murray, J. W. (2012). Sequence variation at the oxygen-evolving centre of photosystem II: a new class of 'rogue' cyanobacterial D1 proteins. *Photosyn. Res.* 110, 177–184. doi: 10.1007/s11120-011-9714-5

Murray, J. W., Duncan, J., and Barber, J. (2006). CP43-like chlorophyll binding proteins: structural and evolutionary implications. *Trends Plant Sci.* 11, 152–158. doi: 10.1016/j.tplants.2006.01.007

Musa-Aziz, R., Chen, L. M., Pelletier, M. F., and Boron, W. F. (2009). Relative CO_2/NH_3 selectivities of AQP1, AQP4, AQP5, AmtB, and RhAG. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5406–5411. doi: 10.1073/pnas.0813231106

Nowack, S. (2014). *Niche Character in a Temporally Varying Environment*. Ph.D. dissertation, Montana State University, Bozeman.

Nowack, S., Olsen, M. T., Schaible, G. A., Becraft, E. D., Shen, G., Klapper, I., et al. (2015). The molecular dimension of microbial species: 2. *Synechococcus* strains representative of putative ecotypes inhabiting different depths in the Mushroom Spring microbial mat exhibit different adaptive and acclimative responses to light. *Front. Microbiol.* 6:626. doi: 10.3389/fmicb.2015.00626

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226

Peary, J. A., and Castenholz, R. W. (1964). Temperature strains of a thermophilic blue-green alga. *Nature* 202, 720–721.

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077

Rabenstein, A., Rethmeier, J., and Fischer, U. (1995). Sulphite as intermediate sulfur compound in anaerobic sulphide oxidation to thiosulphate by marine cyanobacteria. *Z. Naturforsch.* 50c, 769–774.

Ramsing, N. B., Ferris, M. J., and Ward, D. M. (2000). Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter-thick photic zone of a hot spring cyanobacterial mat. *Appl. Environ. Microbiol.* 66, 1038–1049. doi: 10.1128/AEM.66.3.1038-1049.2000

Rocap, G., Larimer, F. W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N. A., et al. (2003). Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* 424, 1042–1047. doi: 10.1038/nature01947

Sakamoto, T., and Bryant, D. A. (2001). Requirement of nickel as an essential micronutrient for the utilization of urea in the marine cyanobacterium *Synechococcus* sp. PCC 7002. *Microbes Environ.* 16, 177–184. doi: 10.1264/jsme2.2001.177

Seki, T., Chung, C., Mikami, H., and Oshima, Y. (1978). Deoxyribonucleic acid homology and taxonomy of the genus *Bacillus*. *Int. J. Syst. Bacteriol.* 28, 182–189. doi: 10.1099/00207713-28-2-182

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., et al. (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1053–1058. doi: 10.1073/pnas.1217107110

Sidler, W. A. (2004). "Phycobilisome and phycobiliproteins structures," in *The Molecular Biology of Cyanobacteria, Advances in Photosynthesis and Respiration,* Vol. 1, ed D. A. Bryant (Dordrecht: Kluwer Academic), 139–216.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high–quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75

Smith, K. S., and Ferry, J. G. (2000). Prokaryotic carbonic anhydrases. *FEMS Microbiol. Rev.* 24, 335–366. doi: 10.1111/j.1574-6976.2000.tb00546.x

Solomon, C. M., Collier, J. L., Berg, G. M., and Glibert, P. M. (2010). Role of urea in microbial metabolism in aquatic systems: a biochemical and molecular review. *Aquat. Microb. Ecol.* 59, 67–88. doi: 10.3354/ame01390

Stackebrandt, E., and Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiol. Today* 33, 152–155.

Stackebrandt, E., and Goebel, B. M. (1994). Taxonomic note: a place for DNA:DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol.* 44, 846–849. doi: 10.1099/00207713-44-4-846

Steunou, A. S., Bhaya, D., Bateson, M. M., Melendrez, M. C., Ward, D. M., Brecht, E., et al. (2006). *In situ* analysis of nitrogen fixation and metabolic switching in

unicellular thermophilic cyanobacteria inhabiting hot spring microbial mats. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2398–2403. doi: 10.1073/pnas.0507513103

Steunou, A. S., Jensen, S. I., Brecht, E., Becraft, E. D., Bateson, M. M., Kilian, O., et al. (2008). Regulation of *nif* gene expression and the energetics of N$_2$ fixation over the diel cycle in a hot spring microbial mat. *ISME J.* 2, 364–378. doi: 10.1038/ismej.2007.117

Straus, N. A. (1994). "Iron deprivation: physiology and gene regulation," in *The Molecular Biology of Cyanobacteria*, ed D. A. Bryant (Dordrecht: Kluwer Academic), 731–750.

Tank, M., and Bryant, D. A. (2015). Nutrient requirements and growth physiology of the photoheterotrophic Acidobacterium, *Chloracidobacterium thermophilum*. *Front. Microbiol.* 6:226. doi: 10.3389/fmicb.2015.00226

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Ciufo, S., and Li, W. (2013). *Prokaryotic Genome Annotation Pipeline*. (Bethesda, MD: National Center for Biotechnology Information (US)).

Umena, Y., Kawakami, K., Shen, J.-R., and Kamiya, N. (2011). Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* 473, 55–60. doi: 10.1038/nature09913

van der Meer, M. T. J., Schouten, S., Bateson, M. M., Nübel, U., Wieland, A., Kühl, M., et al. (2005). Diel variations in carbon metabolism by green nonsulfur-like bacteria in alkaline silicious hot spring microbial mats from Yellowstone National Park, U.S.A. *Appl. Environ. Microbiol.* 71, 3978–3986. doi: 10.1128/AEM.71.7.3978-3986.2005

Wang, Z., and Wu, M. (2013). A phylum-level bacterial phylogenetic marker database. *Mol. Biol. Evol.* 30, 1258–1262. doi: 10.1093/molbev/mst059

Ward, D. M., Bateson, M. M., Ferris, M. J., Kühl, M., Wieland, A., Koeppel, A., et al. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 1997–2008. doi: 10.1098/rstb.2006.1919

Ward, D. M., Castenholz, R. W., and Miller, S. R. (2012). "Cyanobacteria in geothermal habitats," in *Ecology of Cyanobacteria, 2nd Edn.*, ed B. A. Whitton (Heidleberg: Springer Science+Business Media), 39–63.

Ward, D. M., and Cohan, F. M. (2005). "Microbial diversity in hot spring cyanobacterial mats: pattern and prediction," in *Geothermal Biology and Geochemistry in Yellowstone National Park*, eds B. Inskeep and T. McDermott (Bozeman, MT: Thermal Biology Institute, Montana State University), 185–201.

Woese, C. F., and Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5088–5090. doi: 10.1073/pnas.74.11.5088

Zhu, Y., Graham, J. E., Ludwig, M., Xiong, W., Alvey, R. M., Shen, G., et al. (2010). Roles of xanthophyll carotenoids in protection against photoinhibition and oxidative stress in the cyanobacterium *Synechococcus sp.* strain PCC 7002. *Arch. Biochem. Biophys.* 504, 86–99. doi: 10.1016/j.abb.2010.07.007

frontiers
in Microbiology

Time dynamics of the *Bacillus cereus* exoproteome are shaped by cellular oxidation

Jean-Paul Madeira[1, 2, 3], Béatrice Alpha-Bazin[3], Jean Armengaud[3] and Catherine Duport[1, 2]*

[1] UMR408, Sécurité et Qualité des Produits d'Origine Végétale, Université d'Avignon, Avignon, France, [2] INRA, UMR408, Sécurité et Qualité des Produits d' Origine Végétale, Avignon, France, [3] Commissariat à l'énergie Atomique et aux Énergies Alternatives (CEA), Direction des Sciences du Vivant (DSV), IBEB, Li2D, Bagnols sur Cèze, France

At low density, *Bacillus cereus* cells release a large variety of proteins into the extracellular medium when cultivated in pH-regulated, glucose-containing minimal medium, either in the presence or absence of oxygen. The majority of these exoproteins are putative virulence factors, including toxin-related proteins. Here, *B. cereus* exoproteome time courses were monitored by nanoLC-MS/MS under low-oxidoreduction potential (ORP) anaerobiosis, high-ORP anaerobiosis, and aerobiosis, with a specific focus on oxidative-induced post-translational modifications of methionine residues. Principal component analysis (PCA) of the exoproteome dynamics indicated that toxin-related proteins were the most representative of the exoproteome changes, both in terms of protein abundance and their methionine sulfoxide (Met(O)) content. PCA also revealed an interesting interconnection between toxin-, metabolism-, and oxidative stress–related proteins, suggesting that the abundance level of toxin-related proteins, and their Met(O) content in the *B. cereus* exoproteome, reflected the cellular oxidation under both aerobiosis and anaerobiosis.

Keywords: exoproteome, *Bacillus cereus*, shotgun proteomics, methionine oxidation, toxins

Introduction

The gram-positive, motile bacterium, *Bacillus cereus*, is a well-known agent of gastrointestinal (GI) tract infection (Stenfors Arnesen et al., 2008; Bishop et al., 2010). The critical step of infection occurs in the small intestine, where *B. cereus* encounters carbohydrate starvation conditions and changing oxygenation and oxidoreduction potential (ORP) conditions (Guyton, 1977; Moriarty-Craige and Jones, 2004; Fabich et al., 2008; Marteyn et al., 2010). During the course of infection, the survival and growth of *B. cereus* depend on the secretion and release into the extracellular compartment of multiple proteins (Gilois et al., 2007; Gohar et al., 2008). The *B. cereus* ATCC 14579 exoproteome, which comprises the secreted proteins and all the other released proteins found in the pathogen's extracellular surroundings (Armengaud et al., 2012), was recently established for cells grown under conditions considered to mimic those encountered in the human intestine, i.e., low-ORP anoxic conditions, high-ORP anoxic conditions, and oxic conditions, in pH-regulated culture using glucose as the sole carbohydrate source (Clair et al., 2010). The *B. cereus* exoproteome is dominated by toxin-related proteins (∼35% of the exoproteome, as estimated by spectral count) and degradative enzymes plus adhesins (∼35% of the exoproteome), which are all recognized as major virulence factors (Stenfors Arnesen et al., 2008; Ingmer and Brondsted, 2009; Kamar et al., 2013;

Ramarao and Sanchis, 2013). The other components of the *B. cereus* exoproteome comprise components of the flagellar apparatus (∼15% of the exoproteome), as well as an important number of proteins that lack export signal sequences, accounting for 15% of the exoproteome. These proteins, found more abundantly in the cytoplasm, include metabolic enzymes (mainly glycolytic enzymes), translation-related proteins, molecular chaperones, and antioxidant enzymes such as catalase, hydroperoxide reductase, and superoxide dismutase. Several studies have reported the moonlighting activities of these proteins, which are involved in bacterial virulence. Most enzymes in the glycolytic pathway, tricarboxylic acid (TCA) cycle and glyoxylate cycle have adhesive properties that aid in interacting with the host extracellular matrix. The most common moonlighting activity of bacterial molecular chaperones is to activate (or inhibit) mononuclear phagocyte cytokine synthesis. Antioxidants produced by *Mycobacterium bovis* suppress host immune response (Sadagopal et al., 2009; Vellasamy et al., 2009; Henderson and Martin, 2011).

B. cereus adjusts its primary metabolism to grow efficiently under aerobic respiratory and anaerobic fermentative conditions and to adapt to low-ORP conditions (Duport et al., 2006; Clair et al., 2012). In addition, as for all other bacteria, *B. cereus* undergoes a major metabolic switch from primary metabolism (exponential growth) to secondary metabolism (stationary phase) in response to nutrient starvation or oxidative stress (Nieselt et al., 2010). Aerobic respiration relies on dioxygen to drive ATP production via the respiratory chain (Duport et al., 2006). One caveat is that this process is accompanied by a major production of reactive oxygen species (ROS) (Gonzalez-Flecha and Demple, 1995; Brynildsen et al., 2013; Imlay, 2013). In addition to the respiratory chain, endogenous ROS can be generated in response to starvation (nutrient stress) as a secondary stress (Mols and Abee, 2011). Under anaerobiosis, *B. cereus* catabolizes glucose-using, fermentative pathways, which are not recognized as high-ROS-producing pathways under normal conditions. However, low-ORP conditions can induce ROS production in response to reductive stress (Clair et al., 2012). Bacteria use a large spectrum of ROS scavenging systems, including low-molecular-weight molecules, metabolites, and antioxidant enzymes, to maintain ROS at non-toxic levels and to prevent macromolecule damage (Chi et al., 2011; Mailloux et al., 2011). Amino acid residues in proteins represent one of the major targets of ROS and cellular oxidants. The two amino acids that are the most prone to oxidative attack by ROS are cysteine and methionine (Met), both of which contain susceptible sulfur atoms. However, Met residues are the most susceptible to oxidation by almost all forms of ROS (Vogt, 1995; Stadtman et al., 2005). Met oxidation produces a stable product, methionine sulfoxide, Met(O), which can be detected readily by mass spectrometry through a mass increase of 15.9949 atomic mass units. Thus, Met oxidation might serve as a sensitive marker for proteins oxidized by ROS.

The objective of the present study was to define the exoproteome time dynamics of *B. cereus* grown in three ORP conditions, and to assess by tandem mass spectrometry the oxidation level of the secreted proteins, which should be correlated with the cellular oxidation level. For this purpose, we collected *B. cereus* supernatant at three points of the time-growth curve, i.e., during early exponential growth phase (EE), at the late exponential growth phase (LE) signifying the transition between exponential and stationary phases, and during the stationary phase (S). This was performed for cells grown under aerobiosis, as well as under high- and low-ORP anaerobiosis. Time-course changes in terms of exoprotein abundance level and the Met(O) peptide content of exoproteins were assessed by high-throughput nanoLC-MS/MS (Clair et al., 2010). The repertoire of experimentally confirmed exoproteins of *B. cereus* presented here is the largest ever reported, and more interestingly provides new insights into the interplay between toxin-related protein secretion and intracellular ROS production.

Materials and Methods

B. cereus Growth Conditions

B. cereus ATCC 14579 cells were grown in a batch bioreactor on MOD medium supplemented with 30 mM glucose as the carbon source (Rosenfeld et al., 2005) and buffered at pH 7.2 with 2 M KOH. The bioreactor was an autoclavable 3-liter glass BioFlo®/CelliGen®115 (New Brunswick Scientific) with a working volume of 2 liters. It was equipped with a polarographic oxygen electrode (Mettler Toledo), a pH electrode (Mettler Toledo), and a redox-combined electrode (AgCl, Mettler Toledo). Sterile gas was fed through the culture at a constant flow set to 20 mL/h. For oxic conditions, oxygen saturation was maintained at 100% by automatic adjustment of the stirring speed. For anoxic conditions, a dissolved oxygen tension value (pO_2) of 0% was obtained with a constant flow of pure nitrogen (high- ORP condition) or hydrogen gas (low-ORP condition). Each bioreactor was inoculated with a subculture grown for 8 h (exponential growth phase) in glucose-containing MOD medium under aerobiosis or anaerobiosis. Cells from the inocula were harvested by centrifugation ($7000 \times g$ for 5 min at room temperature), washed in fresh medium, and then diluted to achieve an initial optical culture density at 600 nm of 0.02. Batch cultures were carried out at 37°C under a 300 rpm agitation speed.

Exoproteome Preparations and Trypsin In-Gel Proteolysis

For each of the three growth conditions, three independent growth cultures in a fermenter were carried out, resulting in biological samples in triplicate for each time point. Optical density, ORP, and pO_2 were monitored every 30 min during the bacterial growth. The growth rate was determined from the absorbance data. A 200-mL sample of the culture was systematically taken at the exponential, transition, and stationary phases for the nine bioreactor cultures. Cell pellets and extracellular media were separated by centrifugation at 10,000 × g for 10 min at 4°C. The extracellular media were successively filtered through acetate membrane filters (Sartorius) with pore sizes of 0.85, 0.45, and 0.20 μm, respectively. Proteins from the 27 samples were precipitated by adding 10 mL trichloroacetic acid solution at 100% (w/v) to 40 mL filtered solution. The precipitated material was recovered after overnight incubation at 4°C by centrifugation at 7000 × g for 15 min at 4°C, and the extracellular proteins in the

resulting pellet were then dissolved in 100 μL NUPAGE® LDS (Lithium dodecyl sulfate) sample buffer 1X (Invitrogen) supplemented with β-mercaptoethanol. Samples were boiled for 5 min at 95°C, sonicated for 5 × 5 s in a transonic 780H sonicator and loaded on NuPAGE® Novex 4–12% Bis-Tris gels (Invitrogen) that were run for a short 5-min migration at 200 V using NuPAGE® MES supplemented with NuPAGEantioxidant as the running buffer (Hartmann and Armengaud, 2014). This avoids any artifactual protein oxidation. Gels were stained with Simply Blue SafeStain, a ready-to-use Coomassie G-250 stain from Invitrogen. After overnight destaining, the single band of each gel lane was cut and divided into 2 fractions, each corresponding to a $3 \times 4 \, mm^2$ polyacrylamide band. The 54 resulting polyacrylamide gel pieces were processed for further destaining, reduction and iodoacetamide treatments, and in-gel proteolysis with trypsin (Roche) in the presence of ProteaseMax additive (Promega), as previously described (De Groot et al., 2009; Clair et al., 2010). The two digests obtained from the same sample were pooled as a single peptide mixture. Exponential phase samples were injected without being diluted, due to their lower protein content, while the samples collected at the transition and stationary phases were diluted 1:50 in 0.1% trifluoroacetic acid prior to nanoLC-MS/MS analysis.

Tandem Mass Spectrometry

NanoLC-MS/MS experiments were performed using an LTQ-Orbitrap XL hybrid mass spectrometer (ThermoFisher) coupled to an UltiMate 3000 nRSLC system (Dionex ThermoFisher), in similar conditions to those previously described (Dedieu et al., 2011). Peptide mixtures were loaded and desalted on-line on a reverse-phase precolumn (Acclaim PepMap 100 C18, 5 μm bead size, 100 Å pore size, 300 μm i.d × 5 mm (Dionex-ThermoFisher). Peptides were then resolved on a Dionex nanoscale Acclaim Pepmap100 C18 capillary column (3 μm bead size, 100 Å pore size, 75 μm i.d. × 15 cm) at a flow rate of 0.3 μL/min using a 90 min. gradient from 4 to 40% solvent B (0.1% HCOOH/100% CH₃CN) prior to injection into the mass spectrometer. Solvent A was 0.1% HCOOH/100% H₂O. Full-scan mass spectra were measured from m/z 300 to 1800 with the LTQ-Orbitrap XL mass spectrometer in data-dependent mode using TOP3 strategy. In brief, a scan cycle was initiated with a full scan of high mass accuracy in the Orbitrap, followed by MS/MS scans in the linear ion trap on the three most abundant precursor ions, with 60 s dynamic exclusion of previously selected ions.

Protein Identification

Peak lists from the tandem mass spectrometry raw data were generated with the MASCOT DAEMON software (version 2.3.2) from Matrix Science using the extract_msn.exe data import filter from the Xcalibur FT package (version 2.0.7) proposed by ThermoFisher. Data import filter options were set as follows: at 400 (minimum mass), 5000 (maximum mass), 0 (grouping tolerance), 0 (intermediate scans), and 1000 (threshold). Using the MASCOT search engine (version 2.3.02) from Matrix Science, we searched all MS/MS spectra against an in-house polypeptide sequence database containing the sequences of all annotated proteins encoded by the *B.*

cereus ATCC 14579 chromosome (NC_004722) and plasmid, pBClin15 (NC_004721), supplemented with 44 new proteins discovered by a previous proteogenomic analysis (unpublished data). This database comprises 5299 polypeptide sequences, totaling 1,464,675 amino acids. Searches for tryptic peptides were performed with the following parameters: full trypsin specificity, a mass tolerance of 5 ppm on the parent ion and 0.6 Da on the MS/MS, static modifications of carboxyamidomethylated Cys (+57.0215), and dynamic modifications of oxidized Met (+15.9949). The maximum number of missed cleavages was set at 2. All peptide matches with a peptide score below a *p*-value of

FIGURE 1 | Growth curves of *B. cereus* in pH-regulated batch culture under aerobiosis, high-ORP anaerobiosis, and low-ORP anaerobiosis. The results from biological triplicate curves are indicated. Optical densities (OD₆₀₀ₙₘ) and ORP values are shown in blue and red, respectively. Samples for exoproteomic analyses were taken during the exponential growth phase (EE), late growth phase (LE), and stationary growth phase (S), as indicated by black arrows.
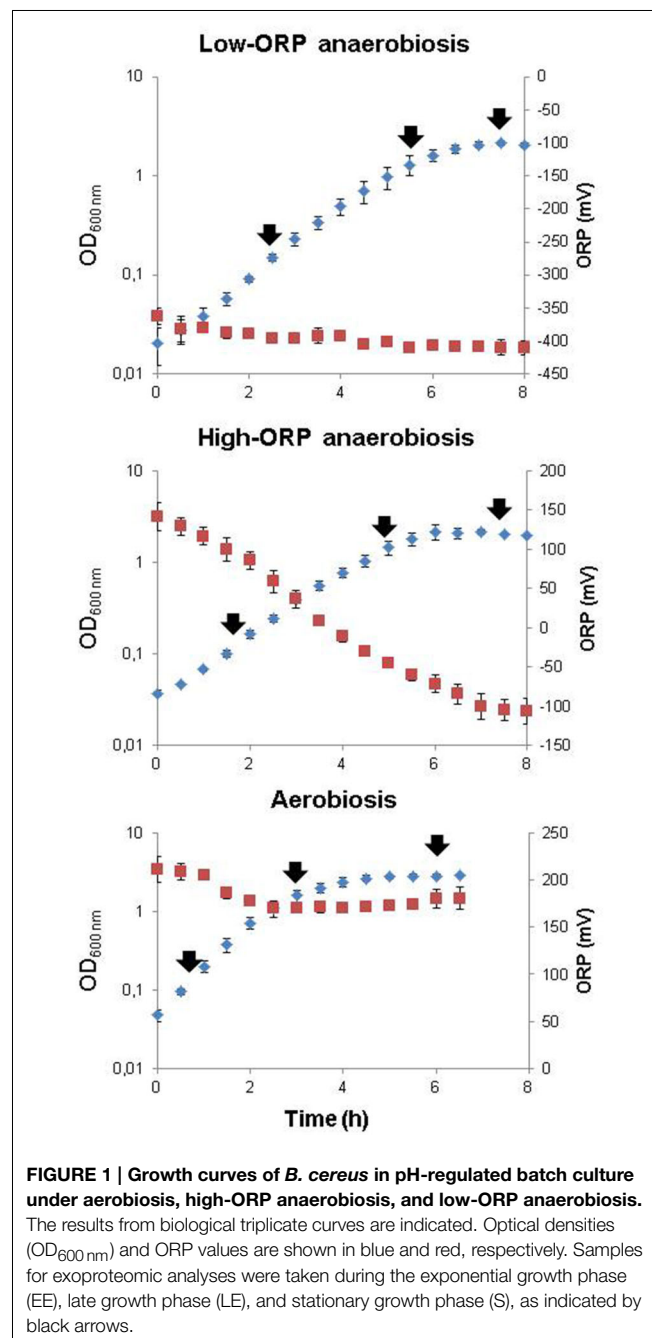
TABLE 1 | Novel exoproteins identified in this study.

Gene	Accession n° (NP)	Protein name	Functional annotation	Secretion signals and functional domains[a]	Low-ORP anaerobiosis EE	LE	S	High-ORP anaerobiosis EE	LE	S	Aerobiosis EE	LE	S
GROUP A[c]													
BC1177	NA		Tryptophanyl-tRNA synthetase		0.00	0.00	**0.19**	0.00	0.00	**0.14**	0.00	0.00	**0.27**
NA	NA	ArgC	Gamma-glutamylphosphate reductase		**0.06**	0.00	**0.06**	0.00	0.00	**0.03**	0.00	0.00	**0.09**
BC3716	NA	EntD	Enterotoxin-like	S, SH3	**0.19**	**0.15**	**0.19**	**0.65**	**0.27**	**0.23**	**0.19**	**0.27**	**0.34**
BC4122	NA		2′, 3′-cyclic nucleotide 2′-phosphodiesterase		**0.09**	**0.05**	**0.02**	**0.23**	**0.08**	**0.05**	**0.20**	**0.01**	0.00
BC5138	NA	Pgk	Phosphoglycerate kinase		**0.90**	**0.66**	**1.74**	**0.49**	**1.15**	**2.00**	**0.40**	**0.07**	**1.25**
BC5137	NA	TpiA	Triosephosphate isomerase,		**1.36**	**0.94**	**2.87**	**1.36**	**1.25**	**2.98**	**0.42**	**0.23**	**1.32**
BC1649	NA		Unknown		**0.02**	0.00	**0.02**	0.00	0.00	0.00	**0.46**	**0.02**	0.00
NA	NA	CnaE	Collagen adhesion protein	S	**0.95**	**0.44**	**0.51**	**1.27**	**0.76**	**0.32**	**0.70**	**0.76**	**0.44**
NA	NA	SipB	Signalpeptidase	S	**0.66**	**0.53**	**0.79**	**0.99**	**0.92**	**0.79**	**0.73**	**0.92**	**1.25**
BC3705	NA	GlnA	Glutamine synthetase		**1.00**	**0.30**	**0.74**	**0.48**	**0.40**	**0.78**	**0.78**	**0.38**	**1.20**
BC3763	NA		Cellwallhydrolase	S	0.00	**1.55**	**1.59**	**0.17**	0.00	0.00	**0.97**	**0.55**	0.00
GROUP B													
BC0602	NP_830419	NprA	Neutralprotease	S	0.00	0.00	0.00	0.00	0.00	**0.06**	0.00	0.00	**0.07**
BC1161	NP_830947	PrsA2	Peptidylprolylisomerase		0.00	0.00	**0.06**	0.00	0.00	0.00	0.00	0.00	**0.12**
BC1641	NP_831419	FgjB	Flagellar basalbody rod protein		0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.20**	0.00
BC1642	NP_831420	FgjC	Flagellar basalbody rod protein		0.00	**0.13**	**0.27**	0.00	**0.27**	0.00	0.00	**0.47**	**0.13**
BC1687	NP_831462		Unknown		0.00	0.00	0.00	0.00	0.00	**0.09**	0.00	0.00	**0.09**
BC1901	NP_831673		Unknown		0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.17**	**0.17**
BC4363	NP_834075		Ferrichrome-binding protein	S	0.00	0.00	0.00	0.00	0.00	**0.06**	0.00	**1.15**	**1.44**
BC4546	NP_834253		Ferrichrome-binding protein	S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.19**
BC5359	NP_835020		Aminopeptidase Y	S	0.00	0.00	**0.06**	0.00	**0.02**	0.00	0.00	0.00	**0.10**
GROUP C													
BC1669	NP_831739	LytR3	LytR family transcriptionalregulator		0.00	0.00	**0.08**	0.00	**0.05**	0.00	**0.03**	**0.13**	**0.13**
BC4549	NP_834256	IsdC	Cellsurface protein	S, srtB	0.00	0.00	0.00	0.00	0.00	0.00	**0.04**	**0.19**	**0.27**
BC2473	NP_832233		Beta-lactamase	S	0.00	0.00	0.00	0.00	**0.03**	0.00	**0.06**	**0.06**	**0.03**
BC1893	NP_831666		Scaffold protein		0.00	0.00	0.00	0.00	0.00	0.00	**0.07**	**0.07**	**0.35**
BC3221	NP_832962		Surface protein	S, Fib-alpha	**0.32**	**0.30**	**0.62**	**0.32**	**0.67**	**0.81**	**0.09**	**0.42**	**0.99**
BC1862	NP_831635		Unknown		0.00	0.00	0.00	0.00	0.00	0.00	**0.10**	0.00	0.00
BC1634	NP_831412		UDP-N-acetylenolpyruvoylglucosamine reductase		0.00	0.00	0.00	0.00	0.00	0.00	**0.11**	0.00	0.00
BC2186	NP_831951	FtsK	Celldivision protein	S	**0.05**	**0.09**	**0.09**	**0.05**	**0.14**	**0.14**	**0.14**	**0.45**	**0.23**
BC1660	NP_831437		Soluble lytic murein transglycosylase	S, LT_GEWL	**0.04**	0.00	0.00	**0.04**	0.00	0.00	**0.14**	0.00	0.00
BCp0009	NP_829897		DNA packaging		0.00	**0.08**	**0.08**	0.00	**0.16**	**0.08**	**0.16**	0.00	**0.08**
BCp0018	NP_829906		Unknown		0.00	**0.17**	**0.58**	0.00	**0.21**	**0.41**	**0.38**	**0.07**	**0.14**
BC4548	NP_834255		Cellsurface protein	S	**0.17**	**0.17**	**1.34**	0.00	**0.17**	**1.40**	**0.45**	**3.98**	**3.81**

[a]Domains that are inherently related to the functional annotation of the proteins are not indicated. S, export signal peptide; SH3, SRC homology 3 domain; srtB, sortase B cell surface sorting signal; Fib-alpha, fibrinogen alpha/beta chain family; LTGEWL, lytic transglycosylase and goose egg-white lysozyme (GEWL) domain.

[b]NSAF values are given at early growth phase (EE), declining growth phase (LE), and stationary growth phase under aerobiosis, high-ORP anaerobiosis, and low-ORP anaerobiosis.

[c]Group A, proteins not hitherto annotated (NA); Group B, proteins not detected in EE; Group C, proteins not detected in previous studies.

NA, not annotated.

Bold values are greater than zero.

0.05 were parsed using the IRMa 1.28.0 software (Dupierris et al., 2009). A protein was considered to be validated when at least two different peptides were detected in the same sample. The false-positive rate for protein identification was estimated using the appropriate decoy database as below 0.1% with these parameters.

Label-free Protein Quantification and Statistical Analysis

The number of MS/MS spectra per protein (spectral counts) was extracted for the 27 samples and used for protein quantitation. The normalized spectral abundance factor (NSAF) was calculated by dividing the spectral count for each observed protein by the polypeptide theoretical mass, as described previously (Christie-Oleza et al., 2012). Principal component analysis (PCA) was carried out with R version 3.0.1 (http://cran.r-project.org/bin/windows/base/old/3.0.1/). The data analyses were performed with "FactoMineR," a package written in R dedicated to multivariate exploration data analysis (Lê et al., 2008). PCA was carried out with biological replicates of each growth phase as individuals and the spectral counts of proteins as quantitative variables. The correlation coefficients between the variable and the coordinates of the individuals on the axis were calculated for all the variables, dimension by dimension. The significance of each correlation coefficient was calculated using a Student's *t*-test. Variables, for which the *p*-value associated with this test was smaller than 0.05, are reported in Table S4 in Supplementary Material.
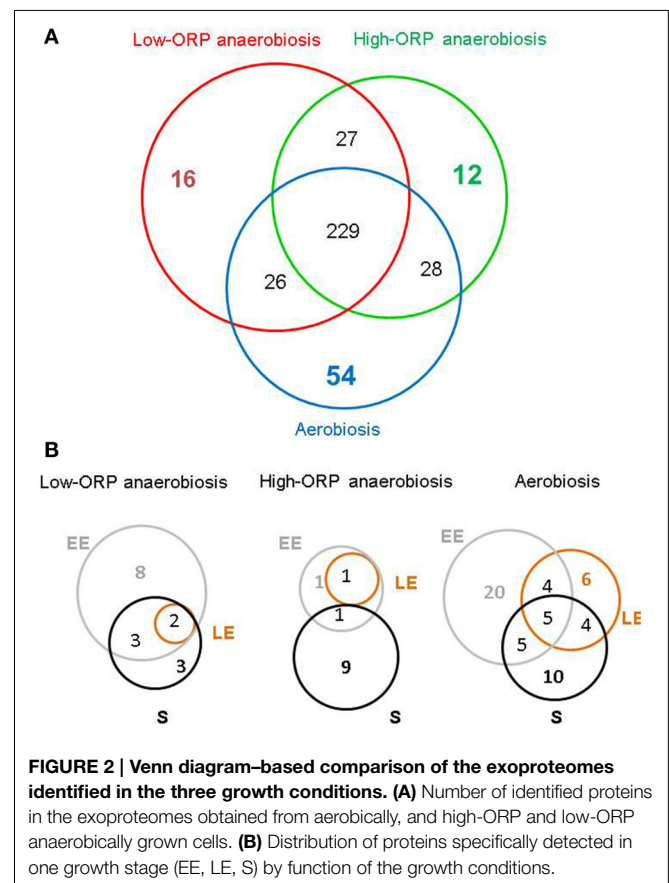
Proteomic Data Repository

The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (http://www.ebi.ac.uk/pride/), with the dataset identifier PXD001482 and DOI 10.6019/PXD001482.

Results and Discussion

Comparative Exoproteome, Large Survey
Growth Kinetics of *B. cereus* ATCC 14579

Bacteria were grown in pH- and temperature-regulated bioreactors using glucose as the sole carbon source (pH 7, 37°C, 30 mM glucose). Growth was investigated under aerobiosis (pO$_2$ =100%) and anaerobiosis (pO$_2$ =0%). Two different ORP conditions were obtained under anaerobiosis: a high-ORP anoxic condition (initial ORP =130 ± 20 mV) and a low-ORP anoxic condition (iORP = −390 ± 35 mV), this latter condition being achieved under flux of hydrogen, a non-toxic reducing agent. Three biological replicates were performed per culture condition. **Figure 1** shows the *B. cereus* growth curves and the extracellular ORP profiles established for the three culture conditions. As reported previously (Clair et al., 2012), *B. cereus* cells grew more slowly and produced less biomass in anoxic fermentative conditions than in oxic respiratory conditions. Changes in the initial extracellular ORP did not alter the growth rate and biomass production under fermentative anoxic conditions (Table S1 in Supplementary Material). However, the extracellular ORP profile differed significantly in the three conditions.

Under aerobiosis (initial ORP = 210 ± 13 mV), the ORP dropped rapidly to its minimal value (final ORP = 184 ± 11 mV). This reflects the rapid consumption of dissolved oxygen through respiration, to generate ATP for growth (Rosenfeld et al., 2005). The ORP measured under high-ORP anoxic fermentative conditions (iORP = 130 ± 20 mV) decreased concomitantly with the biomass increase to a reach a minimal value of −106 ± 16 mV, while under low-ORP conditions the ORP remained constant (iORP = −390 ± 35 mV and fORP = −410 ± 10 mV). Clearly, the reducing capacity of *B. cereus* cells is higher under high-ORP anaerobiosis than under low-ORP anaerobiosis (Le Lay et al., 2015). To examine the changes in exoproteome profiles associated with growth, samples were taken at the time points indicated by the arrows in **Figure 1**, i.e., during early exponential growth phase (EE), late exponential growth phase (LE), and stationary phase (S). Proteins from the 27 filtered supernatants were concentrated by precipitation with trichloroacetic acid. The resulting samples were then dissolved into NuPAGE LDS sample buffer supplemented with β-mercaptoethanol to prevent protein oxidation. Samples were loaded on NUPAGE® precast gels that were run for a short migration time only (Hartmann and Armengaud, 2014). NUPAGE® antioxidant was added in the upper buffer chamber to maintain the reduced state of the proteins during the run and avoid any protein oxidation. Each sample was excised from the gel as a polyacrylamide band. Trypsin proteolysis was carried out *in*-gel. The resulting peptides were analyzed

FIGURE 2 | Venn diagram–based comparison of the exoproteomes identified in the three growth conditions. (A) Number of identified proteins in the exoproteomes obtained from aerobically, and high-ORP and low-ORP anaerobically grown cells. **(B)** Distribution of proteins specifically detected in one growth stage (EE, LE, S) by function of the growth conditions.

by shotgun tandem mass spectrometry (Clair et al., 2010). A total of 120,470 MS/MS spectra were detected when considering the three biological repeats. Among them, 50,828 were assigned to *B. cereus* peptide sequences (Table S2 in Supplementary Material). A total of 392 proteins were identified based on the confident detection of at least two different peptides (Table S3 in Supplementary Material).

New Mass Spectrometry–Identified Exoproteins

Compared to previous large shotgun proteomic studies on exoproteomes from *B. cereus* ATCC 14579 (Clair et al., 2010; Laouami et al., 2014), a total of 32 proteins were detected for the first time. These 32 new mass spectrometry–certified proteins account for 11% of the exoproteome, as assessed by the global sum of their normalized spectral count abundance factors (NSAF) cumulated over the 27 samples (Table S3 in Supplementary Material). **Table 1** shows the sequence similarity–based functional annotation of these proteins and their abundances under aerobiosis, high-ORP- and low-ORP anaerobiosis. The 32 proteins could be categorized into three groups. Group A comprises 11 proteins that were not annotated in the first annotation report of the genome (Ivanova et al., 2003), but have been indicated by a proteogenomic study (unpublished data). Group B comprises 9 proteins that did not accumulate in EE growth phase in all the conditions tested, which explains why they were not detected in our previous study focused on this growth stage (Clair et al., 2010; Laouami et al., 2014). The protocol used in the present study probably favored the detection of the 12 other proteins (group C), which were found in very poor abundance. Among the new proteins identified, we identified a protein exhibiting high sequence similarity with the three putative enterotoxins, EntA, EntB, and EntC (Clair et al., 2010), and that we named EntD (unpublished results). Like EntD, 13 proteins comprised a predicted peptide signal. These were classified into cell-wall/cell-surface biogenesis, degradation/adhesion, and transport functional groups on the basis of data available in the literature and/or using the information

available in the Kegg classification (Table S3 in Supplementary Material). The other proteins did not contain typical peptide signals and were classified as flagella components (BC1641 and BC1642), enzymes of the central glycolytic pathway (TpiA-BC5137 and Pgk-BC5138), enzymes of amino acid–related metabolic pathways (ArgC and GlnA), chaperones (BC1161-PrsA2), translation/transcription-associated proteins (BC1177), and proteins with unknown functions (BC4122 and BC1649).

Insights into the Core-exoproteome of *B. cereus*

Figure 2A, shows a Venn diagram comparing the exoproteomes identified in the three different growth conditions. In this case, 229 of the 392 proteins identified were found to accumulate in the extracellular milieu, whatever the redox growth conditions. Regarding this feature from a quantitative perspective, this core proteome accounts for 89% of the total NSAF. Besides this core exoproteome, 54, 12, and 16 proteins were found exclusively in aerobically, high-ORP- and low-ORP–anaerobically grown cells, respectively. Globally, these proteins are poorly abundant, explaining why some of them were detected in the EE growth phase and not in the LE and S growth phases, especially under aerobiosis (20/54) and low-ORP anaerobiosis (8/16), as shown in **Figure 2B**. However, 5 and 2 proteins may be considered as fully representative of oxic and low-ORP anoxic conditions, respectively, because they were systematically detected in the three growth phases. The five aerobiosis-specific proteins are: the β-subunit of pyruvate dehydrogenase E1 (PdhB; BC3972), which catalyzes the decarboxylation of pyruvate into acetyl-CoA in oxic conditions; a ribosomal protein (RpsH, BC0145); a putative cell-surface protein (BC4549); a scaffold protein (BC1893); and a putative ferrichrome ABC transporter substrate-binding protein (BC5380). The two proteins that specifically accumulated under low-ORP anaerobiosis are a putative D-3-phosphoglycerate dehydrogenase (BC3248) and a putative nucleoside-binding protein (BC3791). No protein was found to be specifically assigned to high-ORP anoxic conditions.
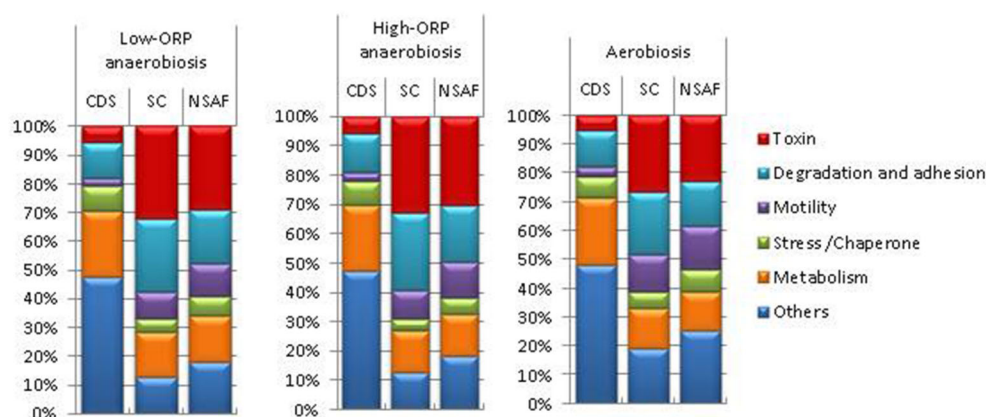

FIGURE 3 | Relative abundance of MS/MS-detected polypeptides under aerobiosis, high-ORP anaerobiosis, and low-ORP anaerobiosis. Ratio expressed in percentage of the total numbers of polypeptides (CDS), spectral counts (SC), and protein abundance normalized by the corresponding molecular weight (NSAF) per functional category are represented graphically by stacked bars for each condition.

Functional Insights into the Pan-Exoproteome of *B. cereus*

Figure 3 shows the whole set of exoproteins that were detected for the three growth phases in each growth condition and were classified into six main functional categories. The group "Others" comprises non-classical secreted proteins (translation, transcription, cell division, rod shape–related proteins), extracellular component of transport systems, proteins that are usually anchored to the bacterial membrane, and proteins with no function yet identified. Remarkably, more than 40% of the identified exoproteins (CDS) were classified in this group. Among these, 27 did not show any significant similarities with any known proteins, as determined by BLAST searches against the NCBInr database. Therefore, these could be considered as lineage-specific proteins for the *B. cereus* species (for more details see Table S3 in Supplementary Material). The number of CDS assigned to the toxin-related group is much lower (10-fold) than to the "Others" group, but the toxin-related group was more highly represented in terms of spectral counts (SC) and NSAF, and thus abundant whatever the condition. Toxin-related group represented the largest ratio of the MS/MS-detected peptides, with a range from 26 to 33%. Like the toxin-related group, the motility and stress/chaperone-related groups contain a low number of proteins. However, these two groups represent a lower abundance fraction of the exoproteome than the toxin-related group in the three conditions. Flagella components, usually anchored to the membrane, are the main contributors to the motility group (Table S3 in Supplementary Material). Their presence in the exoproteome could be explained by their fragility. When shaking the culture or removing cells by filtration or centrifugation, they can be easily broken into small pieces. Like the flagella components, the proteins belonging to the group comprising stress-
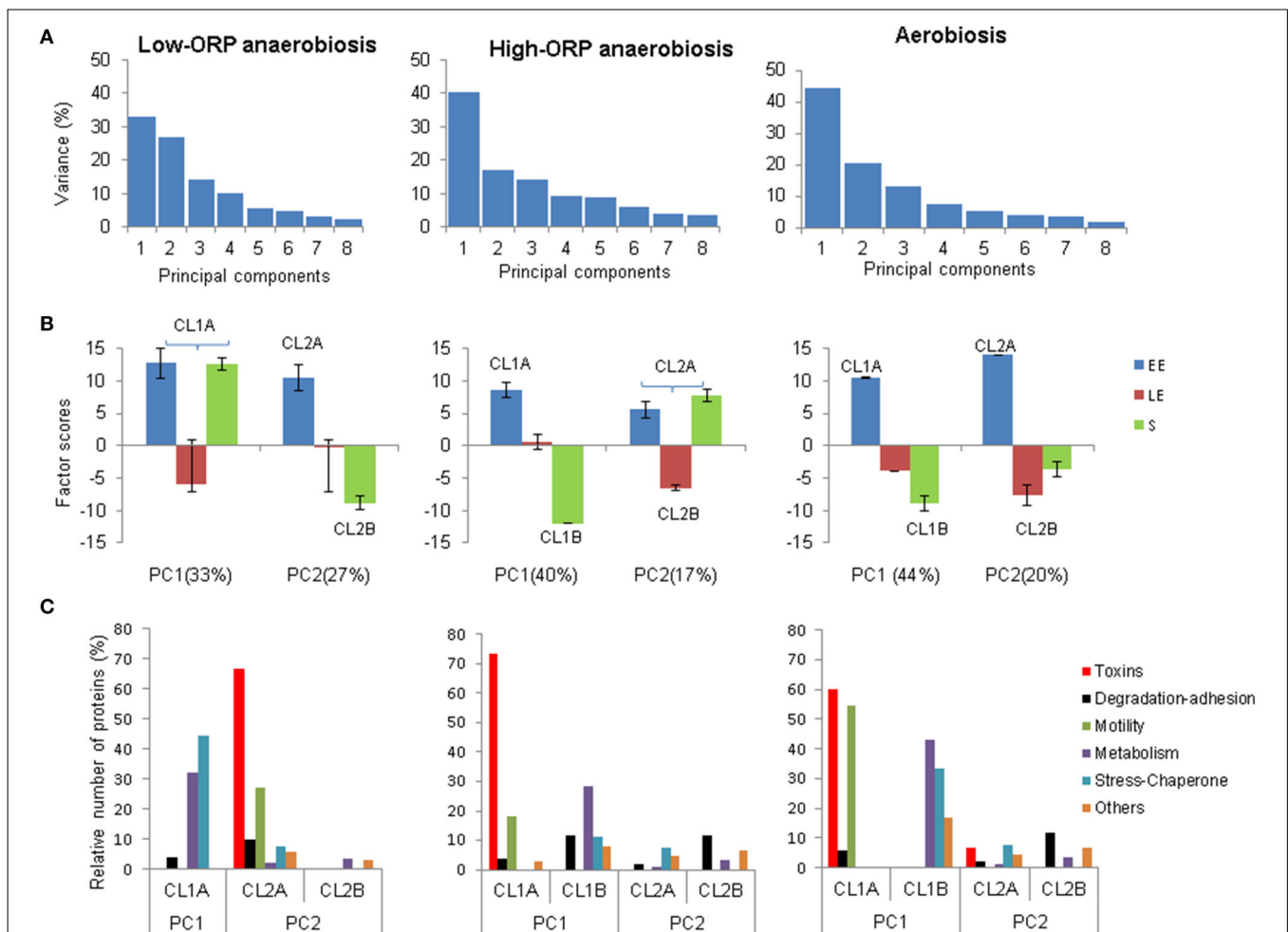
FIGURE 4 | Principal component analysis of the *B. cereus* exoproteome. (A) Fractions of the variances borne by axes 1–8. **(B)** Growth phase contributions to the first two principal components (PC1 and PC2), under low-ORP anaerobiosis, high-ORP anaerobiosis, and aerobiosis. Protein clusters assigned to growth phases were indicated by (i) the same capital letter **(A)** when they did not show abundance level change in these growth phases, or (ii) different capital letters **(A,B)** when the proteins showed negative correlation with abundance level changes. **(C)** Relative number of proteins assigned to toxins, degradation/adhesion, motility, metabolism, stress/chaperone, and "others" functional groups in protein clusters determined by PCA. Each functional group is represented by a color.

and chaperone-related proteins (such as catalases, superoxide dismutase, GroEL, Dnak, etc.) did not comprise any typical peptide signal. However, they are known as typical components of the exoproteome of pathogens (Armengaud et al., 2012). Adhesion and degradative proteins belong to an abundant fraction of the *B. cereus* exoproteome in the three conditions. The number of proteins dedicated to adhesion functions was lower than those assigned to degradation and the adhesion-related group was also less detected in terms of SC (Table S3 in Supplementary Material). The metabolism group comprises proteins related to central, amino acid, lipid, and fatty acid metabolism. The former subgroup is the most abundant and the latter the least abundant in terms of spectral counts (Table S3 in Supplementary Material). Specifically, **Figure 3** shows that the percentages of proteins belonging to the stress/chaperone-related and motility-related groups were higher under aerobiosis than under anaerobiosis, especially under high-ORP anaerobiosis. In contrast, the percentages of toxin-, degradative- and adhesion-related proteins were higher under anaerobiosis than under aerobiosis. The genes/operons involved in flagellum biosynthesis, enzymatic defenses against stress, and virulence factors are known to be tightly regulated in response to the presence or absence of dioxygen (Evans et al., 2011). This may contribute to the changes observed in the exoproteome.

Principal Component Analysis of *B. cereus* Exoproteome Dynamics

PCA was carried out to simplify the exoproteome time-course data of *B. cereus* (Ivosev et al., 2008; Jayapal et al., 2008), following a previous procedure (Clair et al., 2013). We chose to exclude

from the original datasets (259 proteins, Table S2 in Supplementary Material) the proteins found in less than two out of the three replicates for each growth phase sample in each condition. Considering the three growth phase–related observations (EE, LE, and S) and the three biological replicates for each observation, datasets for PCA comprised 9 readouts for 88 proteins under low-ORP anaerobiosis, 106 proteins under high-ORP anaerobiosis, and 114 proteins under aerobiosis. These datasets and analytical details are given in Table S4 in Supplementary Material.

Overview of Exoproteome Dynamics

PCAs extracted two principal components (PC1 and PC2), which explained ~60% of the total variance in the three conditions (**Figure 4A**). Scores and loadings of PC1 and PC2 are different in the three growth conditions (**Figure 4B**). This indicates that PCA extracted two time-course clusters (represented by PC1 and PC2) that did not contribute equally to the dynamics of the exoproteome in each condition. **Figure 4B** shows that, under low-ORP anaerobiosis, PC1 represented the tendency of some proteins (co-clustered in CL1A) to be similarly abundant in the EE and S growth phases. PC2 negatively correlates the abundance level decrease of some proteins (CL2A) between the EE and S growth phases with the abundance level increase of other proteins (CL2B). Under high-ORP anaerobiosis PC1 showed the same features as PC2 under low-ORP anaerobiosis and identified two protein clusters, named CL1A and CL1B. PC2 negatively correlates the absence of abundance level change of some proteins (CL2A) between the EE and S growth phases with the abundance level decrease of some proteins (CL2A) between the EE and LE growth phases. Under aerobiosis, PC1 represented the

TABLE 2 | Clustering of toxin-related proteins during *B. cereus* growth under low- and high-ORP anaerobiosis and aerobiosis.

		Clustering of toxin-related proteins[a]										
		Low-ORP anaerobiosis			High-ORP anaerobiosis				Aerobiosis			
Gene	Protein name	CL1A[b]	CL2A	CL2B	CL1A	CL1B	CL2A	CL2B	CL1A	CL1B	CL2A	CL2B
BC1110	CytK		■		■							
BC5239	EntA				■							
BC2952	EntB		■						■			
BC0813	EntC				■							
BC3716	EntD											
BC1953	EntFM											
BC3101	HblB'		■		■				■			
BC3102	HblB		■		■				■			
BC3103	Hbl1		■		■				■			
BC3104	Hbl2		■		■				■			
BC5101	HlyI										■	
BC3523	HlyII								■			
BC1809	NheA		■		■				■			
BC1810	NheB		■		■				■			
BC1811	NheC		■		■				■			

[a]Background colors identify proteins that are co-clustered.
[b]Clusters extracted from PCA and contributing to PC1 and PC2 were indicated as CL1 and CL2. The capital letters indicate sub-clusters of CL1 and CL2.

same features as PC1 and PC2 under high- and low-ORP anaerobiosis, respectively and identified two clusters of proteins CL1A and CL1B. PC2 negatively correlates the decrease in abundance level of some proteins (CL2A) with the increase in abundance level of other proteins (CL2B) between the EE and S growth phases.

Distribution of Functional Groups inside Kinetic Clusters of Proteins

All proteins contributing to the CL clusters extracted from PC1 and PC2 were assigned to one of the six functionally distinguished groups established in **Figure 3**. **Figure 4C** shows that, under low-ORP anaerobiosis, stress/chaperone- and metabolism-related proteins preferentially contributed to CL1A and toxin- and motility-related proteins to CL2A. Under both high-ORP anaerobiosis and aerobiosis, toxin-, motility-, metabolism-, and stress/chaperone-related proteins preferentially contributed to CL1A. However, CL1A co-clustered a higher number of toxin-related proteins under high-ORP anaerobiosis while it clustered a higher number of motility-, metabolism-, and stress-related proteins under aerobiosis. Taken together, the results show that toxin-related proteins displayed the highest functional-group homogeneity compared to other functionally related proteins in the three growth conditions. Specifically, PCA revealed that the decrease in abundance level of the majority of toxin-related proteins between EE and S growth phases was (i) uncorrelated with the change in abundance level of the majority of metabolism- and stress-related proteins under low-ORP anaerobiosis, (ii) negatively correlated with the increase in abundance level of less than ∼30% of metabolism-related proteins under high-ORP anaerobiosis, and (iii) negatively correlated with the increase in abundance level of more than 40 and 30% of metabolism- and stress-related proteins, respectively, under aerobiosis. Studies of metabolic network structures have shown that connected functional groups of proteins may contribute to a common cellular process (Ravasz et al., 2002). Our data raise the question of the role of toxins in *B. cereus* active growth, i.e., in primary metabolism and possibly in cellular protection against metabolism-related oxidative stress in respiring aerobic cells.
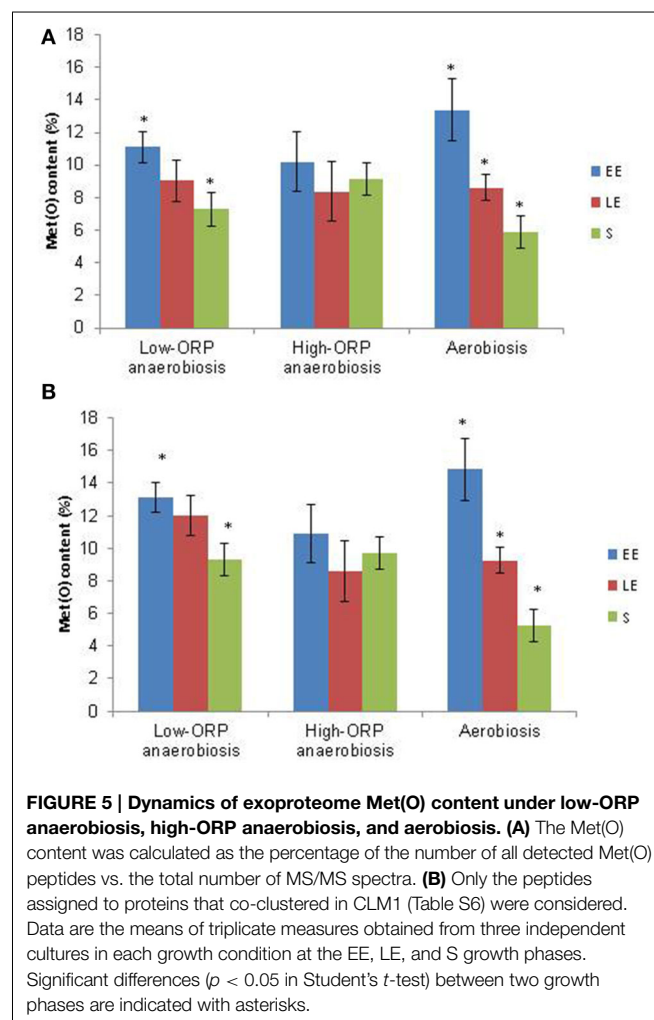
Focus on the Dynamics of Toxin-Related Proteins

Table 2 lists the toxin-related proteins that contributed to CL2A under low-ORP anaerobiosis and CL1A under high-ORP anaerobiosis and aerobiosis. The data show that the three hemolysin BL (Hbl) components (HblL1, HblL2, and HblB) co-clustered with HblB', which is encoded by the *hblB* gene located downstream of the *hblCDA* operon (Clair et al., 2010), in the three conditions. Co-clustering was also observed for the three non-hemolytic enterotoxin (Nhe) components, which are encoded by the *nheABC* operon (Lindback et al., 2004). Hbl and Nhe components also co-clustered with (i) hemolysin II (HlyII) under aerobiosis, (ii) EntB under both aerobiosis and low-ORP anaerobiosis, (iii) EntA and EntC under high-ORP anaerobiosis, and (iv) cytotoxin K (CytK) and Hly I under both high- and low-ORP anaerobiosis. In conclusion, Hbl and Nhe components may constitute the core of the toxin-related clusters and the other proteins

constitute the growth condition variance with (i) HlyII representative of aerobic respiratory condition, (ii) CytK and HlyI representatives of the anaerobic fermentative conditions, (iii) EntA and EntC representatives of classical anoxic conditions (high-ORP anaerobiosis), and (iv) EntB representative of both aerobic respiration and low-ORP anaerobic fermentation. These two latter conditions generate endogenous oxidative stress, which is counteracted by antioxidant systems. Among these, OhrRA was found to regulate EntB (Clair et al., 2012). Consequently, EntB could be a marker of oxidative stress–generating conditions.

Dynamics of the Met(O) Content of the *B. cereus* Exoproteome

In all Gram-positive bacteria, the majority of extracellular proteins need to remain unfolded to be translocated across the plasma membrane, the plasma membrane being known to support the highest level of ROS production in the cell (Fisher, 2009; Schneewind and Missiakas, 2014). On the other hand, Met residues in polypeptidic chains are more sensitive to oxidation than Met residues in mature proteins, as Met residues are usually located in the hydrophobic core of proteins (Fliss et al.,

FIGURE 5 | Dynamics of exoproteome Met(O) content under low-ORP anaerobiosis, high-ORP anaerobiosis, and aerobiosis. (A) The Met(O) content was calculated as the percentage of the number of all detected Met(O) peptides vs. the total number of MS/MS spectra. **(B)** Only the peptides assigned to proteins that co-clustered in CLM1 (Table S6) were considered. Data are the means of triplicate measures obtained from three independent cultures in each growth condition at the EE, LE, and S growth phases. Significant differences ($p < 0.05$ in Student's *t*-test) between two growth phases are indicated with asterisks.

FIGURE 6 | Characteristics of the protein cluster CLM1 determined by PCA. Number of proteins assigned to toxins, degradation/adhesion, motility, metabolism, stress/chaperone, and "others" functional groups that co-clustered in CLM1 under low-ORP anaerobiosis, high-ORP anaerobiosis, and aerobiosis. The number of proteins with correlated (c) and uncorrelated (nc) abundance levels and Met(O) content changes is indicated for each growth condition.
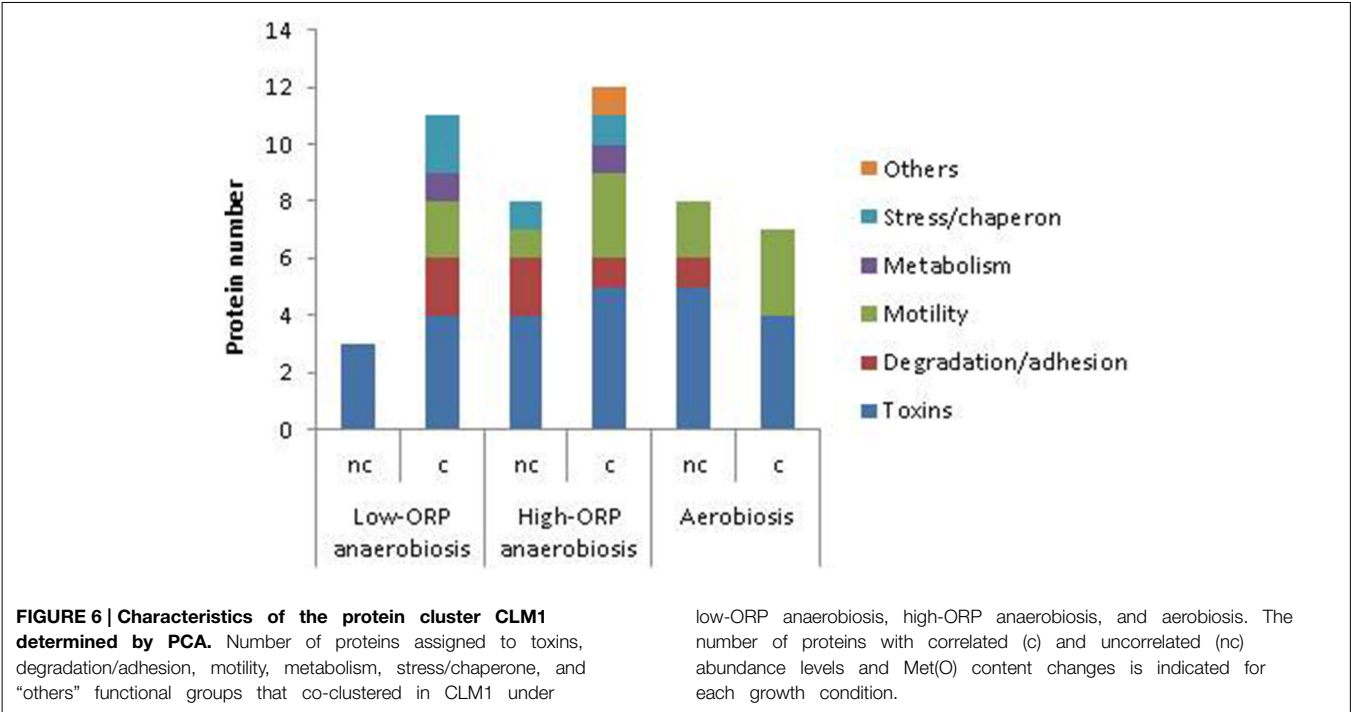
TABLE 3 | Co-clustering of toxin-related proteins in CLM1 under low-ORP anaerobiosis, high-ORP anaerobiosis, and aerobiosis.

Gene	Protein name	Clustering of toxin-related proteins[a]		
		Anaerobiosis		Aerobiosis
		Low-ORP	High-ORP	
BC1110	CytK			
BC5239	EntA		c[b]	nc
BC2952	EntB	nc	nc	c
BC0813	EntC		c	nc
BC1953	EntFM		nc	
BC3102	HblB	nc	nc	nc
BC3103	HblL1	c		c
BC3104	HblL2	c	c	c
BC3101	HblB'	nc	nc	nc
BC5101	HlyI			
BC1809	NheA	c	c	nc
BC1810	NheB	c	c	c
BC1811	NheC			

[a]*Background colors identify proteins that are co-clustered.*
[b]*The symbols c and nc indicate that the Met(O) peptide content change of a protein is correlated or uncorrelated, respectively, with its abundance level change during growth.*

1983; Drazic and Winter, 2014). For these reasons, intracellular ROS may cause significant oxidation of exoproteins prior to their translocation. Insofar as Met(O) residues are not reduced back to Met, and there is no ROS source in the extracellular medium, the Met(O) content of the exoproteome might directly reflect endogenous ROS oxidation. To test this hypothesis, we used nanoLC-MS/MS to assess Met(O) content in all the proteins identified in the exoproteome. We analyzed their time-course dynamics in aerobically grown cells and in anaerobically grown cells for this specific parameter.

Overview of Methionine Oxidation

A total of 4532 peptides containing oxidized Met residue(s) (Met(O) peptides) were identified along the 27 nanoLC-MS/MS runs (Table S1 in Supplementary Material). A total of 211 different Met(O) peptides were listed (Table S5 in Supplementary Material), a significant number of them being detected reproducibly. The Met(O) peptide content of the *B. cereus* exoproteome was estimated as a percentage of the total number of peptides identified in each of the three biological samples obtained for each growth phase sample under low- and high-ORP anaerobiosis and aerobiosis. **Figure 5A** shows that the Met(O) peptide content of the *B. cereus* exoproteome decreased significantly during growth under low-ORP anaerobiosis and aerobiosis, to reach its minimum in the stationary phase. However, aerobiosis sustains a higher decrease along this kinetic compared to low-ORP anaerobiosis. Strikingly, no significant change was observed under high-ORP anaerobiosis. Similar results were obtained by comparing the number of Met(O) to the total number of Met (Figure S1 in Supplementary Material). The level of Met oxidation as assessed here is a complex result of the balance between endogenous ROS generation on the one hand and the ability of the cell to repair Met on the other. Oxidized Met can be repaired by antioxidant systems (Drazic and Winter, 2014). Under aerobiosis, the high Met(O) peptide content of the EE exoproteome compared to the S exoproteome could reflect either a surplus of ROS generated by the activity of the respiratory chain (Seaver and Imlay, 2001) or a higher activity of the antioxidant

systems in S growth phase (Alamuri and Maier, 2006; Vekaria and Chivukula, 2010). Under anaerobiosis, and in the absence of final electron acceptors for respiratory electron processes, *B. cereus* cells ferment glucose (Zigha et al., 2007). Fermentative pathways do not produce ROS as typical metabolic by-products under classical anaerobic conditions (Landolfo et al., 2008). This may explain why there is no change in the Met(O) peptide content of the *B. cereus* exoproteome during growth under high-ORP anaerobiosis. We reported previously that reductive stress, such as is encountered under low-ORP anaerobiosis, caused intracellular redox imbalance at the EE growth phase, and generated a secondary oxidative stress response (Mols and Abee, 2011; Clair et al., 2013). This could increase the ability of anaerobic cells to repair oxidized Met and explain why S growth phase sustains a lower Met(O) content under low-ORP anaerobiosis than under high-ORP anaerobiosis.

Identification of Proteins with Differential Abundance Levels and Met(O)-Content Dynamics

To identify proteins exhibiting differences in abundance level and Met(O)-content dynamics, we conducted a second PCA using both abundance (in terms of total number of peptides) and Met(O) peptide content (number of Met(O)-containing peptides) to define proteins in each growth condition. For a robust analysis of the variability in terms of Met(O) peptide content, we considered the proteins containing at least one Met(O) peptide identified in at least two biological replicates. A total of 43 proteins were confidently listed as being oxidized with this criterion (Table S6 in Supplementary Material). Among these, 13 proteins are toxin-related proteins. Remarkably, EntD and HlyII are the only components from the list of detected toxins reported in **Table 2** that are not post-translationally modified. The other oxidized proteins are degradative enzymes and adhesins (10), and to a lesser extent, flagella (6), stress-related proteins (4), metabolism-related proteins (7), and uncharacterized proteins (3). PCA extracted 3 Met(O)-related groups (CLM1-3) under low-ORP anaerobiosis, high-ORP anaerobiosis, and aerobiosis (Table S6 in Supplementary Material). **Figure 5B** shows that CLM1 is representative of the variability of the Met(O) peptide content of the *B. cereus* exoproteome during growth in the three conditions tested. When analyzing the correlation between Met(O) peptide content and abundance level, proteins with differential abundance levels and Met(O)-content dynamics were highlighted. These represent 27, 40, and 53% of proteins co-clustered in CLM1 under low- and high-ORP anaerobiosis, and aerobiosis, respectively (**Figure 6**). This suggests that oxidation of

```
MKKTLITGLLVTAVSTSCFIPVSAYAKEGQTEVKTVYAQNVIAPNTLSNS 50
IRMLGSQSPLIQAYGLIILQQPDIKVNAMSSLTNHQKFAKANVREWIDEY 100
NPKLIDLNQEMMRYSTRFNSYYSKLYELAGNVNEDQQAKADFMSAYGKLQ 150
LQVQSIQESMEQDLLELNRFKTVLDKDSNNLSIKADEAIKTLQGSSGDIV 200
KLREDIKRIQGEIQAELTTILNRPQEIIKGSINIGKQVFTITNQTAQTKT 250
IDFVSIGTLSNEIVNAADSQTREAALRIQQKQKELLPLIQKLSQTEAEAT 300
QITFVEDQVNSFTELIDRQITTLETLLTDWKVLNNNMIQIQTNVEEGTYT 350
DSSLLQKHFNQIKKVSDEMNKQTNQFEDYVTNVEVH             386
```

FIGURE 7 | Amino acid sequence of NheA. Peptides detected by LC-MS/MS are shown in red and are underlined. Met residues are shown in bold.

TABLE 4 | List of NheA peptides containing oxidized and non-oxidized Met residues.

		LC-MS/MS identification		
		Met oxidation		
Peptides detected by LC-MS/MS	Met[a]	**Anaerobiosis**		**Aerobiosis**
		Low-ORP	**High-ORP**	
MLGSQSPLIQAYGLIILQQPDIK	M53	M53(O)	M53(O)	nd[b]
		M111(O)	M111(O)	nd
LIDLNQE**MM**R	M111	M111(O) M112(O)	M111(O) M112(O)	M111(O)M112(O)
	M112	M112(O)	M112(O)	M112(O)
ADF**M**SAYGK	M143	nd	Nd	nd
LQLQVQSIQES**M**EQDLLELNR	M160	nd	Nd	nd
VLNNN**M**IQIQTNVEEGTYTDSSLLQK	M337	nd	Nd	nd
VSDE**M**NKQTNQFEDYVTNVEVH	M369	nd	Nd	nd

[a]*Methionine residues (Met) and oxidized Met residues Met(O) were identified by their position in the protein sequence (**Figure 7**).*
[b]*Nd indicates that no oxidized Met residue was detected.*
Methionine residues are indicated in bold in peptides detected by LC-MS/MS.

Gene	Protein name	Number of Met residues		Number of Met(O) residues		
		Total[a]	Detected[b]	Low-ORP anaerobiosis	High-ORP anaerobiosis	Aerobiosis
BC1110	CytK	5	5	1	2	0
BC5239	EntA	3	2	2	2	2
BC2952	EntB	4	2	1	1	1
BC0813	EntC	2	2	2	2	2
BC1953	EntFM	1	1	0	1	0
BC3101	HblB'	13	1	1	1	1
BC3102	HblB	7	6	6	5	4
BC3103	HblL1	7	4	4	4	3
BC3104	HblL2	7	6	6	5	4
BC5101	HlyI	5	2	0	0	2
BC1809	NheA	8	7	3	3	2
BC1810	NheB	4	3	2	2	2
BC1811	NheC	10	2	0	1	0

[a]*The number of Met residues was calculated from the sequence of the mature form of the protein (without peptide signal).*
[b]*The numbers reported in this column are the numbers of Met residues detected in our study by LC-MS/MS.*

Met residues may be more specific under aerobiosis than under anaerobiosis. **Figure 6** shows that CLM1 comprises a significant subset of Met(O) toxin–related proteins whatever the conditions (7, 9, and 9 under low- and high-ORP anaerobiosis, and aerobiosis, respectively). **Table 3** lists the toxin-related proteins that contributed to CML1 and differentiates proteins with similar abundance levels and Met(O)-content dynamics from proteins with differential abundance levels and Met(O)-content dynamics. The data show that HblB, HblL2, HblB', NheA, NheB, and EntB may constitute the core of the toxin-related sub clusters and HblL1, EntA, EntC, and EntFM constitute the growth condition variance with EntFM representative of high-ORP aerobiosis. **Table 3** also shows that aerobiosis may sustain higher specific oxidation of Met residues in NheA compared to anaerobiosis. To further strengthen this latter observation, we analyzed the peptides specifically assigned to NheA (**Figure 7**). Among the 7 Met residues detected in the 6 NheA-assigned peptides reported in **Figure 7**, four were never detected as oxidized (**Table 4** and Supplementary Table S6). This indicates that all NheA-bound methionines are not equally susceptible to oxidation. This may be due to their neighboring amino acids (Ghesquiere et al., 2011). Secondly, NheA contains one Met residue (M53) that is oxidized under anaerobiosis but not under aerobiosis. In addition, NheA contains two adjacent Met residues at positions 111 and 112, which are differentially oxidized under aerobiosis compared to anaerobiosis: oxidation of the first Met residue (M111) occurred only when the second (M112) was oxidized under aerobiosis, while oxidation of M111 did not depend on M112 oxidation under anaerobiosis. Therefore, NheA contains Met residues that respond differently to oxidation under anaerobiosis and

aerobiosis. This is also the case for CytK, EntFM, HblB, HblL1, HblL2, and NheC, which all contain one Met residue oxidized under anaerobiosis but not under aerobiosis (**Table 5**). Thus, anaerobiosis increases the oxidation susceptibility of methionine in toxin-related proteins. This may due to the presence of a different pattern of oxidants in fermentative cells (Mahawar et al., 2012). Taken together, our data indicate that toxin-related proteins contain Met residues that are not equally susceptible to oxidation and Met residue selectivity is a factor that may contribute to Met oxidation under aerobiosis.

Conclusion

We used nanoLC-MS/MS data to analyze global changes in the *B. cereus* exoproteome during growth in glucose-containing medium under controlled conditions of pH and pO$_2$. We have shown that PCA can identify groups of exoproteins that are coordinately controlled at the growth phase level. The results indicated that proteins belonging to the toxin-related group define characteristic kinetic profiles correlated with the physiological state of the culture in respiring, as in fermenting, cells. The majority of toxin-related proteins accumulated during the exponential growth phase, whatever the conditions. However, their dynamics differ significantly under aerobiosis and anaerobiosis if we consider how their patterns in terms of metabolism, oxidative stress–related proteins and the time dynamics of their Met(O) content are interconnected. Several studies have reported that Met residues of proteins may act as ROS scavengers (Luo and Levine, 2009). It is thus possible that Met residues in toxin-related proteins may act as endogenous antioxidants before being secreted into the extracellular medium. High-level secretion of toxins during the exponential phase may thus contribute to the protection of *B. cereus* cells against cellular oxidation and maintain redox homeostasis by keeping endogenous ROS at bay, especially under aerobiosis. Evidently further studies should be now conducted to confirm these hypotheses. The consequences of methionine oxidation on proteins may vary from structural alterations leading to altered activity and/or altered signal events to protein degradation (Levine et al., 2000). This raises questions about the role of Met oxidation in *B. cereus* virulence, and especially in *B. cereus* cytotoxicity. Indeed, our study demonstrated that the major cytotoxins of the *B. cereus* exoproteome, such as Nhe and Hbl (Sastalla et al., 2013), contain oxidizable methionines, and the effect of oxidation on their biological activity is worthy of documentation.

Acknowledgments

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.00342/abstract

References

Alamuri, P., and Maier, R. J. (2006). Methionine sulfoxide reductase in *Helicobacter pylori*: interaction with methionine-rich proteins and stress-induced expression. *J. Bacteriol.* 188, 5839–5850. doi: 10.1128/JB.00430-06

Armengaud, J., Christie-Oleza, J. A., Clair, G., Malard, V., and Duport, C. (2012). Exoproteomics: exploring the world around biological systems. *Expert Rev. Proteomics* 9, 561–575. doi: 10.1586/epr.12.52

Bishop, B. L., Lodolce, J. P., Kolodziej, L. E., Boone, D. L., and Tang, W. J. (2010). The role of anthrolysin O in gut epithelial barrier disruption during *Bacillus anthracis* infection. *Biochem. Biophys. Res. Commun.* 394, 254–259. doi: 10.1016/j.bbrc.2010.02.091

Brynildsen, M. P., Winkler, J. A., Spina, C. S., Macdonald, I. C., and Collins, J. J. (2013). Potentiating antibacterial activity by predictably enhancing endogenous microbial ROS production. *Nat. Biotechnol.* 31, 160–165. doi: 10.1038/nbt.2458

Chi, B. K., Gronau, K., Maeder, U., Hessling, B., Becher, D., and Antelmann, H. (2011). S-bacillithiolation protects against hypochlorite stress in *Bacillus subtilis* as revealed by transcriptomics and redox proteomics. *Mol. Cell. Proteomics.* 10:M111.009506. doi: 10.1074/mcp.M111.009506

Christie-Oleza, J. A., Pina-Villalonga, J. M., Bosch, R., Nogales, B., and Armengaud, J. (2012). Comparative proteogenomics of twelve *Roseobacter* exoproteomes reveals different adaptive strategies among these marine bacteria. *Mol. Cell. Proteomics* 11:M111.013110. doi: 10.1074/mcp.M111.013110

Clair, G., Armengaud, J., and Duport, C. (2012). Restricting fermentative potential by proteome remodeling: an adaptive strategy evidenced in *Bacillus cereus*. *Mol. Cell. Proteomics* 11:M111.013102. doi: 10.1074/mcp.M111.013102

Clair, G., Lorphelin, A., Armengaud, J., and Duport, C. (2013). OhrRA functions as a redox-responsive system controlling toxinogenesis in *Bacillus cereus*. *J. Proteomics* 94, 527–539. doi: 10.1016/j.jprot.2013.10.024

Clair, G., Roussi, S., Armengaud, J., and Duport, C. (2010). Expanding the known repertoire of virulence factors produced by *Bacillus cereus* through early secretome profiling in three redox conditions. *Mol. Cell. Proteomics* 9, 1486–1498. doi: 10.1074/mcp.M000027-MCP201

Dedieu, A., Gaillard, J. C., Pourcher, T., Darrouzet, E., and Armengaud, J. (2011). Revisiting iodination sites in thyroglobulin with an organ-oriented shotgun strategy. *J. Biol. Chem.* 286, 259–269. doi: 10.1074/jbc.M110.159483

De Groot, A., Dulermo, R., Ortet, P., Blanchard, L., Guerin, P., Fernandez, B., et al. (2009). Alliance of proteomics and genomics to unravel the specificities of Sahara bacterium *Deinococcus deserti*. *PLoS Genet.* 5:e1000434. doi: 10.1371/journal.pgen.1000434

Drazic, A., and Winter, J. (2014). The physiological role of reversible methionine oxidation. *Biochim. Biophys. Acta.* 1844, 1367–1382. doi: 10.1016/j.bbapap.2014.01.001

Dupierris, V., Masselon, C., Court, M., Kieffer-Jaquinod, S., and Bruley, C. (2009). A toolbox for validation of mass spectrometry peptides identification and generation of database: IRMa. *Bioinformatics* 25, 1980–1981. doi: 10.1093/bioinformatics/btp301

Duport, C., Zigha, A., Rosenfeld, E., and Schmitt, P. (2006). Control of enterotoxin gene expression in *Bacillus cereus* F4430/73 involves the redox-sensitive ResDE signal transduction system. *J. Bacteriol.* 188, 6640–6651. doi: 10.1128/JB.00702-06

Evans, M. R., Fink, R. C., Vazquez-Torres, A., Porwollik, S., Jones-Carson, J., McClelland, M., et al. (2011). Analysis of the ArcA regulon in anaerobically grown *Salmonella enterica* sv. *Typhimurium*. *BMC Microbiol.* 11:58. doi: 10.1186/1471-2180-11-58

Fabich, A. J., Jones, S. A., Chowdhury, F. Z., Cernosek, A., Anderson, A., Smalley, D., et al. (2008). Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. *Infect. Immun.* 76, 1143–1152. doi: 10.1128/IAI.01386-07

Fisher, A. B. (2009). Redox signaling across cell membranes. *Antioxid. Redox Signal.* 11, 1349–1356. doi: 10.1089/ars.2008.2378

Fliss, H., Weissbach, H., and Brot, N. (1983). Oxidation of methionine residues in proteins of activated human neutrophils. *Proc. Natl. Acad. Sci. U.S.A.* 80, 7160–7164. doi: 10.1073/pnas.80.23.7160

Ghesquiere, B., Jonckheere, V., Colaert, N., Van Durme, J., Timmerman, E., Goethals, M., et al. (2011). Redox proteomics of protein-bound methionine oxidation. *Mol. Cell. Proteomics* 10:M110.006866. doi: 10.1074/mcp.M110.006866

Gilois, N., Ramarao, N., Bouillaut, L., Perchat, S., Aymerich, S., Nielsen-Leroux, C., et al. (2007). Growth-related variations in the *Bacillus cereus* secretome. *Proteomics* 7, 1719–1728. doi: 10.1002/pmic.200600502

Gohar, M., Faegri, K., Perchat, S., Ravnum, S., Okstad, O. A., Gominet, M., et al. (2008). The PlcR virulence regulon of *Bacillus cereus*. *PLoS ONE* 3:e2793. doi: 10.1371/journal.pone.0002793

Gonzalez-Flecha, B., and Demple, B. (1995). Metabolic sources of hydrogen peroxide in aerobically growing *Escherichia coli*. *J. Biol. Chem.* 270, 13681–13687. doi: 10.1074/jbc.270.23.13681

Guyton, A. C. (1977). "Basic human physiology," in *Normal Function and Mechanisms of Diseases* (Philadelphia, PA: W. B. Saunders Company), 408–423.

Hartmann, E. M., and Armengaud, J. (2014). N-terminomics and proteogenomics, getting off to a good start. *Proteomics.* 14, 2637–2646. doi: 10.1002/pmic.201400157

Henderson, B., and Martin, A. (2011). Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease. *Infect. Immun.* 79, 3476–3491. doi: 10.1128/IAI.00179-11

Imlay, J. A. (2013). The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. *Nat. Rev. Microbiol.* 11, 443–454. doi: 10.1038/nrmicro3032

Ingmer, H., and Brondsted, L. (2009). Proteases in bacterial pathogenesis. *Res. Microbiol.* 160, 704–710. doi: 10.1016/j.resmic.2009.08.017

Ivanova, N., Sorokin, A., Anderson, I., Galleron, N., Candelon, B., Kapatral, V., et al. (2003). Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. *Nature* 423, 87–91. doi: 10.1038/nature01582

Ivosev, G., Burton, L., and Bonner, R. (2008). Dimensionality reduction and visualization in principal component analysis. *Anal. Chem.* 80, 4933–4944. doi: 10.1021/ac800110w

Jayapal, K. P., Philp, R. J., Kok, Y. J., Yap, M. G., Sherman, D. H., Griffin, T. J., et al. (2008). Uncovering genes with divergent mRNA-protein dynamics in *Streptomyces coelicolor*. *PLoS ONE* 3:e2097. doi: 10.1371/journal.pone.0002097

Kamar, R., Gohar, M., Jehanno, I., Rejasse, A., Kallassy, M., Lereclus, D., et al. (2013). Pathogenic potential of *Bacillus cereus* strains as revealed by phenotypic analysis. *J. Clin. Microbiol.* 51, 320–323. doi: 10.1128/JCM.02848-12

Landolfo, S., Politi, H., Angelozzi, D., and Mannazzu, I. (2008). ROS accumulation and oxidative damage to cell structures in *Saccharomyces cerevisiae* wine strains during fermentation of high-sugar-containing medium. *Biochim. Biophys. Acta* 1780, 892–898. doi: 10.1016/j.bbagen.2008.03.008

Laouami, S., Clair, G., Armengaud, J., and Duport, C. (2014). Proteomic evidences for rex regulation of metabolism in toxin-producing *Bacillus cereus* ATCC 14579. *PLoS ONE* 9:e107354. doi: 10.1371/journal.pone.0107354

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R Package for multivariate analysis. *J. Stat. Softw.* 25, 1–18.

Le Lay, J., Bahloul, H., Serino, S., Jobin, M., and Schmitt, P. (2015). Reducing activity, glucose metabolism and acid tolerance response of *Bacillus cereus* grown at various pH and oxydo-reduction potential levels. *Food Microbiol.* 46, 314–321. doi: 10.1016/j.fm.2014.07.007

Levine, R. L., Moskovitz, J., and Stadtman, E. R. (2000). Oxidation of methionine in proteins: roles in antioxidant defense and cellular regulation. *IUBMB Life* 50, 301–307. doi: 10.1080/15216540051081056

Lindback, T., Fagerlund, A., Rodland, M. S., and Granum, P. E. (2004). Characterization of the *Bacillus cereus* Nhe enterotoxin. *Microbiology* 150, 3959–3967. doi: 10.1099/mic.0.27359-0

Luo, S., and Levine, R. L. (2009). Methionine in proteins defends against oxidative stress. *FASEB J.* 23, 464–472. doi: 10.1096/fj.08-118414

Mahawar, M., Atianand, M. K., Dotson, R. J., Mora, V., Rabadi, S. M., Metzger, D. W., et al. (2012). Identification of a novel *Francisella tularensis* factor required for intramacrophage survival and subversion of innate immune response. *J. Biol. Chem.* 287, 25216–25229. doi: 10.1074/jbc.M112.367672

Mailloux, R. J., Lemire, J., and Appanna, V. D. (2011). Metabolic networks to combat oxidative stress in *Pseudomonas fluorescens*. *Antonie Van Leeuwenhoek* 99, 433–442. doi: 10.1007/s10482-010-9538-x

Marteyn, B., West, N. P., Browning, D. F., Cole, J. A., Shaw, J. G., Palm, F., et al. (2010). Modulation of *Shigella* virulence in response to available oxygen *in vivo*. *Nature* 465, 355–358. doi: 10.1038/nature08970

Mols, M., and Abee, T. (2011). Primary and secondary oxidative stress in *Bacillus*. *Environ. Microbiol.* 13, 1387–1394. doi: 10.1111/j.1462-2920.2011.02433.x

Moriarty-Craige, S. E., and Jones, D. P. (2004). Extracellular thiols and thiol/disulfide redox in metabolism. *Annu. Rev. Nutr.* 24, 481–509. doi: 10.1146/annurev.nutr.24.012003.132208

Nieselt, K., Battke, F., Herbig, A., Bruheim, P., Wentzel, A., Jakobsen, O. M., et al. (2010). The dynamic architecture of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* 11:10. doi: 10.1186/1471-2164-11-10

Ramarao, N., and Sanchis, V. (2013). The pore-forming haemolysins of *Bacillus cereus*: a review. *Toxins (Basel).* 5, 1119–1139. doi: 10.3390/toxins5061119

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabasi, A. L. (2002). Hierarchical organization of modularity in metabolic networks. *Science* 297, 1551–1555. doi: 10.1126/science.1073374

Rosenfeld, E., Duport, C., Zigha, A., and Schmitt, P. (2005). Characterization of aerobic and anaerobic vegetative growth of the food-borne pathogen *Bacillus cereus* F4430/73 strain. *Can. J. Microbiol.* 51, 149–158. doi: 10.1139/w04-132

Sadagopal, S., Braunstein, M., Hager, C. C., Wei, J., Daniel, A. K., Bochan, M. R., et al. (2009). Reducing the activity and secretion of microbial antioxidants enhances the immunogenicity of BCG. *PLoS ONE* 4:e5531. doi: 10.1371/journal.pone.0005531

Sastalla, I., Fattah, R., Coppage, N., Nandy, P., Crown, D., Pomerantsev, A. P., et al. (2013). The *Bacillus cereus* Hbl and Nhe tripartite enterotoxin components assemble sequentially on the surface of target cells and are not interchangeable. *PLoS ONE* 8:e76955. doi: 10.1371/journal.pone.0076955

Schneewind, O., and Missiakas, D. (2014). Sec-secretion and sortase-mediated anchoring of proteins in Gram-positive bacteria. *Biochim. Biophys. Acta.* 1843, 1687–1697. doi: 10.1016/j.bbamcr.2013.11.009

Seaver, L. C., and Imlay, J. A. (2001). Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *J. Bacteriol.* 183, 7173–7181. doi: 10.1128/JB.183.24.7173-7181.2001

Stadtman, E. R., Van Remmen, H., Richardson, A., Wehr, N. B., and Levine, R. L. (2005). Methionine oxidation and aging. *Biochim. Biophys. Acta* 1703, 135–140. doi: 10.1016/j.bbapap.2004.08.010

Stenfors Arnesen, L. P., Fagerlund, A., and Granum, P. E. (2008). From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol. Rev.* 32, 579–606. doi: 10.1111/j.1574-6976.2008.00112.x

Vekaria, H. J., and Chivukula, R. P. (2010). Influence of thiol stress on oxidative phosphorylation and generation of ROS in *Streptomyces coelicolor*. *J. Biophys. Chem.* 1, 172–176. doi: 10.4236/jbpc.2010.13020

Vellasamy, K. M., Vasu, C., Puthucheary, S. D., and Vadivelu, J. (2009). Comparative analysis of extracellular enzymes and virulence exhibited by *Burkholderia pseudomallei* from different sources. *Microb. Pathog.* 47, 111–117. doi: 10.1016/j.micpath.2009.06.003

Vogt, W. (1995). Oxidation of methionyl residues in proteins—tools, targets, and reversal. *Free Rad. Biol. Med.* 18, 93–105. doi: 10.1016/0891-5849(94)00158-G

Zigha, A., Rosenfeld, E., Schmitt, P., and Duport, C. (2007). The redox regulator Fnr is required for fermentative growth and enterotoxin synthesis in *Bacillus cereus* F4430/73. *J. Bacteriol.* 189, 2813–2824. doi: 10.1128/JB.01701-06

Advantages of publishing in Frontiers

OPEN ACCESS

Articles are free to read, for greatest visibility

COLLABORATIVE PEER-REVIEW

Designed to be rigorous – yet also collaborative, fair and constructive

FAST PUBLICATION

Average 85 days from submission to publication (across all journals)

COPYRIGHT TO AUTHORS

No limit to article distribution and re-use

TRANSPARENT

Editors and reviewers acknowledged by name on published articles

SUPPORT

By our Swiss-based editorial team

IMPACT METRICS

Advanced metrics track your article's impact

GLOBAL SPREAD

5'100'000+ monthly article views and downloads

LOOP RESEARCH NETWORK

Our network increases readership for your article

Find us on