frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 09 March 2023
DOI 10.3389/fmicb.2023.1093372

Check for updates

OPEN ACCESS

EDITED BY
Penny Lea Morrill,
Memorial University of Newfoundland, Canada

REVIEWED BY
Mark Alexander Lever,
The University of Texas at Austin, United States
Jeffrey M. Dick,
Central South University, China

*CORRESPONDENCE
Lauren M. Seyler
✉ auren.seyler@stockton.edu

SPECIALTY SECTION
This article was submitted to Frontiers in Microbiology Microbiological Chemistry and Geomicrobiology, a section of the journal Frontiers in Microbiology

RECEIVED 08 November 2022
ACCEPTED 23 January 2023
PUBLISHED 09 March 2023

CITATION
Seyler LM, Kraus EA, McLean C, Spear JR, Templeton AS and Schrenk MO (2023) An untargeted exometabolomics approach to characterize dissolved organic matter in groundwater of the Samail Ophiolite. *Front. Microbiol.* 14:1093372. doi: 10.3389/fmicb.2023.1093372

COPYRIGHT
© 2023 Seyler, Kraus, McLean, Spear, Templeton and Schrenk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

An untargeted exometabolomics approach to characterize dissolved organic matter in groundwater of the Samail Ophiolite

Lauren M. Seyler[1,2,3]*, Emily A. Kraus[4,5], Craig McLean[6], John R. Spear[4], Alexis S. Templeton[7] and Matthew O. Schrenk[1,8]

[1]Department of Earth and Environmental Sciences, Michigan State University, East Lansing, MI, United States, [2]Biology Program, School of Natural Sciences and Mathematics, Stockton University, Galloway, NJ, United States, [3]Blue Marble Space Institute of Science, Seattle, WA, United States, [4]Department of Civil and Environmental Engineering, Colorado School of Mines, Golden, CO, United States, [5]Department of Environmental Engineering, University of Colorado, Boulder, Boulder, CO, United States, [6]Massachusetts Institute of Technology, Cambridge, MA, United States, [7]Department of Geological Sciences, University of Colorado, Boulder, Boulder, CO, United States, [8]Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI, United States

The process of serpentinization supports life on Earth and gives rise to the habitability of other worlds in our Solar System. While numerous studies have provided clues to the survival strategies of microbial communities in serpentinizing environments on the modern Earth, characterizing microbial activity in such environments remains challenging due to low biomass and extreme conditions. Here, we used an untargeted metabolomics approach to characterize dissolved organic matter in groundwater in the Samail Ophiolite, the largest and best characterized example of actively serpentinizing uplifted ocean crust and mantle. We found that dissolved organic matter composition is strongly correlated with both fluid type and microbial community composition, and that the fluids that were most influenced by serpentinization contained the greatest number of unique compounds, none of which could be identified using the current metabolite databases. Using metabolomics in conjunction with metagenomic data, we detected numerous products and intermediates of microbial metabolic processes and identified potential biosignatures of microbial activity, including pigments, porphyrins, quinones, fatty acids, and metabolites involved in methanogenesis. Metabolomics techniques like the ones used in this study may be used to further our understanding of life in serpentinizing environments, and aid in the identification of biosignatures that can be used to search for life in serpentinizing systems on other worlds.

KEYWORDS

metabolomics, serpentinization, metagenomics, DOM, alkaliphiles, methanogenesis, biosignatures

Introduction

Serpentinization is the geochemical process by which ultramafic rock reacts with water, resulting in the production of serpentine minerals and hydrogen gas (H_2) (Sleep et al., 2004). High concentrations of H_2 can further drive the reduction of dissolved inorganic carbon (DIC) to reduced carbon compounds including formate, methane, and carbon monoxide (McCollom and Seewald, 2001; Charlou et al., 2002; Proskurowski et al., 2008; McCollom, 2013). Chemolithotrophic microorganisms are able to use the products of these reactions as sources of carbon and reducing power (Schrenk et al., 2004, 2013; McCollom, 2007; Amend et al., 2011; Brazelton et al., 2011, 2012, 2013; Lang et al., 2012; Cardace et al., 2015). Serpentinization may be a widespread process in other locations in our solar system, including the subsurface of Mars, Enceladus, and Europa, making it a potential source of chemical disequilibria that may be harnessed by extraterrestrial life (Schulte et al., 2006; Vance et al., 2007; Klein et al., 2019; McCollom et al., 2022). Environments associated with serpentinization host microbial communities capable of a variety of metabolic strategies, including methanogenesis (Miller et al., 2016; Brazelton et al., 2017; Rempfert et al., 2017; Fones et al., 2019; Kraus et al., 2021), methane oxidation (Miller et al., 2016; Brazelton et al., 2017; Seyler L. M. et al., 2020; Kraus et al., 2021), acetogenesis (Brazelton et al., 2012; Rempfert et al., 2017; Suzuki et al., 2018; Fones et al., 2019; Seyler L. M. et al., 2020; Coleman et al., 2022), and sulfate reduction (Sabuda et al., 2020; Glombitza et al., 2021; Nothaft et al., 2021; Templeton et al., 2021). However, the highly alkaline conditions resulting from the serpentinization reaction, coupled with the precipitation of available DIC to calcite at high pH (Barnes et al., 1978), creates a challenging environment for microbial life (Schrenk et al., 2013). The physiological strategies used by microbial communities to overcome these challenges are not well understood.

Metabolomics is a rapidly advancing tool in the field of microbial ecology (Bauermeister et al., 2021; Oyedeji et al., 2021; van Santen et al., 2021). The detection and annotation of metabolites using mass spectrometry has allowed researchers to demonstrate the existence of complete metabolic pathways not commonly detected by genome annotation (Tang et al., 2009), to describe the function of theoretical pathways (Peyraud et al., 2009), and to discover entirely novel pathways (Fürch et al., 2009; Liu et al., 2016). When integrated with other -omics approaches, including metagenomics and metatranscriptomics, metabolomics is a powerful tool for describing metabolic activity in microbial communities (Turnbaugh and Gordon, 2008; Beale et al., 2016; Porcar et al., 2018). Metabolomic profiling of microbial communities in extreme (e.g., hyperalkaline) environments can assist in understanding the metabolic pathways used by microbial communities to adapt to extreme conditions, and the mechanisms and degree to which metabolites may be preserved under these conditions (Blachowicz et al., 2019). Metabolomics thus holds great potential in the search for biomarkers with astrobiological applications (Seyler L. et al., 2020).

The extracellular mixture of compounds that constitutes the dissolved organic matter pool contains growth factors, substrates, and chemical signals that represent the fundamental relationship between microbes and their environment (Kujawinski, 2011; Douglas, 2020). Most primary metabolites tend to be retained within cells due to their charge at physiological pH values (Bar-Even et al., 2011). These metabolites are usually involved in multiple metabolic pathways and are common across clades of organisms (Peregrín-Alvarez et al., 2009). Secondary metabolites, however, are often more species-specific, and more likely to diffuse into the surrounding environment (Breitling et al., 2013; Covington et al., 2016). Secreted metabolites also act as a chemical language for microorganisms, whose transmission is in constant flux due to evolutionary and environmental pressures (Soldatou et al., 2019; Douglas, 2020). The secretion of metabolites into the surrounding environment is a clear reflection of microbial activity, as the production of these compounds relies on the functional operation of metabolic pathways (Kell et al., 2005). Exometabolomics, or "metabolomic footprinting" (Kell et al., 2005), thus allows for the direct characterization of the molecular interaction between microbes and their environment (Sogin et al., 2019, Douglas, 2020). Metabolomics techniques, in conjunction with metagenomic data, have previously been used to describe the pool of intracellular and extracellular metabolites produced by microbial communities in the Coast Range Ophiolite Microbial Observatory (CROMO), a series of wells that provide access to a serpenization-influenced aquifer in the Coast Range of northern California (Seyler L. M. et al., 2020).

In this study, we used exometabolomics to characterize dissolved organic matter derived from microbial metabolic activity in the Samail Ophiolite groundwaters in the Sultanate of Oman, which is the largest and best exposed ophiolite in the world. The Samail Ophiolite consists of a large (~15,000 km^3) block of ocean crust and upper mantle (Nicolas et al., 2000) that was rapidly emplaced onto the Arabian continental plate approximately 70 million years ago (Glennie and Clarke, 1973; Coleman, 1981; Tilton et al., 1981; Hacker et al., 1996). Here, mafic and ultramafic rocks that once underwent alteration in seafloor hydrothermal systems continue to be altered by active serpentinization on the continent (Canovas et al., 2017; Keleman et al., 2021). We used an untargeted metabolomics approach (Patti et al., 2012) to characterize compounds in the dissolved organic matter (DOM) pool of groundwaters collected from varied geochemical conditions. Metagenomic data acquired from biomass filtered from the same samples was used to verify the annotation of metabolomic features. To our knowledge, this is the first such study attempting to describe biomolecules that make up the DOM pool in the Samail Ophiolite groundwater.

Materials and methods

Site description

The Samail Ophiolite provides a complete cross-section through 7 km of oceanic crust and 15 km of underlying upper mantle rocks (Boudier and Coleman, 1981; Coleman and Hopson, 1981; Lippard et al., 1986; Glennie, 1996; Nicolas et al., 2000). Partially serpentinized peridotite in the ophiolite actively undergoes hydration and carbonation reactions at low temperatures (≤60?), producing hyperalkaline fluids enriched in dissolved methane and hydrogen gas in the subsurface (Barnes et al., 1978;

Neal and Stanger, 1983, 1985; Clark and Fontes, 1990; Kelemen and Matter, 2008; Kelemen et al., 2011, 2013; Paukert et al., 2012; Streit et al., 2012; Miller et al., 2016). The ophiolite's size, exposure, and accessibility create unique opportunities for studying a broad range of geologic and hydrologic conditions supporting subsurface microbial life (Rempfert et al., 2017). Previous geochemical studies of the Samail Ophiolite groundwater observed a strong partitioning between two subgroups of peridotite-hosted fluids and classified them as "Type I/alkaline" (pH 8–10) or "Type II/hyperalkaline" (pH > 10), while wells positioned at the "contact" between peridotite and gabbro displayed more geochemical variability (Rempfert et al., 2017; **Supplementary Figure 1**). Alkaline waters have greater availability of oxidants and DIC and more meteoric influences than hyperalkaline, while contact waters appear to be a mixture of the two (Nothaft et al., 2021). We continue to use these designations to refer to well types throughout this study.

Sampling

In February 2017, a submersible pump was used to collect water samples from five previously drilled wells in the Samail Ophiolite (**Table 1**). Approximately 100 L of flush water were pumped from at least 20 M below the air-water interface in each well prior to sample collection. Sample volume varied according to the amount of water that could be acquired from each well while continuously pumping: 20 L were collected from NSHQ14, 7 L from WAB71, 18 L from WAB55, 4 L from WAB104, and 20 L from WAB105. Each sample was filtered through 0.3 μM glass fiber filters in a muffled stainless steel inline filter holder using PTFE tubing. Approximately 30 ml filtrate was quenched and stored in liquid nitrogen for transport back to the laboratory, where total organic carbon analysis was performed on a Shimadzu total organic carbon analyzer. The rest of the filtrate was acidified to pH 3 using concentrated HCl for solid phase extraction in the field. Biomass for DNA extraction was collected using in-line 0.2 μM polycarbonate filters as described previously (Fones et al., 2019). A total of 10 ml samples of unfiltered waters from each well were preserved in the field for enumeration of planktonic cells by adding formaldehyde to a final concentration of 10% vol/vol, and transported back to the laboratory for direct cell counting *via* 4′,6-diamidino-2-phenylindole (DAPI) staining on black polycarbonate 0.2 μM filters. Water temperature, conductivity, pH, oxidation-reduction potential (ORP), and dissolved oxygen concentration were measured during sampling using a Hach HQ40D Portable Multi Meter (Fones et al., 2019).

Metabolomics

Dissolved organics were captured from the acidified filtrate on a solid phase extraction Bond Elut-PPL cartridge (Agilent, Santa Clara, CA, USA) (Dittmar et al., 2008; Fiore et al., 2015; Seyler L. M. et al., 2020). 1/8″ × 1/4″ PTFE tubing was used to pull the supernatant into the cartridge to minimize the possibility of contamination from plastic leaching, and discarded flow-through was removed *via* Viton tubing with a peristaltic pump (flow rate not exceeding 40 ml/min). The cartridges were then rinsed with at least two cartridge volumes of 0.01 M HCl to remove salt, and the sorbent dried with air for 5 min. Dissolved organic matter (DOM) was eluted with 2 ml 100% methanol *via* gravity flow into muffled amber glass vials. These extracts were then shipped back to the laboratory at Michigan State University, where they were dried down *via* vacuum centrifugation, and resuspended in 495 μL 95:5 water:acetonitrile and 5 μL 5 μg/ml biotin-(ring-6,6-d_2) standard (Rabinowitz and Kimball, 2007; Fiore et al., 2015). This protocol was repeated for 4 L milliQ water as an extraction blank. In addition, an extraction solvent blank was run consisting of 495 μL of 95:5 water:acetonitrile plus 5 μL 5 μg/ml biotin-(ring-6,6-d_2).

Samples were analyzed *via* quadrupole time-of-flight liquid chromatography tandem mass spectrometry (QToF-LC/MS/MS) on a Waters Xevo G2-XS UPLC/MS/MS instrument at the MSU Metabolomics Core facility. Each sample was separated chromatographically in an Acquity UPLC BEH C18 column (1.7 μM, 2.1 mm × 50 mm) using a polar/non-polar gradient made up of 10 mm TBA and 15 mm acetic acid in 97:3 water:methanol (solvent A), and 100% methanol (solvent B). The gradient was run at 99.9/0.1 (% A/B) to 1.0/99.0 (% A/B) over 9 min, held an additional 3 min at 1.0/99.0 (% A/B), then reversed to 99.9/0.1 (% A/B) and held another 3 min. At a rate of 400μL/min, the sample was fed into a quadrupole time-of-flight mass spectrometer using a Waters Xevo G2-XS MS/MS in negative ion mode using a data independent collection method (m/z acquisition range 50–1,500 Da). Mass spec data were then imported into Progenesis QI software for statistical analysis and correlation. Quality control was performed by checking ion intensity maps and producing alignments of all sample runs. An aggregate data set was then created from the aligned runs to produce a map of compound ions that was applied to each sample for peak matching. Compound ions were deconvoluted and ion abundance measurements normalized so that runs could be compared. This approach yielded a table of peak areas that are identified by a unique mass to charge (m/z) and retention time (rt) pair referred to as features. Multivariate statistics was then performed on the data including principal components

TABLE 1 Geochemical data of wells sampled for metabolomics.

Site	UTM easting	UTM northing	Lithology	Pump depth (m)	pH	T (°C)	Cond (μS/cm)	ORP (mV)	Cell counts (per ml)	TOC (mg/L)	SPE volume (L)	TOC of sample (mg)
NSHQ14	675,495	2,529,716	Peridotite	50	11.1	34.4	493	−415	3.04E + 05	1.4470	20	28.94
WAB71	670,322	2,533,981	Peridotite	50–70	10.6	34.5	1803	−86	2.79E + 05	2.5130	7	17.591
WAB55	634,777	2,506,101	Contact	30	9.2	34.7	1171	110	6.79E + 04	0.3560	18	6.408
WAB104	643,099	2,541,124	Peridotite	70	8.5	33.4	493	180	2.38E + 05	0.6146	20	12.292
WAB105	644,678	2,536,524	Peridotite	50	8.3	31.6	448	178	1.80E + 05	0.5756	4	2.3024

analysis and correlation analysis. A table of relative abundances of features was generated for further analysis. Features found in the blanks were subtracted from the sample data (Broadhurst et al., 2018). In order to be considered present in a sample, a feature had to have at least a two-fold greater abundance in the sample than in both blanks. Features were putatively annotated in Progenesis QI against the ChemSpider database using mass, charge, retention time, and fragmentation spectra. Further annotation of features was achieved using mummichog (Li et al., 2013), which predicts functional activity by compiling metabolic networks using mass spec data. Whenever possible, we used autonomous METLIN-Guided In-source Fragment Annotation (MISA) (Domingo-Almenara et al., 2019), an algorithm that compares detected features against low-energy fragments from MS/MS spectra in the METLIN database (Smith et al., 2005), to increase our confidence in the putative annotation of representative metabolites; however, most of the putative metabolites did not have MS/MS spectral data available in the database. Raw spectrometry data are available in the MetaboLights database (Haug et al., 2020) under study identifier MTBLS6081.

Metagenome binning

DNA extraction, generation of sequencing libraries, and analysis of metagenomic sequence data, including binning, was described previously (Fones et al., 2019; Kraus et al., 2021).

Bins were taxonomically classified using GTDB-Tk (Chaumeil et al., 2020). Bin quality was assessed using CheckM version 1.0.11 (Parks et al., 2015); bins with 50% or greater completion were selected for further downstream analyses. The relative abundance of each bin in the metagenomes was calculated using a competitive recruitment approach. A bowtie2 index was built from a concatenated file of all bin contigs using bowtie2-build (Langmead and Salzberg, 2012). Quality-filtered metagenome reads were mapped against the bin bowtie2 index using bowtie2 with the –no-unal flag. The resulting SAM files (containing only the reads which mapped) were converted to BAM files, then BamM[1] was used to remove all alignments with <95% identity and <75% alignment coverage. The number of reads in each remaining BAM file was counted using Binsanity-profile (v0.3.3) (Graham et al., 2017). The read counts were then used to calculate the normalized relative fraction of the metagenomes that mapped to the bins using the following equation:

$$\frac{\frac{Reads}{bp}\,per\ genome}{\sum \frac{Reads}{bp}\,all\ genome} \times \frac{\sum Recruited\ reads\ to\ genomes}{total\ reads} \times 100$$

Z-scores of the normalized relative fractions were calculated using the average and sample standard deviation of the relative fraction of each bin across all metagenomes. The bin relative abundance data was run through the Non-negative Matrix Factorization (NMF) package (Gaujoux and Seoighe, 2010) in R to run a standard non-negative matrix factorization algorithm based on the Kullback–Leibler divergence (Brunet et al., 2004). This algorithm generated a heatmap of the bins across samples and a dendrogram using a Pearson correlation distance function. Bins that were representative of particular wells or geochemical conditions were selected using the method defined by Carmona-Saez et al. (2006). MetaSanity (Neely et al., 2020) was used to annotate functional orthologies against the KEGG database (Kanehisa and Goto, 2000); these annotations were run through KEGGDecoder (Graham et al., 2018) to determine the completeness of metabolic pathways of interest. Genes and pathways not detected by KEGGDecoder were searched for in the bins using protein sequences available in UniProt (UniProt Consortium, 2021) and the Conserved Domain Database (Lu et al., 2020), the MUSCLE alignment tool (Edgar, 2004), and the hmmsearch function in HMMER v 3.3.2 (Eddy, 2008).

Cross-referencing metabolomic and metagenomic data

Metabolomic feature abundance and bin abundance were standardized for ordination according to Legendre and Gallagher (2001) using the decostand function in the R vegan package (v2.4-2).[2] Regularized Discriminant Analysis (RDA) was performed on the standardized data using the klaR package (v0.6-14).[3] These RDAs were then subjected to Procrustes rotation using function Procrustes with symmetric scaling.

Feature annotations acquired using mummichog were cross-referenced with intermediate compounds from pathways identified in representative bins. As for the bins, NMF analysis was performed, and a heatmap of features across samples and a dendrogram using a Pearson correlation distance function were generated. Features that were representative of particular wells or geochemical conditions were extracted using the procedure described by Kim and Park (2007).

Results

Total organic carbon (TOC) was highest in WAB71 (2.5130 mg/L), followed by NSHQ14 (1.4470 mg/L), WAB104 (0.6146 mg/L), WAB105 (0.5756 mg/L), and WAB55 (0.3560 mg/L) (**Table 1**). A total of 2,483 features were detected in the mass spectrometry data across all wells (**Supplementary Table 1**). Principal coordinate analysis of the mass spectrometry data grouped WAB71, WAB55, WAB104, and WAB105 together and set NSHQ14 apart from the other wells (**Figure 1**). A Pearson correlation of feature abundance across wells followed by NMF analysis of feature abundance using the standard algorithm (Brunet et al., 2004) and a factorization rank of three clustered the wells into three "levels": NSHQ14, WAB71, and the alkaline/contact wells (WAB104, WAB105, WAB55) (**Figure 2**). A factorization rank of two resulted in WAB71 being clustered with WAB104, WAB105, and WAB55 (data not shown). NMF analysis also identified features which were representative of (i.e., most abundant in and most likely to be found in) each level. In our analyses, 48 features were designated as representative of NSHQ14, only one

1 https://github.com/Ecogenomics/BamM

2 https://www.rdocumentation.org/packages/vegan

3 https://www.rdocumentation.org/packages/klaR

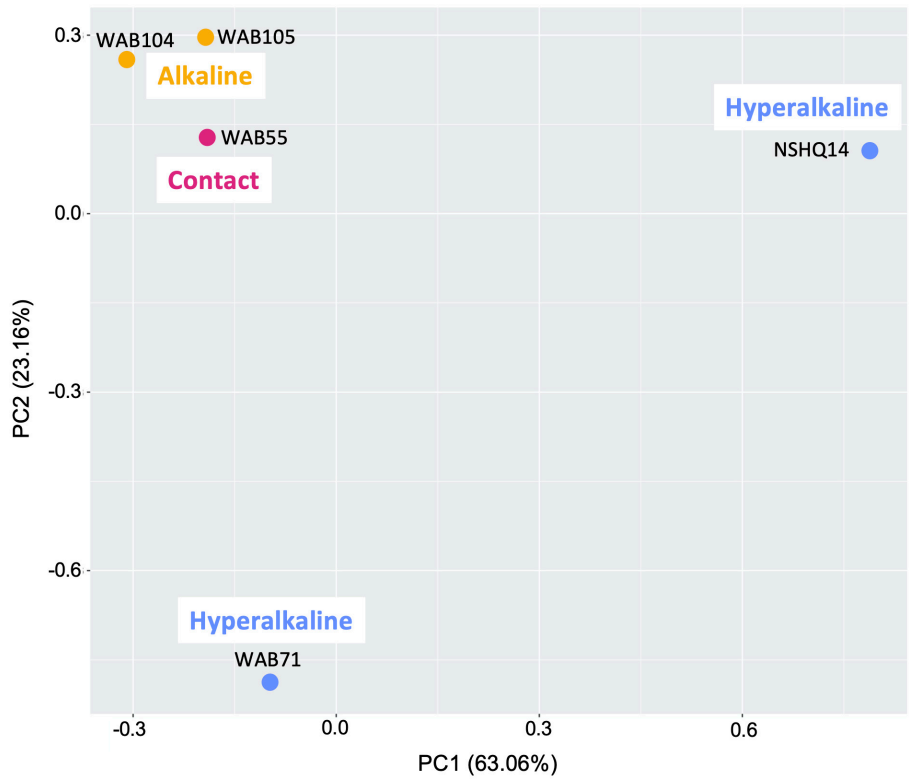

FIGURE 1
Principal component analysis comparing metabolomes of five sampled wells and blanks. The gray area is a cloud of all metabolomic features across all samples.

of which could be annotated by mummichog; two features were representative of WAB71, neither of which could be annotated by either Progenesis QI or mummichog; and the remaining wells had 90 representative features, eight of which could be annotated by mummichog (Supplementary Table 2). We also searched the mass spectrometry data for features which were only detected in one well (unique features). NSHQ14 had the greatest number (200) of unique features, WAB71 had 49, WAB104 had 15, WAB105 had five, and WAB55 had no unique features; none of these unique features could be annotated by either Progenesis QI or mummichog (data not shown).

In contrast to the metabolomic features, metagenomic bins tended to cluster by well (Figure 3). NMF analysis clustered the bins according to fluid type, and identified bins which were representative of (i.e., most abundant in and most likely to be found in) the hyperalkaline (NSHQ14 and WAB71), alkaline (WAB104 and WAB105), or contact (WAB55) wells. Superimposing the ordination plots of feature abundance and bin abundance data using Procrustes analysis showed that variance in the DOM pool could be attributed to variance in bin abundance and that the wells clustered according to fluid type (Procrustes sum of squares = 0.5033, correlation coefficient = 0.7048, $p = 0.375$), with the exception of WAB71, whose DOM pool clustered with the alkaline wells (Figure 4). Procrustes analysis also distinguished the DOM pool of WAB55 from those of NSHQ14 and WAB71, WAB 104, and WAB105 (Figure 4).

A total of 1,625 of the 2,483 features in the mass spectrometry data were putatively annotated by Progenesis QI, by matching feature mass to compounds in the ChemSpider database. After removing features that appeared in the extraction blanks, the list of features was reduced to 953, 206 of which could be annotated by Progenesis QI. In contrast, 119 of these 953 features were able to be annotated by mummichog, which annotates metabolomic data by using network analysis and metabolite prediction to resolve ambiguity in metabolite annotation. Because Progenesis QI relies on mass matches to the ChemSpider database, which includes synthetic and theoretical compounds in addition to naturally-occurring metabolites, we primarily relied on the mummichog annotations for our analyses (We have included the Progenesis QI annotations of representative compounds in Supplementary Table 2). We cross-referenced metabolic pathways annotated in the representative bins by KEGGDecoder with the features that could be annotated by mummichog to select features in the data on which to further focus our analyses. We then used the MetaCyc database (Caspi et al., 2014) to identify specific reaction pathways involving these metabolites and confirmed the presence of the relevant genes in the annotated MAGs. Pathways of interest and their annotated features in the mass spectrometry data are outlined below.

Carbohydrates and organic acids

Putative annotations of mass spectrometry features included glycan sugars (L-fuculose, L-rhamnose, L-rhamnulose) and their degradation products (L-fucono-1,5-lactone and 2-dehydro-3-deoxy-L-rhamnonate), sugar alcohols, polysaccharides, butanoate,

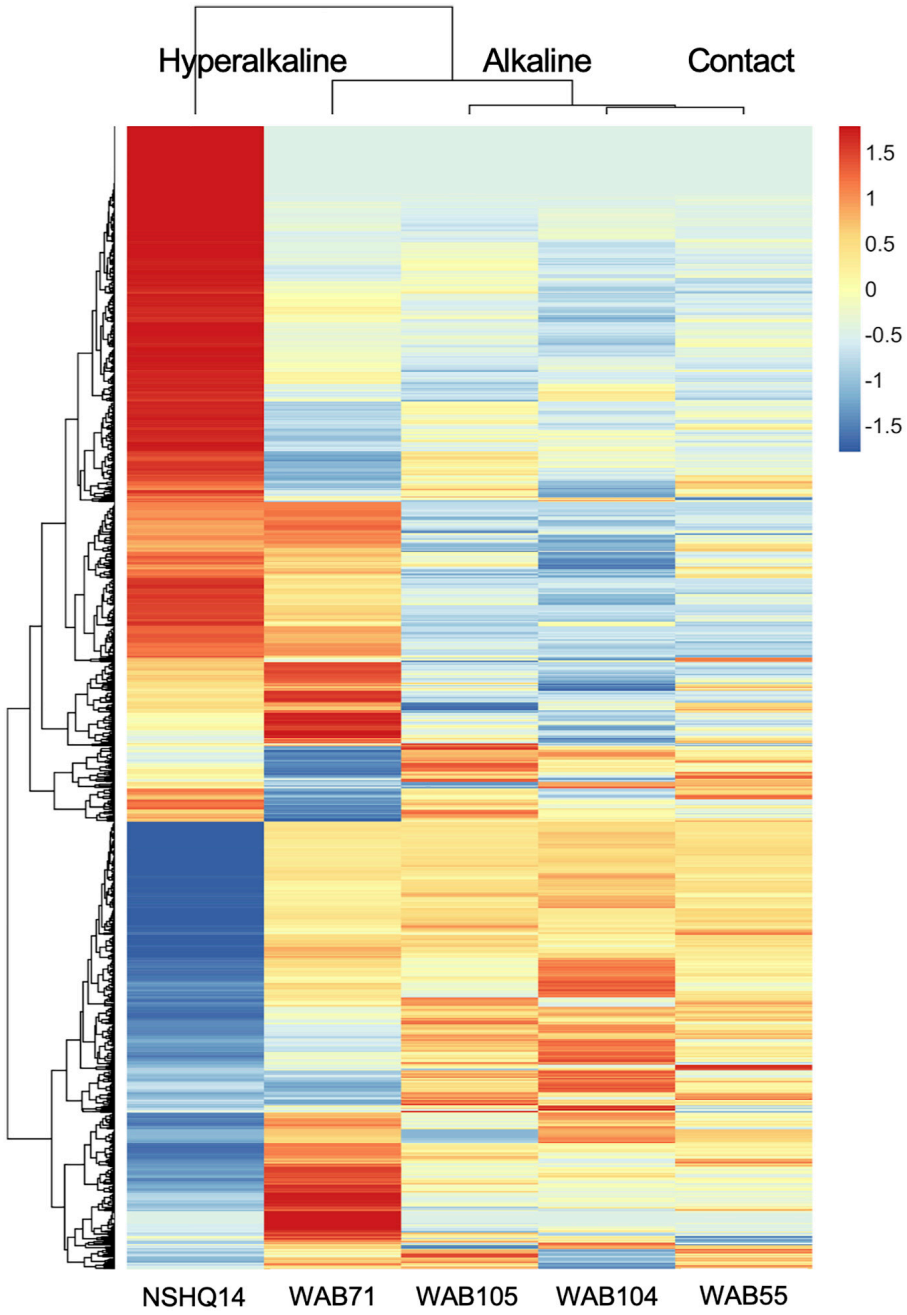

FIGURE 2
Z-scores of metabolomic feature abundance across wells clustered using a Pearson correlation, indicate peaks in the mass spectrometry data are differentially more abundant in some wells than in others. In particular, NSHQ14 had many unique features, most of which could not be identified.

and isocitrate (Figure 5). We further confirmed the annotation of isocitrate using MISA.

Isocitrate was more abundant in the hyperalkaline wells than in the contact or alkaline wells. Isocitrate is an intermediate of multiple metabolic pathways, most notably the citric acid (TCA) cycle, the reverse TCA (rTCA) carbon fixation pathway, the glyoxylate cycle, and coenzyme B biosynthesis in methanogens. A partial TCA cycle was annotated in nearly every bin generated from the metagenomic data, but the rTCA pathway was not identified in any of the bins (Figure 6). The glyoxylate cycle, which uses five of the eight enzymes of the TCA cycle to synthesize macromolecules from two-carbon compounds (e.g., acetate), was annotated in 25 bins, three of which were representative bins of the alkaline wells, three which were representative of the contact well, and one which was representative of the hyperalkaline wells (Figure 6). A complete pathway for coenzyme B biosynthesis was detected in only one bin, a *Methanobacterium* that was representative of NSHQ14 (Figure 6).

Butanoate was also most abundant in the hyperalkaline wells. Butanoate may be produced by the fermentation of pyruvate, but while fermentation pathways were found in 31 (23 representative) bins (Figure 6), the complete set of enzymes required for pyruvate

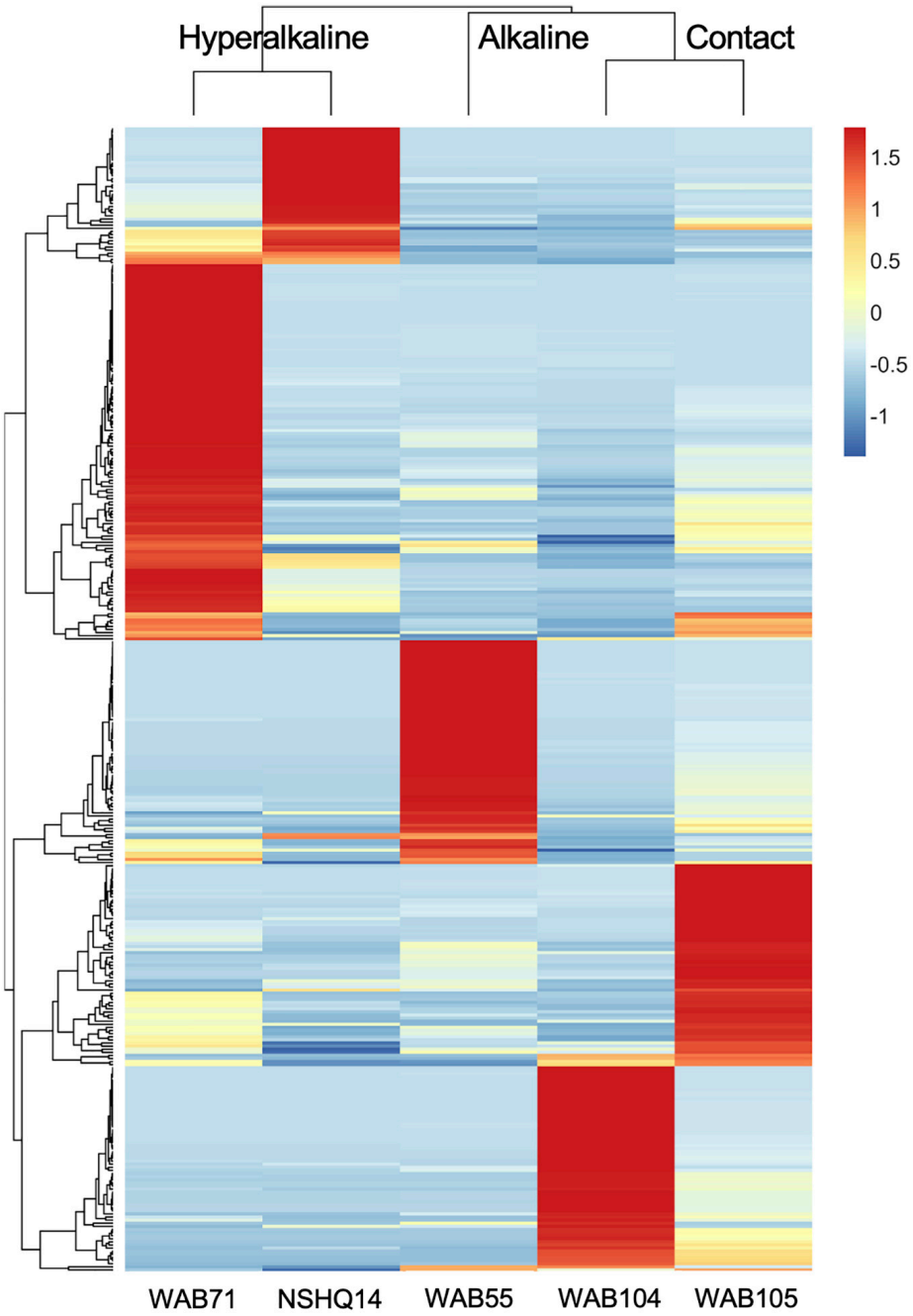

FIGURE 3
Z-scores of bin abundance across wells, clustered using a Pearson correlation. Wells clustered according to fluid type: hyperalkaline (NSHQ14 and WAB71), alkaline (WAB104 and WAB105), and contact (WAB55). WAB71 had the greatest number of bins which were more abundant in this well than any other well, while NSHQ14 had the fewest.

fermentation to butanoate were not detected in any of the bins (Supplementary Table 3).

A complete pathway for L-fuculose or L-rhamnose degradation was not annotated in any of the bins (Supplementary Table 3). The gene for D-threo-aldose 1-dehydrogenase, which produces L-fucono-1,5-lactone during L-fucose degradation, was only annotated in three bins, but the gene for L-fuconolactonase, which further degrades L-fucono-1,5-lactone to L-fuconate, was annotated in fifteen bins (Supplementary Table 3). Similarly, the gene for the L-rhamnonate dehydratase enzyme that produces 2-dehydro-3-deoxy-L-rhamnonate was only annotated in two bins, but the enzyme that acts on this product, 2-dehydro-3-deoxy-L-rhamnonate dehydrogenase, was annotated in seven bins (Supplementary Table 3).

Chitobiose, a disaccharide produced by the hydrolysis of chitin, was also annotated in the mass spectrometry data (Figure 5). MISA further confirmed the annotation of chitobiose; however, we question the validity of the annotation as the peaks eluted very far apart, suggesting that they did not arise from the same parent molecule (Supplementary Figure 1). Genes

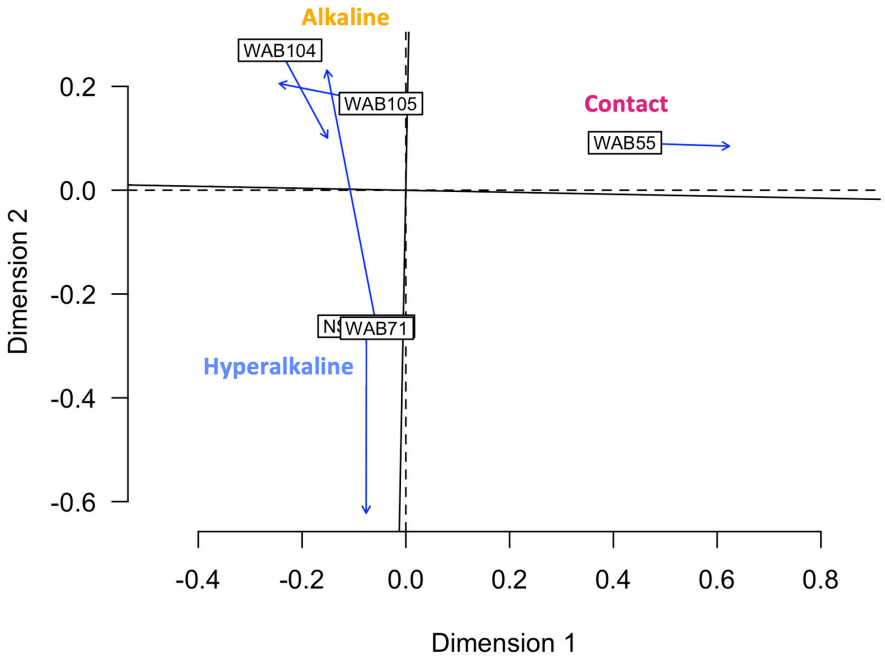

FIGURE 4
Procrustes analysis in which the metabolomic feature abundance and bin abundance ordinations are optimally superimposed through rotating and rescaling, showing that metabolomic feature abundance and bin abundance are correlated. Hyperalkaline wells and alkaline wells cluster together, while the contact well is distinct from both. Labeled boxes indicate the position of the samples in the first ordination, and arrows point to their positions in the target ordination. The dashed lines represent the axes of the first ordination (bin abundance); the solid lines represent the axes of the target ordination (metabolomic feature abundance).

for chitinase and diacetylchitobiose deacetylase, enzymes that hydrolyze chitin and chitobiose, were identified in 11 and 16 bins, respectively, though only two bins contained both (Figure 6). Beta-*N*-acetylhexosaminidase, an exoglycosidase that catalyzes the hydrolysis of chitobiose, was detected in 54 out of 128 total bins, 14 of which were representative bins (six of the seven alkaline representative bins, three hyperalkaline representative bins, and five contact representative bins) (Figure 6). Genes for enzymes involved in the catabolism of other polysaccharides, including alpha-amylase, glucoamylase, beta-glucosidase, D-galacturonate epimerase, D-galacturonate isomerase, oligogalacturonide lyase, and pullulanase were also detected in the metagenomic data, and tended to be associated with bins that displayed the greatest abundance in the hyperalkaline wells (Supplementary Table 3).

Methanogenesis/propanoate degradation

Five features were putatively annotated as either *cis*-homoaconitate, an intermediate in coenzyme B biosynthesis, or *cis*-2-methylaconitate/2-methyl-trans-aconitate, intermediates in propanoate degradation to pyruvate and succinate *via* the 2-methylcitrate cycle (Figure 5). These features were more abundant (by up to eight orders of magnitude) in the hyperalkaline wells compared to the contact and alkaline wells. An intermediate of coenzyme M production, (2R)-phosphosulfolactate, was also putatively annotated in the mass spectrometry data and was detected in all of the wells, though it was most abundant in the alkaline wells and was designated a representative feature of the alkaline/contact wells by NMF analysis (Figure 5). A feature putatively annotated as 2-hexaprenyl-6-methoxyphenol, an intermediate in methanogenesis from methoxylated aromatic compounds, was almost exclusively found in NSHQ14 (Figure 5). The enzyme that produces this intermediate has not been characterized.

Complete pathways for methanogenesis *via* CO_2 and coenzyme M reduction to methane were found in two out of the 128 bins analyzed (Figure 6; Fones et al., 2019). Both of these bins were taxonomically annotated as Euryarchaeota belonging to the Methanobacteriales; one bin was designated a representative bin for NSHQ14, and was one of the most abundant bins in that well (Figure 6; Fones et al., 2019). Both bins also contained the acetoclastic methanogenesis pathway, as well as a nearly complete pathway for methanogenesis from methoxylated aromatic compounds, except for the gene for the first enzyme in the pathway, *mtoB* (Figure 6 and Supplementary Table 3). The *mtoB* gene was detected in 51 other bins; 80% of these bins displayed the greatest abundance in the hyperalkaline or contact wells (Supplementary Table 3). A complete 2-methylcitrate cycle for propanoate degradation was detected in four bins, including a representative bin of WAB71 (a member of the Microtrichales) and a representative bin of WAB105 (*Acinetobacter junii*) (Figure 6).

Porphyrins, pigments, and quinones

Fourteen features in the mass spec data were annotated as intermediates of pigment, porphyrin, and quinone biosynthesis

	Mass/Charge	Retention Time (s)	Putative Annotation(s)	Pathways	NSHQ-14 (Hyperalkaline)	WAB-71 (Hyperalkaline)	WAB-55 (Contact)	WAB-104 (Alkaline)	WAB-105 (Alkaline)
Carbohydrates and Organic Acids	173.0090522	413.99	Isocitrate*	TCA cycle, Glyoxylate cycle, Coenzyme B biosynthesis					
	181.071362	23.209	Mannitol, D-Sorbitol	Sugar alcohols					
	213.1126111	455.914	Butanoic acid	Fermentation product of polysaccharides					
	213.1128715	469.475	Butanoic acid	Fermentation product of polysaccharides					
	251.0769583	533.946	Dextrin, Maltotriose, Manninotriose	Polysaccharides					
	143.0344967	335.43	L-Fucono-1,5-lactone 2-Dehydro-3-deoxy-L-rhamnonate	L-fuculose degradation L-rhamnose degradation					
	143.034647	428.701	L-Fucono-1,5-lactone 2-Dehydro-3-deoxy-L-rhamnonate	L-fuculose degradation L-rhamnose degradation					
	145.0502616	162.542	L-Fuculose, L-Rhamnose, L-Rhamnulose 2-Dehydropantoate (S)-2-Aceto-2-hydroxybutanoate	Glycan sugars Pantothenate and CoA biosynthesis Isoleucine biosynthesis					
	425.140347	489.908	Chitobiose*	Amino sugar metabolism					
Methanogenesis / Propanoate Degradation	169.0135249	456.972	cis-Homoaconitate cis-2-Methylaconitate, 2-Methyl-trans-aconitate	Coenzyme B biosynthesis Propanoate degradation					
	169.0136102	323.892	cis-Homoaconitate cis-2-Methylaconitate, 2-Methyl-trans-aconitate	Coenzyme B biosynthesis Propanoate degradation					
	169.0137395	476.347	cis-Homoaconitate cis-2-Methylaconitate, 2-Methyl-trans-aconitate	Coenzyme B biosynthesis Propanoate degradation					
	169.0137904	466.924	cis-Homoaconitate cis-2-Methylaconitate, 2-Methyl-trans-aconitate	Coenzyme B biosynthesis Propanoate degradation					
	169.0137981	300.815	cis-Homoaconitate cis-2-Methylaconitate, 2-Methyl-trans-aconitate	Coenzyme B biosynthesis Propanoate degradation					
	248.945677	**463.752**	(2R)-Phosphosulfolactate	Coenzyme M biosynthesis					
	265.2063864	462.694	2-Hexaprenyl-6-methoxyphenol	Methanogenesis from methoxylated aromatic compounds					
Porphyrins, Pigments, and Quinones	280.1222442	54.125	Protoporphyrin IX	Heme and chlorophyll biosynthesis					
	329.1495666	542.312	Coproporphyrinogen III, Coproporphyrinogen I	Precursor to protoporphyrin IX					
	163.0331047	490.965	2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate	Menaquinone biosynthesis					
	239.0555698	280.382	(1R,6R)-6-Hydroxy-2-succinylcyclohexa-2,4-diene-1-carboxylate	Menaquinone biosynthesis					
	137.0239831	383.605	4-Hydroxybenzoate	Ubiquinol biosysnthesis					
	356.2346793	564.86	(2'S)-Deoxymyxol 2'-alpha-L-fucoside	Carotenoid biosynthesis					
	593.3469851	567.504	Glycosyl-4,4'-diaponeurosporenoate	Staphyloxanthin (carotenoid) biosynthesis					
	457.4201912	**558.08**	Tetraprenyl-beta-curcumene	Sesquarterpene found in Bacillus spp.					
	449.3408343	345.911	Phylloquinone	Photosystem I electron transport chain					
	451.356674	598.854	Phylloquinol	Phylloquinone biosynthesis					
	565.2085338	402.452	Magnesium protoporphyrin	Chlorophyll biosynthesis					
	595.2181431	414.519	13(1)-Hydroxy-magnesium-protoporphyrin IX 13-monomethyl ester	Chlorophyll biosynthesis					
	196.0613066	**306.102**	3,4-Dihydroxy-L-phenylalanine, N,N-Dihydroxy-L-phenylalanine	Plant/fungi pigments					
	455.2351908	458.557	Phytyl diphosphate	Geranylgeranyl diphosphate biosynthesis					
Fatty Acids and Sterols	325.0977032	463.752	Dolichyl diphosphate	N-Glycan biosynthesis					
	325.0979834	476.347	Dolichyl diphosphate	N-Glycan biosynthesis					
	325.0978361	500.389	Dolichyl diphosphate	N-Glycan biosynthesis					
	365.091095	507.261	Dolichyl diphosphate	N-Glycan biosynthesis					
	219.174499	**539.14**	30-oxolanosterol	Sterol biosynthesis (methylotrophs)					
	181.1591554	531.303	Dodecanoic acid	Fatty acid biosynthesis					
	181.159206	424.472	Dodecanoic acid	Fatty acid biosynthesis					
	183.1380953	**552.793**	Dodecanoic acid	Fatty acid biosynthesis					
	199.1697832	**565.918**	Dodecanoic acid	Fatty acid biosynthesis					
	221.1534627	**559.045**	Dodecanoic acid	Fatty acid biosynthesis					
	259.1908329	543.37	Dodecanoic acid	Fatty acid biosynthesis					
	253.2166859	576.399	(9Z)-Hexadecenoic acid (palmitoleate)	Fatty acid biosynthesis					
	193.123585	**544.956**	12-Oxo-9(Z)-dodecenoic acid	alpha-Linolenic acid metabolism					
	231.1232484	455.386	9-Oxononanoic acid, 1,2-dibutyrin	alpha-Linolenic acid metabolism					
	183.101987	504.089	Prostaglandin G2	Arachidonic acid metabolism					
	349.2015361	554.908	Prostaglandin G2	Arachidonic acid metabolism					
	333.2081752	559.045	Prostaglandin H2*, Prostaglandin D2*	Arachidonic acid metabolism					
	351.2190235	560.631	Prostaglandin H2*, Prostaglandin D2*	Arachidonic acid metabolism					
Amino Acid and Polyamine Metabolism	**259.2374446**	**543.37**	Agmatine	Putrescine biosynthesis / Arginine degradation					
	297.0648519	403.509	5'-S-Methyl-5'-thioinosine	Methionine salvage / Aromatic amino acid biosynthesis					
	170.9969508	**329.178**	Isethionic acid	Sulfoacetaldehyde degradation					
	170.9967908	**334.901**	Isethionic acid	Sulfoacetaldehyde degradation					
	173.0450841	228.071	Shikimate 2-Isopropylmaleate, (2S)-2-Isopropyl-3-oxosuccinate	Chorismate biosynthesis Leucine biosynthesis					
	173.0451555	151.004	Shikimate 2-Isopropylmaleate, (2S)-2-Isopropyl-3-oxosuccinate	Chorismate biosynthesis Leucine biosynthesis					
	157.0501747	282.497	alpha-Isopropylmalate, 2-Isopropylmaleate, 3-Isopropylmalate	Leucine biosynthesis					
	157.0501959	239.609	alpha-Isopropylmalate, 2-Isopropylmaleate, 3-Isopropylmalate	Leucine biosynthesis					
	152.0349219	375.239	2,3,4,5-Tetrahydrodipicolinate 3-Hydroxyanthranilate	Lysine biosynthesis Tryptophan degradation					
	138.0191247	290.863	2-Aminomuconate	Tryptophan degradation					
	135.0445777	401.923	Phenylacetic acid	Phenylalanine metabolism					
	135.0445681	209.224	Phenylacetic acid	Phenylalanine metabolism					
	165.0552258	240.574	3-(3-Hydroxyphenyl)propanoic acid	Phenylalanine metabolism					
	151.0395899	317.64	3-(3-Hydroxyphenyl)propanoic acid	Phenylalanine metabolism					
Miscellaneous	159.0657347	223.313	Pimelate	Biotin biosynthesis					
	193.0366249	383.076	7-Carboxy-7-carbaguanine	preQ0 biosynthesis					

FIGURE 5

Features cross-referenced between KEGG results and mummichog annotations. Features are grouped by putative annotation/function. Diagnostic features appear in bold. Feature annotations confirmed by MISA are indicated with an asterisk (*). Feature abundance is expressed as the log of the normalized relative abundance.

pathways (Figure 5). A feature putatively annotated as protoporphyrin IX, a central metabolite in the production of heme and chlorophyll, was detected in all wells and was particularly abundant (3–5 orders of magnitude greater) in NSHQ14 (Figure 5). Another feature, putatively annotated as coproporphyrinogen III (a precursor to protoporphyrin IX) was detected in every well but displayed the lowest abundance in NSHQ14, by more than two orders of magnitude (Figure 5). Complete pathways for the biosynthesis of heme from glutamate and/or glutamyl-tRNA were annotated in bins detected in every well; however, bins containing complete or nearly complete (≥90%) heme pathways were generally more abundant in the hyperalkaline and contact wells (Figure 6).

Three features annotated as intermediates in ubiquinol (4-hydroxybenzoate) and menaquinone synthesis (2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate and (1R,6R)-6-hydroxy-2-succinylcyclohexa-2,4-diene-1-carboxylate, intermediates in the synthesis of menaquinol and phylloquinol from isochorismate) were detected in all of the wells but were most abundant in NSHQ14 (Figure 5). Genes for the enzymes

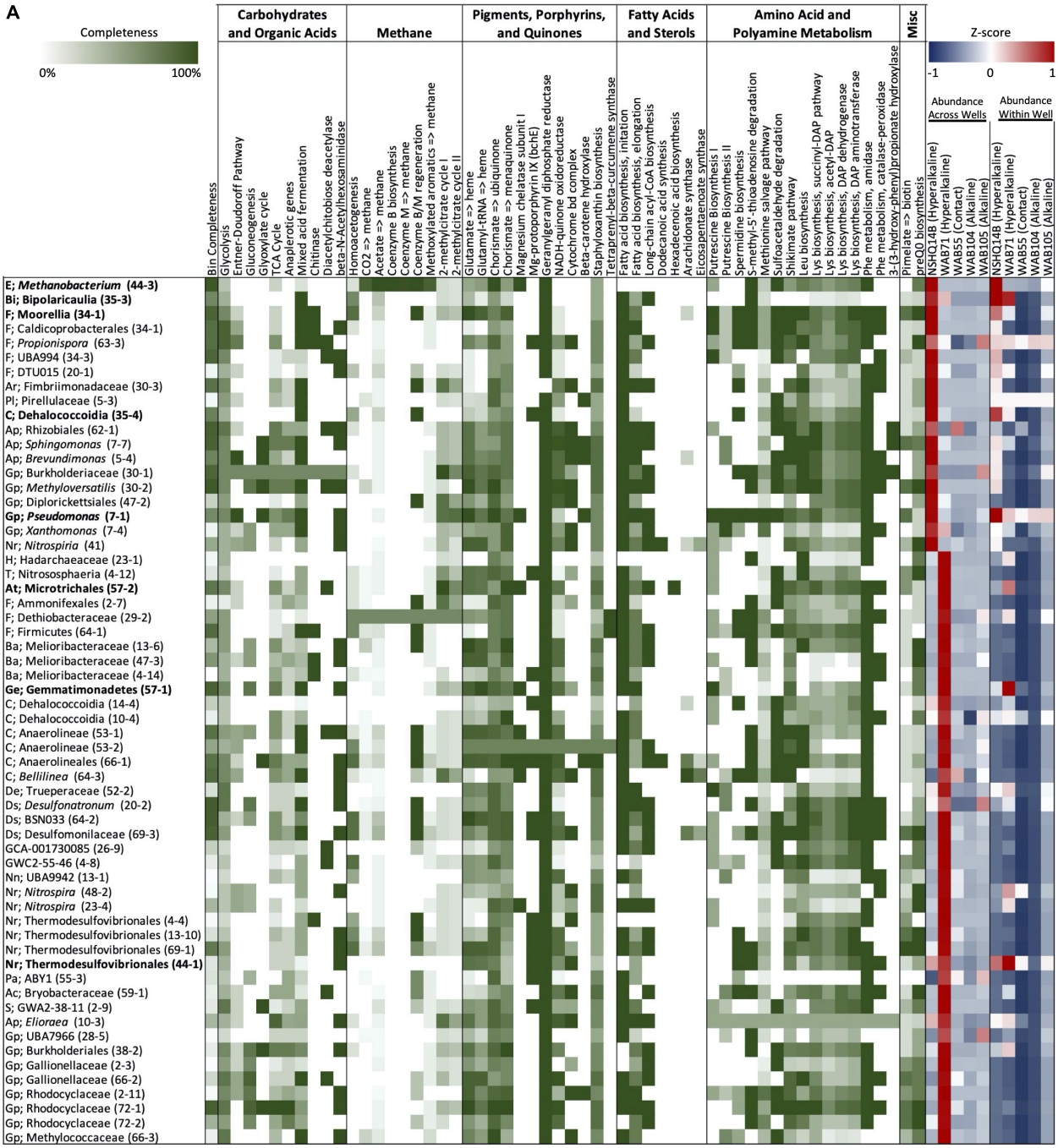

FIGURE 6
(Continued)

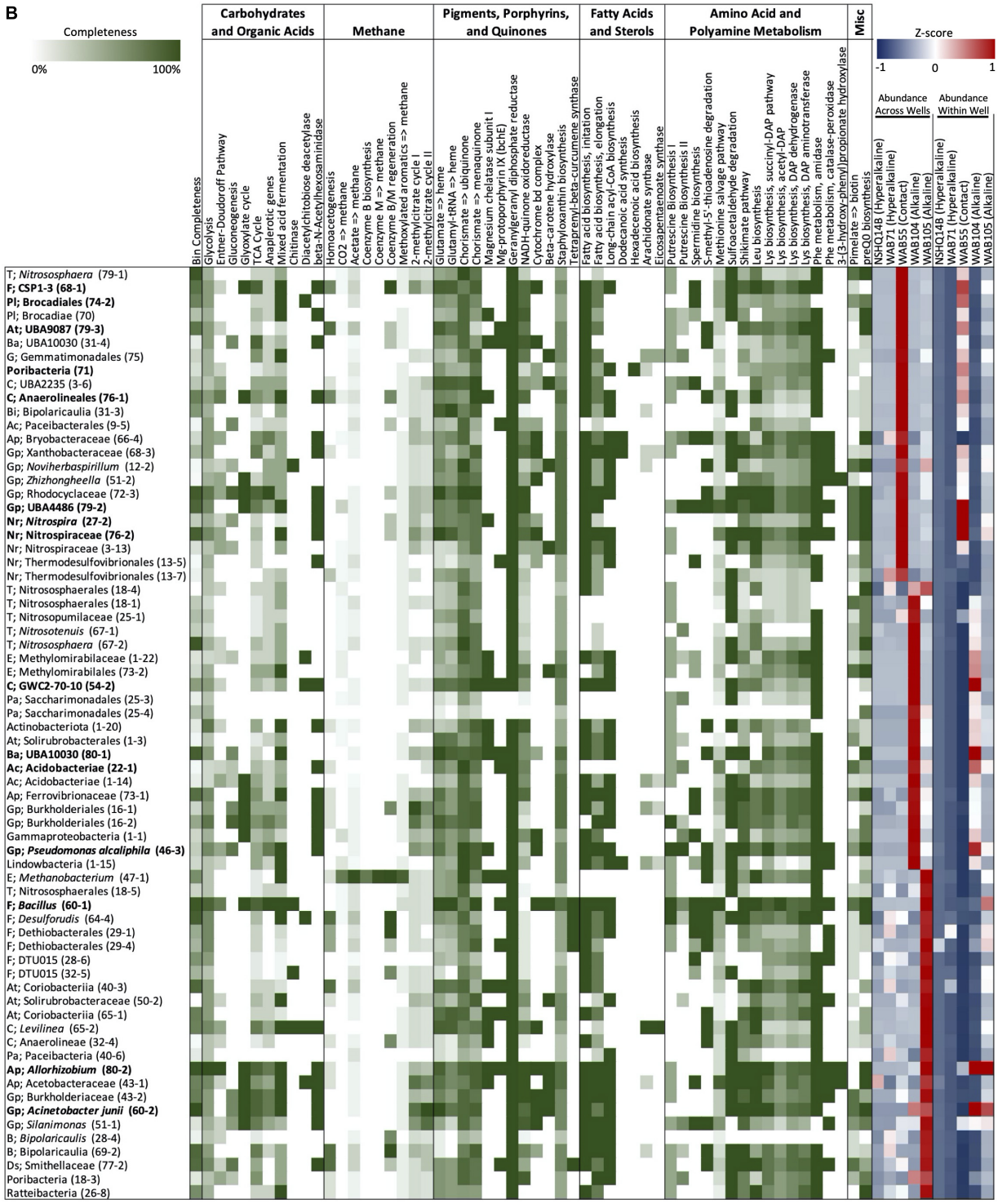

FIGURE 6
(A,B) Metabolic genes and pathways of interest, annotated in each bin using KEGGDecoder and HMMER. Percentage completion of each pathway is indicated in shades of green. Z-scores of bin abundance within each well and across all wells are indicated in a blue/red heatmap at the top of the figure (negative z-scores in blue, positive z-scores in red).

that produce these intermediates were detected in 14, 11, and 6 bins, respectively; these bins tended to be most abundant in the hyperalkaline or contact wells (Supplementary Table 3). The complete pathway for ubiquinone biosynthesis was detected in thirteen bins, including two representative bins of NSHQ14 (*Methanobacterium* and *Pseudomonas*) and two representative bins of WAB104 (a member of the Acidobacteria, and *Pseudomonas alcaliphila*) (Figure 6). The complete pathway for menaquinone biosynthesis was detected in eleven bins, including a representative bin of every well except NSHQ14 (although one representative bin

of that well, a member of the Moorellia, had 90% of the pathway) (Figure 6). The most commonly found cytochrome genes in the bins encoded for the cytochrome bd complex and NADH-quinone oxidoreductase. A total of 41 of the 128 bins contained ≥90% of the genes encoding one or both of these cytochromes (Figure 6).

Features putatively annotated as (2′S)-deoxymyxol 2′-alpha-L-fucoside and glycosyl-4,4′-diaponeurosporenoate, metabolites produced in the biosynthesis of the carotenoids myxol 2′-dimethyl-fucoside and staphyloxanthin, respectively, were also detected in all the wells (Figure 5). The biosynthesis of (2′S)-deoxymyxol 2′-alpha-L-fucoside is a multi-step reaction that has not been fully characterized, but we were able to annotate the enzyme that catalyzes the conversion of (2′S)-deoxymyxol 2′-alpha-L-fucoside to myxol 2′-dimethyl-fucoside, beta-carotene hydroxylase, in 18 bins (Figure 6). Genes in the pathway for staphyloxanthin biosynthesis were annotated in nearly every bin, but the complete pathway was only detected in one bin, a member of the Anaerolineales that was most abundant in WAB71 (Figure 6).

A feature found exclusively in NSHQ14 and designated as a representative feature of that well was putatively annotated as tetraprenyl-beta-curcumene, an intermediate in the biosynthesis of a sesquiterpene found in *Bacillus* spp. (Figure 5). The gene for tetraprenyl-beta-curcumene synthase was detected in seven bins, two of which were most abundant in WAB71, and the rest of which were most abundant in WAB105 (Figure 6).

Two features that were among the most abundant features in the mass spec data were annotated as intermediates of chlorophyll metabolism: magnesium protophorin and 13(1)-hydroxy-magnesium-protoporphyrin IX 13-monomethyl ester (Figure 5). Other putatively annotated features associated with photosynthesis and chlorophyll metabolism include phylloquinone, phytyl diphosphate, and phylloquinol (Figure 5). A feature annotated as either 3,4-dihydroxy-L-phenylalanine or N,N-dihydroxy-L-phenylalanine, precursors to pigments/cytochromes in fungi and plants, was also detected in all of the wells and was a representative feature of the alkaline/contact wells (Figure 5). Most of these features were relatively more abundant in the alkaline and contact wells. However, the feature putatively annotated as phylloquinone was exclusively detected in the hyperalkaline wells and was particularly abundant in NSHQ14 (Figure 5). Because the pathway for phylloquinol/phylloquinone synthesis and menaquinol synthesis involve the same/highly homologous enzymes, we could not distinguish between these pathways in the bins.

The complete pathway for chlorophyll biosynthesis was not detected in any bin (Supplementary Table 3). However, genes for two enzymes in this pathway, magnesium chelatase subunit I and anaerobic magnesium-protoporphyrin IX monomethyl ester cyclase, were annotated in 27 bins and 29 bins, respectively (Figure 6). Magnesium chelatase is a three-subunit enzyme responsible for inserting a Mg^{+2} into protoporphyrin IX in the first step of chlorophyll biosynthesis; genes for the other two subunits (H and D) were found in two and five bins, respectively, and none of the bins contained the genes for all three subunits of the enzyme (Supplementary Table 3). Anaerobic magnesium-protoporphyrin IX monomethyl ester cyclase produces 13(1)-hydroxy-magnesium-protoporphyrin IX 13-monomethyl ester in the chlorophyll biosynthesis pathway found in anaerobic photosynthetic bacteria. None of the bins contained the complete suite of genes required for photosynthesis, but two bins (one diagnostic) in the alkaline wells did contain the genes for an anoxygenic type-II reaction center (Supplementary Table 3). Thirteen bins contained the genes for the RuBisCo enzyme, nine of which also contained the complete Calvin–Benson–Bassham pathway for carbon fixation, but none of which included the anoxygenic type-II reaction center (Supplementary Table 3).

A feature putatively annotated as phytyl diphosphate, an intermediate of geranylgeranyl diphosphate biosynthesis, was also detected in all the wells. Geranylgeranyl diphosphate is a precursor to phylloquinol, chlorophyll, archaeol, and multiple pigments across all domains of life.

Fatty acids and sterols

Eighteen features were annotated as fatty acids and intermediates in fatty acid pathways, five of which were designated characteristic features of the alkaline and contact wells by NMF analysis. These included four peaks annotated as dolichyl diphosphate, lipids involved in N-glycan production in archaea, which were detected in all the wells but were most abundant in NHSQ14 and WAB71. We also detected a feature putatively annotated as 12-oxo-9(Z)-dodecenoic acid (an intermediate of alpha-linolenic acid metabolism), a feature putatively annotated as 30-oxolanosterol (an intermediate of sterol biosynthesis in methylotrophs), and three features annotated as dodecanoic acid (Figure 5). There were also three features found in all the wells but particularly abundant in WAB71, which were annotated as prostaglandins by mummichog. MISA further confirmed the annotation of the latter two features as prostaglandin H2 or prostaglandin D2 (Figure 5 and Supplementary Figure 1). However, we suspect that this annotation is incorrect and that these features are more likely intermediates or secondary metabolites of arachidonic acid and/or eicosapentaenoic acid (see section "Discussion").

Genes for polyunsaturated fatty acid (PUFA) synthases (arachidonate and/or eicosapentaenoate synthase) were annotated in 18 of the bins analyzed in this study; however, a complete PUFA synthase operon (*pfaABCD*) was only found in two bins, one taxonomically annotated as belonging to the genus *Bellilinea* and the other as genus *Levilinea*, which are in the family Anaerolineaceae (phylum Chloroflexota, class Anaerolineae, order Anaerolineales) (Figure 6 and Supplementary Table 3). The *pfaE* gene that encodes the phosphopantetheinyl transferase required for catalytic activation of acyl carrier protein domains was not present in these bins, but was found in two other bins belonging to phylum Chloroflexota, class Dehalococcoidia (Supplementary Table 3). Genes for fatty acid synthesis initiation and elongation (*fabB, fabD, fabF, fabG, fabH, fabI, fabZ*) and long chain acyl-CoA biosynthesis (*ACSL, fadD*), which produce the precursors to arachidonate and eicosapentaenoate, were annotated in all but two of the representative bins (Figure 6).

Amino acid and polyamine metabolism

Thirteen features were putatively annotated as intermediates of amino acid/polyamine biosynthesis or degradation pathways

(Figure 5). Most of these features were detected in all of the wells, with a few exceptions. A feature annotated as agmatine, an intermediate of arginine degradation and putrescine/spermidine biosynthesis, was present in every well except NSHQ14, and was designated a representative feature of the alkaline/contact wells by NMF analysis (Figure 5). Complete putrescine biosynthesis pathways were annotated in 12 bins; interestingly, nearly half of these bins displayed the greatest abundance in NSHQ14 (Figure 6). Conversely, a feature annotated as 5′-S-methyl-5′-thioinosine, a metabolite produced by the deamination of 5′-methylthioadenosine (an intermediate of the methionine salvage cycle and spermidine biosynthesis), was detected in all the wells but was particularly abundant in NHSQ14 (Figure 5). The gene for enzyme that performs the deamination reaction, 5-methylthioadenosine/S-adenosylhomocysteine deaminase, was detected in 50 of the 128 bins analyzed in this study, including four representative bins of NSHQ14, one representative bin of WAB71, three representative bins of WAB55, two representative bins of WAB104, and one representative bin of WAB105 (Figure 6). This enzyme has also been implicated in the recycling of 5′-deoxyadenosine and the biosynthesis of aromatic amino acids in methanogens (Miller et al., 2014). The methionine salvage cycle is also linked to the production of putrescine and spermidine. Only four bins contained ≥90% of the methionine salvage pathway, but 60 bins contained the portion of the pathway that results in 5′-methylthioadenosine/spermidine biosynthesis (Figure 6).

Two representative features of NSHQ14 were putatively annotated as isethionic acid, a product of sulfoacetaldehyde degradation. These features were found only in NSHQ14 and WAB105, but at eight orders of magnitude greater abundance in the former versus the latter well (Figure 5). Sulfoacetaldehyde is both a product of taurine degradation and an intermediate of coenzyme M biosynthesis. The gene for NADH-dependent sulfoacetaldehyde reductase, the enzyme that catalyzes the conversion of sulfoacetaldehyde to isethionic acid, was detected in 77 bins, including four representative bins from NSHQ14, one from WAB71, five from WAB55, three from WAB104, and two from WAB105 (Figure 6).

Two features annotated as either shikimate, 2-isopropylmaleate, or (2S)-2-isopropyl-3-oxosuccinate were detected in all of the wells but were most abundant in the alkaline and contact wells (Figure 5). Shikimate is an intermediate of the shikimate pathway for the production of chorismate, a precursor to aromatic amino acid and menaquinone/ubiquinone biosynthesis. 2-Isopropylmaleate and (2S)-2-Isopropyl-3-oxosuccinate are intermediates of leucine biosynthesis. Two other features, annotated as alpha-isopropylmalate/2-isopropylmaleate/3-isopropylmalate, all of which are intermediates in leucine biosynthesis, were also detected in all the wells but most abundant in the alkaline wells (Figure 5). Complete shikimate and leucine biosynthesis pathways were detected in bins in every well (Figure 6). Nine of the 15 representative bins of the contact and alkaline wells contained one or both of these pathways, but only one of the eight representative hyperalkaline bins contained these pathways (Figure 6).

Other putatively annotated features related to amino acid metabolism include intermediates of lysine biosynthesis (2,3,4,5-tetrahydrodipicolinate), tryptophan degradation (3-hydroxyanthranilate and 2-aminomuconate), and phenylalanine metabolism [phenylacetic acid and 3-(3-hydroxyphenyl)propanoic acid] (Figure 5). Lysine biosynthesis pathways were detected in multiple bins (Figure 6). The complete pathway for tryptophan degradation *via* 3-hydroxyanthranilate to 2-aminomuconate was annotated in only one bin, which was taxonomically annotated as a member of the Bipolaricaulia (Supplementary Table 3). This bin was more abundant in WAB105 than any other well, but had a low relative abundance compared to other bins in that well. The gene *amiE*, which encodes the amidase enzyme that converts 2-phenylacetamide to phenylacetic acid, was detected in almost every bin; the gene *katG*, which encodes the catalase-peroxidase enzyme that converts phenylalanine to 2-phenylacetamide was detected in 43 bins (Figure 6). The enzyme in phenylalanine metabolism that produces 3-(3-hydroxyphenyl)propanoic acid has not been characterized, but the gene for 3-(3-hydroxy-phenyl)propionate hydroxylase, which converts 3-(3-hydroxy-phenyl)propionate acid to 3-(2,3-dihydroxyphenyl)propanoate in the same pathway, was detected in seven bins, only one of which was designated a representative bin of a well (WAB105) (Figure 6).

Miscellaneous

A feature putatively annotated as pimelate, a precursor of biotin, was detected in every well, with a slightly greater relative abundance in the alkaline wells (Figure 5). The complete pathway for biotin synthesis from pimelate was annotated in only one bin, a member of the Thermodesulfovibrionales that displayed the greatest relative abundance in WAB71 (Figure 6). Twenty-one bins possessed four out of the five genes necessary for the pathway but were missing the gene for the enzyme 6-carboxyhexanoate-CoA ligase (*bioW*), which initiates the pathway by converting pimelate to pimeloyl-CoA (Supplementary Table 3). Two other bins contained this gene but otherwise lacked the complete pathway (Supplementary Table 3). The mechanism for the biosynthesis of pimelate is currently unknown, though it is speculated to begin with malonyl-CoA (Manandhar and Cronan, 2017). This same study found that deletion of the *bioW* gene results in a biotin auxotrophic phenotype in *Bacillus subtilis*.

Another feature, putatively annotated as 7-carboxy-7-carbaguanine (the precursor to preQ0 in preQ0 biosynthesis), was also detected in every well but was most abundant in WAB55 and WAB105 (Figure 5). A total of 36 bins contained the complete pathway for preQ0 biosynthesis, including one representative bin for NSHQ14, four representative bins for WAB55, two representative bins for WAB104, and three representative bins for WAB105 (Figure 6).

Representative and abundant bins

Bins we identified as representative of alkaline wells belonged to the Bacillales (family Bacillaceae), Rhizobiales (family Rhizobiaceae), Pseudomonadales (*Acinetobacter junii* and *Pseudomonas alcaliphila*), Acidobacteria, Bacteroidota (UBA10030), and Chloroflexota A (Figure 6). The most

abundant bins in the alkaline wells also included members of the Thaumarchaeota, Actinobacteriota, Methylomirabilota, and Saccharimonadales (**Figure 6**). Representative hyperalkaline bins included members of the Thermodesulfovibrionales, Bipolaricaulia (formerly candidate phylum OP1), Chloroflexota (Dehalococcoidia), Moorellia, Methanobacteriales, Microtrichales, Gemmatimonadetes, and Pseudomonadales (*Pseudomonas aeruginosa*) (**Figure 6**). Members of the Deinococcales, Burkholderiales, and Nitrospirales were also among the most abundant bins in the hyperalkaline wells (**Figure 6**). Bins representative of WAB55 (the only contact well sampled in this study) included Nitrospirales, Brocadiales, Chloroflexota (Anaerolineales), Gammaproteobacteria (UBA4486), Firmicutes (CSP1-3), and Poribacteria (**Figure 6**). A complete list of the bins, including bin completeness, contamination, strain heterogeneity, and phylogenetic classification, is presented in **Supplementary Table 4**.

Discussion

Previous studies have demonstrated clear trends correlating microbial diversity and metabolic capability to rock/fluid type in the Samail Ophiolite (Rempfert et al., 2017; Fones et al., 2019, 2021; Kraus et al., 2021). The chemistry of the subsurface fluids reflects both hydrologic context (i.e., the extent of equilibration with the atmosphere or host rock, or potential for mixing with other fluids) and the geology of the host rock (Rempfert et al., 2017; Leong et al., 2021). Multitracer experiments indicate that the higher pH (>9.3) fluids are older than their less alkaline counterparts, and their chemistry reflects a greater extent of serpentinization in the host rock (Paukert Vankeuren et al., 2019). Microbial community composition correlates with fluid geochemistry, with gabbro and alkaline peridotite fluids displaying a greater richness of microorganisms than the low-diversity communities found in hyperalkaline peridotite (Rempfert et al., 2017). Due to the covariance of multiple geochemical parameters, it is difficult to determine which of these parameters drives these correlations (Rempfert et al., 2017). Here, we present a broad snapshot of the DOM pool in the Samail Ophiolite groundwater and identify some of the secondary metabolites produced by the microbial communities found there. The annotation of these features was supported by comparing putative compounds to the genes identified in the metagenomic bins.

Given that both microbial community structure and protein-coding gene abundance correlate with rock and fluid type (Rempfert et al., 2017; Fones et al., 2019, 2021; Kraus et al., 2021), we hypothesized that DOM composition would likewise correlate to these characteristics. Principal components analysis (**Figure 1**), NMF analysis (**Figure 2**), and procrustes analysis (**Figure 4**) comparing mass spec data across five wells lend support to this hypothesis. The DOM pools of NSHQ14, WAB71, and WAB55 are each distinguished from the other wells in these analyses, while the two alkaline wells analyzed (WAB104 and WAB105) contain DOM pools that are very similar to each other. TOC was highest in the hyperalkaline wells, lower in the alkaline wells, and lowest in the contact wells (**Table 1**). Correlating TOC to cell count resulted in an R^2 value of 0.503, indicating that TOC is only partially driven by the amount of biomass present. Assuming an average carbon content per cell of 65 fg C $cell^{-1}$ in the subsurface (Parkes et al., 1994), intact cells account for 0.00072–0.0025% of TOC; lower estimates of 14 fg C $cell^{-1}$ from subsurface marine sediments (Braun et al., 2016) yield a TOC percentage of 0.00016–0.00054% from intact cells.

We hypothesize that the relative shortage of bioavailable electron acceptors in the hyperalkaline wells leads to dissolved organics being degraded and remineralized more slowly in these environments (Fones et al., 2019). The higher TOC values in the hyperalkaline wells also may reflect the abiotic production of formate and/or small molecular weight organic compounds associated with serpentinization (Proskurowski et al., 2008; McCollom, 2013). For example, the features annotated as butanoic acid, which displayed the greatest abundance in the hyperalkaline wells, may be the result of abiotic organic synthesis (McCollom and Seewald, 2006), as we did not detect the pathway for butanoic acid production *via* fermentation in the metagenomic data. Similarly, relatively high abundances of features annotated as intermediates of propanoate degradation in the hyperalkaline wells raise the intriguing possibility that the abiotic production of propanoate could serve as a carbon source for the microbial communities found there (Shock and Schulte, 1998). Recent carbon analysis of serpentinized rocks obtained in the Oman Drilling Project shows that non-carbonate-carbon loadings are approximately 140–360 ppm (Ternieten et al., 2021). Thus, there may be multiple primary sources of organic carbon in this system: autotrophic carbon fixation and turnover of biomass, heterotrophic utilization of reduced carbon accumulated in the rock during prior water/rock interaction, or ongoing abiotic organic synthesis.

Metagenomic data from the hyperalkaline wells indicates an enrichment of functional genes involved in C1 metabolisms such as CO utilization, methanogenesis, and acetogenesis (Fones et al., 2019; Coleman et al., 2022). Our characterization of the DOC pool generally supports these findings. Features annotated as metabolites associated with methanogenesis tended to display greater relative abundance in the hyperalkaline wells, where both methane and methanogens have been detected by previous studies (Rempfert et al., 2017; Fones et al., 2019; Kraus et al., 2021). The features putatively annotated as intermediates of propanoate degradation (mentioned above) were also putatively annotated as *cis*-homoaconitate (**Figure 5**), an intermediate in the biosynthesis of coenzyme B, which reacts with coenzyme M to release methane during methanogenesis. The detection of isethionic acid in NSHQ14 may be linked to the degradation of sulfoacetaldehyde produced during coenzyme M biosynthesis, while the abundance of dolichyl diphosphate, an intermediate of N-glycan biosynthesis in archaea, may also be a result of the high abundance of methanogenic archaea in NSHQ14 (Fones et al., 2019, 2021).

The feature annotated as isocitrate was three orders of magnitude more abundant in the hyperalkaline wells than in the alkaline wells, and seven orders of magnitude higher in the hyperalkaline wells than in the contact well (**Figure 5**). Isocitrate has been implicated as a metabolic intermediate in some methanogens; for example, Methanosarcinales generates CO_2 and 2-oxoglutarate from isocitrate *via* a partial oxidative TCA cycle (Anderson et al., 2009). This activity may serve as a source of CO_2 for the production of methane by hydrogenotrophic methanogen

such as *Methanobacterium*, which do not contain the pathway for 2-oxoglutarate generation *via* isocitrate (Anderson et al., 2009), but are abundant and active in NSHQ14 (Rempfert et al., 2017; Kraus et al., 2021). A bin annotated as *Methanobacterium* was the most abundant bin in NSHQ14, and designated as a representative bin of NSHQ14 by NMF analysis (**Figure 6**). This bin did not contain the gene for the NAD-dependent isocitrate dehydrogenase required for CO_2 and 2-oxoglutarate generation *via* isocitrate, but 20 other bins did contain this gene, more than half of which were most abundant in the hyperalkaline wells (including one bin belonging to the Moorellia that was designated a representative bin of NSHQ14, and a likely homoacetogen) (**Supplementary Table 3**). A syntrophy relationship may exist between hydrogenotrophic methanogens and bacteria containing the partial oxidative TCA cycle, whereby the CO_2 generated through this pathway fuels methanogenesis. Further studies, perhaps involving tracer/microcosm experiments, will be necessary to confirm if isocitrate generates CO_2 for methanogenesis in the hyperalkaline wells.

Isocitrate is an important central metabolite in many processes, including glyoxylate and dicarboxylate metabolism, 2-oxocarboxylic metabolism, and the biosynthesis of signaling molecules and secondary metabolites. The glyoxylate cycle allows microorganisms to use C2 compounds such as acetate to satisfy cellular carbon requirements when simple sugars such as glucose are not available (Kornberg and Madsen, 1957; LaPorte et al., 1984). Microcosm experiments conducted in parallel to this study found that rates of assimilation/oxidation of acetate were higher than those of bicarbonate and carbon monoxide across all sampled wells, suggesting that acetate may be the preferred source of bioavailable carbon (Fones et al., 2019; Coleman et al., 2022).

Metagenomic data have also shown an enrichment in genes involved in anaerobic metabolisms in hyperalkaline waters, while the alkaline wells were enriched in respiratory and electron transport chain genes (Fones et al., 2019). In contrast to the previously published metagenomic study, we found that features annotated as quinones and their precursors were generally enriched in NSHQ14, and to a lesser extent WAB71, compared to the contact and alkaline wells, with the exception of coproporphyrinogen III and phylloquinol (**Figure 6**). The abundance of these intermediates in the deeper wells may be indicative of slower turnover rates due to the extremophilic nature of this environment, or they may serve as intermediates of a process not yet characterized.

The abundance of features associated with plant and fungi pigments, including chlorophyll, in the alkaline and contact wells may reflect the greater influence of meteoric water in these wells (Nothaft et al., 2021). However, while the complete pathways for chlorophyll biosynthesis were not detected in the metagenomic data, the genes for the production of the intermediates we detected in the metabolome were found in multiple bins (**Figure 6**). In general, features related to the biosynthesis of pigments including carotenoids, quinones, and the chlorophyll intermediates may be linked to oxidative stress. Increased production of carotenoids, for example, has been observed previously in both neutrophilic (Lan et al., 2010) and extremophilic microorganisms (Tian and Hua, 2010; Mandelli et al., 2012) under oxidative stress conditions.

Of particular interest is the detection of several features putatively annotated as prostaglandins. These features were detected in all the wells, but were two to six orders of magnitude more abundant in WAB71 than in the other wells. We immediately questioned the validity of this annotation, as prostaglandins are signaling molecules found in higher eukaryotes. MISA analysis suggested that these features were either prostaglandin H2 or prostaglandin D2, both of which are involved in the human inflammatory response. However, prostaglandin H2 has a half-life of 90–100 s at room temperature (Human Metabolome Database), while prostaglandin D2 is found chiefly in brain and mast cells. These factors make it unlikely that these features are the result of contamination of our samples *via* casual contact. They may be prostaglandins produced by plants, but their abundance in WAB71, one of the most alkaline and least surface-influenced of the wells, makes this likewise unlikely.

We propose that these features may instead be secondary metabolites of arachidonic (ARA) or eicosapentaenoic acid (EPA). These polyunsaturated fatty acids serve as precursors to prostaglandins in higher eukaryotes, but are also found in some (mostly marine) bacteria belonging to the Gammaproteobacteria, Bacteroidetes, and Cyanobacteria. Eicosapentaenoic acid was annotated in the mass spec data during analysis with Progenesis QI, but MISA did not confirm this annotation (data not shown). Bacterial biosynthetic mechanisms of long-chain polyunsaturated fatty acids (LC-PUFAs) are not well understood and the functions of these lipids in bacteria have not been experimentally confirmed (Yoshida et al., 2016). It has been suggested that the primary function of LC-PUFAs is the adjustment of cell membrane fluidity under conditions of low temperature and/or high pressure (Kato and Nogi, 2001). EPA is required for growth at low temperatures and high pressures by some *Shewanella* species (Wang et al., 2004; Kawamoto et al., 2009; Usui et al., 2012), but a study of an EPA-deficient strain of *Photobacterium profundum* showed that monounsaturated or branched chain fatty acids could compensate for the loss of EPA (Allen et al., 1999). EPA and other LC-PUFAs have also been implicated in resistance to reactive oxygen species and oxidative stress (Nishida et al., 2006a,b, 2007, 2010; Okuyama et al., 2008; Tilay and Annapure, 2012).

Four of the bins analyzed in this study contained ≥75% of the protein-coding genes required to produce arachidonate synthase and/or eicosapentaenoate synthase (**Figure 6**). Three of these bins were more abundant in WAB71 than any other well, and included two members of the Chloroflexota, order Anaerolineales (one of which was annotated *Bellilinea* sp.), and one member of the Desulfobacterota, family Desulfomonilaceae. The other bin was most abundant in WAB105, and was also a member of the Anaerolineales, *Levilinea* sp. The gene organization, order of genes, and composition of enzymatic domains in *pfa* gene clusters that produce LC-PUFAs appear to be highly conserved (Okuyama et al., 2007). *Pfa* gene homologs have also been identified in the genomes of 86 bacterial species belonging to 10 phyla, the products of which are largely unknown (Shulse and Allen, 2011). These homologs

can be classified into 20 different gene types, which show a significant correlation with the ecological and physiological characteristics of the bacteria possessing the genes, suggesting that characterizing these secondary lipids may provide valuable insight into the ecological adaptation and evolution of microorganisms (Shulse and Allen, 2011). Lipid biomarkers have previously been characterized in rock samples obtained from the Samail Ophiolite; however, this study did not screen for LC-PUFAs (Newman et al., 2020). The *pfa* genes we detected in these bins may be homologs that produce an as-of-yet uncharacterized fatty acid product (Shulse and Allen, 2011).

Conclusion

Characterization of the DOM pool in the Samail Ophiolite groundwater reveals the presence of numerous metabolic products and intermediates that may provide clues to the survival strategies of microbial communities in environments associated with serpentinization. Among the features we were able to annotate and confirm with metagenomic data, the pigments, porphyrins, fatty acids, sterols, and intermediates/byproducts of pathways associated with methanogenesis (such as coenzyme B and coenzyme M biosynthesis) were the compounds whose abundance was most strongly correlated to fluid type. In particular, pigments are a promising and intriguing biosignature of life in this system, due to both their specificity to particular processes and their longevity in the geological record (Schwieterman et al., 2015).

We chose methods for characterization that we hoped would allow us to capture as many DOM features as possible. These methods naturally have limitations. Many of the metabolic pathways of interest in these wells, in particular those involved with C1 metabolisms, consist of intermediate compounds that are involved in multiple overlapping pathways (Peregrín-Alvarez et al., 2009), many of which are too small to be detected using the liquid chromatography methods used in this study. These compounds are also cycled rapidly and retained within the cell (Bar-Even et al., 2011), precluding them from an exometabolome analysis such as this one. Additionally, many of the features we detected could not be annotated using available metabolite databases- a caveat which reflects both the under-characterization of microbial secondary metabolites on the whole, but also the uniqueness of the environment and the microbial communities that inhabit it. The annotation of features based on mass and fragmentation is also a potential source of error, and structural assignments based on these criteria may not be valid. Metagenomic data can be used to reinforce these annotations but does not provide absolute certainty.

Linking DOM composition to environmental characteristics has exciting implications for the search for potential biomarkers for life fueled by serpentinization, in particular because the greatest numbers of unique features were found in wells that are the most heavily influenced by serpentinization. Metabolomics studies of environments of astrobiological significance, such as serpentinizing rock, hold great promise in the search for biomarkers that may be used as evidence of life on other worlds (Seyler L. et al., 2020). Further efforts to characterize the unique features detected in hyperalkaline fluids in the Samail Ophiolite may thus provide researchers with the means to detect life in serpentinizing systems elsewhere in our Solar System, such as the subsurface of Mars or the icy moons of Jupiter and Saturn. Such techniques hold the possibility of uncovering biomarkers of processes and relationships that are representative of habitats influenced by serpentinization, aiding the search for life in similar environments on other worlds in our Solar System.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.mg-rast.org/mgmain.html?mgpage=overview&metagenome=mgm4795806.3, mgm.4795806.3, https://www.mg-rast.org/mgmain.html?mgpage=overview&metagenome=mgm4795807.3, mgm4795807.3, https://www.mg-rast.org/mgmain.html?mgpage=overview&metagenome=mgm4795808.3, mgm4795808.3, https://www.mg-rast.org/mgmain.html?mgpage=overview&metagenome=mgm4795804.3, mgm4795804.3, https://www.mg-rast.org/mgmain.html?mgpage=overview&metagenome=mgm4795810.3, mgm4795810.3, and https://www.ebi.ac.uk/metabolights/MTBLS6081. The data in the MetaboLights database has just been made publicly available.

Author contributions

LS collected the samples, carried out metabolite extractions, and performed metagenomic and metabolomic data analysis. EK performed metagenomic data analysis. CM performed metabolomic data analysis. JS and AT oversaw fieldwork operations. JS, AT, and MS supervised the work. LS wrote the manuscript with input from all authors. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by a grant from the NASA Astrobiology Institute (NNA15BB02A) to AT, JS, Eric Boyd, and MS and the Deep Carbon Observatory - Deep Life Community (#2015-14084) to MS.

Acknowledgments

We thank Kaitlin Rempfert, Eric Ellison, Laura Beuter, and Eric Boyd for assistance with sample collection, Juerg Matter for help with equipment acquisition, permitting, and sample export, and the Ministry of Regional Municipalities and Water Resources in the Sultanate of Oman for allowing sampling and export of well waters.

We also thank A. Daniel Jones and Anthony Schilmiller for their assistance with performing QToF-LC/MS/MS, and Elizabeth Kujawinski for her guidance regarding data analysis and interpretation.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1093372/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Map of sampled wells in the Samail Ophiolite. Figure adapted from Rempfert et al. (2017).

SUPPLEMENTARY TABLE 1
All features identified as representative of the three well types ("levels") determined by non-negative matrix factorization (NMF) analysis. Annotations acquired by mummichog are included where available. Features highlighted in bold were annotated by Progenesis QI using the ChemSpider Database (**Supplementary Table 2**).

SUPPLEMENTARY TABLE 2
Representative features that could be annotated by Progenesis QI using the ChemSpider database.

SUPPLEMENTARY TABLE 3
KOs for all genes annotated by KEGGDecoder in bins included in this study.

SUPPLEMENTARY TABLE 4
All bins included in this study, including bin completeness, contamination, strain heterogeneity, and phylogenetic annotation.

References

Allen, E. E., Facciotti, D., and Bartlett, D. H. (1999). Monounsaturated but not polyunsaturated fatty acids are required for growth of the deep-sea bacterium *Photobacterium profundum* SS9 at high pressure and low temperature. *Appl. Environ. Microbiol.* 65, 1710–1720. doi: 10.1128/AEM.65.4.1710-1720.1999

Amend, J. P., McCollom, T. M., Hentscher, M., and Bach, W. (2011). Catabolic and anabolic energy for chemolithoautotrophs in deep-sea hydrothermal systems hosted by different rock types. *Geochim. Cosmochim. Acta* 75, 5736–5748. doi: 10.1016/j.gca.2011.07.041

Anderson, I., Ulrich, L. E., Lupa, B., Susanti, W., Porat, I., Hooper, S. D., et al. (2009). Genomic characterization of Methanomicrobiales reveals three classes of methanogens. *PLoS One* 4:e5797. doi: 10.1371/journal.pone.0005797

Bar-Even, A., Noor, E., Flamholz, A., Buescher, J. M., and Milo, R. (2011). Hydrophobicity and charge shape cellular metabolite concentrations. *PLoS Comp. Biol.* 7:e1002166. doi: 10.1371/journal.pcbi.1002166

Barnes, I., O'Neil, J. R., and Trescases, J. J. (1978). Present day serpentinization in New Caledonia, Oman and Yugoslavia. *Geochim. Cosmochim. Acta* 42, 144–145. doi: 10.1016/0016-7037(78)90225-9

Bauermeister, A., Mannochio-Russo, H., Costa-Lotufo, L. V., Jarmusch, A. K., and Dorrestein, P. C. (2021). Mass spectrometry-based metabolomics in microbiome investigations. *Nat. Rev. Microbiol.* 20, 143–160. doi: 10.1038/s41579-021-00621-9

Beale, D. J., Karpe, A. V., and Ahmed, W. (2016). "Beyond metabolomics: a review of multi-omics-based approaches," in *Microbial Metabolomics*, eds D. Beale, K. Kouremenos, and E. Palombo (Cham: Springer), 289–312.

Blachowicz, A., Chiang, A. J., Elsaesser, A., Kalkum, M., Ehrenfreund, P., Stajich, J. E., et al. (2019). Proteomic and metabolomic characteristics of extremophilic fungi under simulated Mars conditions. *Front. Microbiol.* 10:1013. doi: 10.3389/fmicb.2019.01013

Boudier, F., and Coleman, R. G. (1981). Cross section through the peridotite in the Samail Ophiolite. *J. Geophys. Res.* 86, 2573–2592. doi: 10.1029/JB086iB04p02573

Braun, S., Morono, Y., Littman, S., Kuypers, M., Aslan, H., Dong, M., et al. (2016). Size and carbon content of sub-seafloor microbial cells at Landsort Deep, Baltic Sea. *Front. Microbiol.* 7:1375. doi: 10.3389/fmicb.2016.01375

Brazelton, W. J., Mehta, M. P., Kelley, D. S., and Baross, J. A. (2011). Physiological differentiation within a single-species biofilm fueled by serpentinization. *mBio* 2:e00127-11. doi: 10.1128/mBio.00127-11

Brazelton, W. J., Morrill, P. L., Szponar, N., and Schrenk, M. O. (2013). Bacterial communities associated with subsurface geochemical processes in continental serpentinite springs. *Appl. Environ. Microbiol.* 79, 3906–3916. doi: 10.1128/AEM.00330-13

Brazelton, W. J., Nelson, B., and Schrenk, M. O. (2012). Metagenomic evidence for H2 oxidation and H2 production by serpentinite-hosted subsurface microbial communities. *Front. Microbiol.* 2:268. doi: 10.3389/fmicb.2011.00268

Brazelton, W. J., Thornton, C. N., Hyer, A., Twing, K. I., Longino, A. A., Lang, S. Q., et al. (2017). Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy. *PeerJ* 5:e2945. doi: 10.7717/peerj.2945

Breitling, R., Ceniceros, A., Jankevics, A., and Takano, E. (2013). Metabolomics for secondary metabolite research. *Metabolites* 3, 1076–1083. doi: 10.3390/metabo3041076

Broadhurst, D., Goodacre, R., Reinke, S. N., Kuligowski, J., Wilson, I. D., Lewis, M. R., et al. (2018). Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* 14:72. doi: 10.1007/s11306-018-1367-3

Brunet, J. P., Tamayo, P., Golub, T. R., and Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4164–4169. doi: 10.1073/pnas.0308531101

Canovas, P. A., Hoehler, T., and Shock, E. L. (2017). Geochemical bioenergetics during low-temperature serpentinization: an example from the Samail ophiolite, Sultanate of Oman. *J. Geophys. Res. Biogeosci.* 122, 1821–1847. doi: 10.1002/2017JG003825

Cardace, D., Meyer Dombard, D. R., Woycheese, K. M., and Arcilla, C. A. (2015). Feasible metabolisms in high pH springs of the Philippines. *Front. Microbiol.* 6:10. doi: 10.3389/fmicb.2015.00010

Carmona-Saez, P., Pascual-Marqui, R., Tirado, F., Carazo, J., and Pascual-Montano, A. (2006). Biclustering of gene expression data by non-smooth non-negative matrix factorization. *BMC Bioinformatics* 7:78. doi: 10.1186/1471-2105-7-78

Caspi, R., Altman, T., Billington, R., Foerster, H., Fulcher, C. A., and Keseler, I. M. (2014). The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* 44, D471–D480. doi: 10.1093/nar/gkv1164

Charlou, J. L., Donval, J. P., Fouquet, Y., Jean-Baptiste, P., and Holm, N. (2002). Geochemistry of high H2 and CH4 vent fluids issuing from ultramafic rocks at the Rainbow hydrothermal field (36 Deg14'N, MAR). *Chem. Geol.* 191, 345–359. doi: 10.1016/S0009-2541(02)00134-1

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2020). GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36, 1925–1927. doi: 10.1093/bioinformatics/btz848

Clark, I. D., and Fontes, J.-C. (1990). Paleoclimatic reconstruction in northern Oman based on carbonates from hyperalkaline groundwaters. *Quat. Res.* 33, 320–336. doi: 10.1016/0033-5894(90)90059-T

Coleman, R. G. (1981). Tectonic setting for ophiolite obduction in Oman. *J. Geophys. Res.* 86, 2497–2508. doi: 10.1029/JB086iB04p02497

Coleman, R. G., and Hopson, C. A. (1981). Introduction to the Oman Ophiolite special issue. *J. Geophys. Res.* 86, 2495–2496. doi: 10.1029/JB086iB04p02495

Coleman, D. R., Kraus, E. A., Rempfert, K., Templeton, A. S., Spear, J. R., and Boyd, E. S. (2022). Adaptations of deeply-rooted acetogens to serpentinized subsurface fluids. *Proc. Natl. Acad. Sci. U.S.A.* 119:e2206845119. doi: 10.1073/pnas.220684511

Covington, B. C., McLean, J. A., and Bachmann, B. O. (2016). Comparative mass spectrometry-based metabolomics strategies for the investigation of microbial secondary metabolites. *Nat. Prod. Rep.* 34, 6–24. doi: 10.1039/c6np00048g

Dittmar, T., Koch, B., Hertkorn, N., and Kattner, G. (2008). A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr. Methods* 6, 230–235. doi: 10.4319/lom.2008.6.230

Domingo-Almenara, X., Montenegro-Burke, J. R., Guijas, C., Majumder, E. L., Benton, H. P., and Siuzdak, G. (2019). Autonomous METLIN-guided in-source fragment annotation for untargeted metabolomics. *Anal. Chem.* 91, 3246–3253. doi: 10.1021/acs.analchem.8b03126

Douglas, A. E. (2020). The microbial exometabolome: ecological resource and architect of microbial communities. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 375:20190250. doi: 10.1098/rstb.2019.0250

Eddy, S. R. (2008). A probabilistic model of local sequence alignment that simplifies statistical significance estimation. *PLoS Comp. Biol.* 4:e1000069. doi: 10.1371/journal.pcbi.1000069

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Fiore, C. L., Longnecker, K., Kido Soule, M. C., and Kujawinski, E. (2015). Release of ecologically relevant metabolites by the cyanobacterium *Synechococcus elongatus* CCMP 1631. *Environ. Microbiol.* 17, 3949–3963. doi: 10.1111/1462-2920.12899

Fones, E. M., Colman, D. R., Kraus, E. A., Nothaft, D. B., Poudel, S., Rempfert, K. R., et al. (2019). Physiological adaptations to serpentinization in the Samail Ophiolite, Oman. *ISME J.* 13, 1750–1762. doi: 10.1038/s41396-019-0391-2

Fones, E. M., Colman, D. R., Kraus, E. A., Stepanauskas, R., Templeton, A. S., Spear, J. R., et al. (2021). Diversification of methanogens into hyperalkaline serpentinizing environments through adaptations to minimize oxidant limitation. *ISME J.* 15, 1121–1135. doi: 10.1038/s41396-020-00838-1

Fürch, T., Preusse, M., Tomasch, J., Zech, H., Wagner-Döbler, I., Rabus, R., et al. (2009). Metabolic fluxes in the central carbon metabolism of *Dinoroseobacter shibae* and *Phaeobacter gallaeciensis*, two members of the marine *Roseobacter clade*. *BMC Microbiol.* 9:209. doi: 10.1186/1471-2180-9-209

Gaujoux, R., and Seoighe, C. (2010). A flexible R package for nonnegative matrix factorization. *BMC Bioinform.* 11:367. doi: 10.1186/1471-2105-11-367

Glennie, K. W. (1996). The geology of the Oman mountains. *J. Pet. Geol.* 19:125. doi: 10.1111/j.1747-5457.1996.tb00520.x

Glennie, K. W., and Clarke, M. H. (1973). Late *Cretaceous nappes* in Oman Mountains and their geologic evolution. *AAPG Bull.* 57, 5–27. doi: 10.1306/83D912F4-16C7-11D7-8645000102C1865D

Glombitza, C., Putman, L. I., Rempfert, K. R., Kubo, M. D., Schrenk, M. O., Templeton, A. S., et al. (2021). Active microbial sulfate reduction in fluids of serpentinizing peridotites of the continental subsurface. *Commun. Earth Environ.* 2:84. doi: 10.1038/s43247-021-00157-z

Graham, E. D., Heidelberg, J. F., and Tully, B. J. (2017). BinSanity: unsupervised clustering of environmental microbial assemblies using coverage and affinity propagation. *PeerJ* 5:e3035. doi: 10.7717/peerj.3035

Graham, E. D., Heidelberg, J. F., and Tully, B. J. (2018). Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J.* 12, 1861–1866. doi: 10.1038/s41396-018-0091-3

Hacker, B. R., Mosenfelder, J. L., and Gnos, E. (1996). Rapid emplacement of the Oman ophiolite: thermal and geochronologic constraints. *Tectonics* 15, 1230–1247. doi: 10.1029/96TC01973

Haug, K., Cochrane, K., Nainala, V. C., Williams, M., Chang, J., Jayaseelan, K. V., et al. (2020). MetaboLights: a resource evolving in response to the needs of its scientific community. *Nucleic Acids Res.* 48, D440–D444. doi: 10.1093/nar/gkz1019

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28:4. doi: 10.1093/nar/28.1.27

Kato, C., and Nogi, Y. (2001). Correlation between phylogenetic structure and function: examples from deep-sea *Shewanella*. *FEMS Microbiol. Ecol.* 35, 223–230. doi: 10.1111/j.1574-6941.2001.tb00807.x

Kawamoto, J., Kurihara, T., Yamamoto, K., Nagayasu, M., Tani, Y., Mihara, H., et al. (2009). Eicosapentaenoic acid plays a beneficial role in membrane organization and cell division of a cold-adapted bacterium, *Shewanella livingstonensis* Ac10. *J. Bacteriol.* 191, 632–640. doi: 10.1128/JB.00881-08

Keleman, P. B., Leong, J. A., de Obeso, J. C., Matter, J. M., Ellison, E. T., Templeton, A., et al. (2021). Initial results from the Oman drilling project multi-borehole observatory: petrogenesis and ongoing alteration of mantle peridotite in the weathering horizon. *J. Geophys. Res. Solid Earth* 126:e2021JB022729. doi: 10.1029/2021JB022729

Kelemen, P. B., and Matter, J. (2008). *In situ* carbonation of peridotite for CO_2 storage. *Proc. Nat. Acad. Sci. U.S.A.* 105, 17295–17300. doi: 10.1073/pnas.0805794105

Kelemen, P., Rajhi, A. A., Godard, M., Ildefonse, B., Köpke, J., MacLeod, C., et al. (2013). Scientific drilling and related research in the Samail Ophiolite, Sultanate of Oman. *Sci. Dril.* 15, 64–71. doi: 10.2204/iodp.sd.15.10.2013

Kelemen, P. B., Matter, J., Streit, E. E., Rudge, J. F., Curry, W. B., and Blusztajn, J. (2011). Rates and mechanisms of mineral carbonation in peridotite: natural processes and recipes for enhanced, *in situ* CO_2 capture and storage. *Annu. Rev. Earth. Planet. Sci.* 39, 545–576. doi: 10.1146/annurev-earth-092010-152509

Kell, D., Brown, M., Davey, H., Dunn, W. B., Spasic, I., and Oliver, S. G. (2005). Metabolic footprinting and systems biology: the medium is the message. *Nat. Rev. Microbiol.* 3, 557–565. doi: 10.1038/nrmicro1177

Kim, H., and Park, H. (2007). Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. *Bioinformatics* 23, 1495–1502. doi: 10.1093/bioinformatics/btm134

Klein, F., Grozeva, N. G., and Seewald, J. S. (2019). Abiotic methane synthesis and serpentinization in olivine-hosted fluid inclusions. *Proc. Natl. Acad. Sci. U.S.A.* 116, 17666–17672. doi: 10.1073/pnas.1907871116

Kornberg, H. L., and Madsen, N. B. (1957). Synthesis of C4-dicarboxylic acids from acetate by a glyoxylate bypass of the tricarboxylic acid cycle. *Biochim. Biophys. Acta* 24, 651–653. doi: 10.1016/0006-3002(57)90268-8

Kraus, E. A., Nothaft, D., Stamps, B. W., Rempfert, K. R., Ellison, E. T., Matter, J. M., et al. (2021). Molecular evidence for an active microbial methane cycle in subsurface serpentinite-hosted groundwaters in the Samail Ophiolite, Oman. *Appl. Environ. Microbiol.* 87:e02068-20. doi: 10.1128/AEM.02068-20

Kujawinski, E. B. (2011). The impact of microbial metabolism on marine dissolved organic matter. *Annu. Rev. Mar. Sci.* 3, 567–599. doi: 10.1146/annurev-marine-120308-081003

Lan, L., Murray, T. S., Kazmierczak, B. I., and He, C. (2010). *Pseudomonas aeruginosa* OspR is an oxidative stress sensing regulator that affects pigment production, antibiotic resistance and dissemination during infection. *Mol. Microbiol.* 75, 76–91. doi: 10.1111/j.1365-2958.2009.06955.x

Lang, S. Q., Früh-Green, G. L., Bernasconi, S. M., Lilley, M. D., Proskurowski, G., Méhaya, S., et al. (2012). Microbial utilization of abiogenic carbon and hydrogen in a serpentinite-hosted system. *Geochim. Cosmochim. Acta* 92, 82–99. doi: 10.1016/j.gca.2012.06.006

Langmead, B., and Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

LaPorte, D. C., Walsh, K., and Koshland, D. E. Jr. (1984). The branch point effect. Ultrasensitivity and subsensitivity to metabolic control. *J. Biol. Chem.* 259, 14068–14075. doi: 10.1016/S0021-9258(18)89857-X

Legendre, P., and Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia* 129, 271–280. doi: 10.1007/s004420100716

Leong, J. A. M., Howells, A. E., Robinson, K. J., Cox, A., Debes, R. V. II, Fecteau, K., et al. (2021). Theoretical predictions versus environmental observations on serpentinization fluids: lessons from the Samail Ophiolite in Oman. *J. Geophys. Res. Solid Earth* 126:e2020JB020756. doi: 10.1029/2020JB020756

Li, S., Park, Y., Duraisingham, S., Strobel, F. H., Khan, N., Soltow, Q. A., et al. (2013). Predicting network activity from high throughput metabolomics. *PLoS Comput. Biol.* 9:e1003123. doi: 10.1371/journal.pcbi.1003123

Lippard, S. J., Shelton, A. W., and Gass, I. G. (1986). *The Ophiolite of Northern Oman*. London: Geological Society.

Liu, X., Lu, Y. F., Guan, X., Zhao, M., Wang, J., and Li, F. (2016). Characterizing novel metabolic pathways of melatonin receptor agonist agomelatine using metabolomic approaches. *Biochem. Pharmacol.* 109, 70–82. doi: 10.1016/j.bcp.2016.03.020

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., et al. (2020). CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 48, D265–D268. doi: 10.1093/nar/gkz991

Manandhar, M., and Cronan, J. E. (2017). Pimelic acid, the first precursor of the Bacillus subtilis biotin synthesis pathway, exists as the free acid and is assembled by fatty acid synthesis. *Mol. Microbiol.* 104, 595–607. doi: 10.1111/mmi.13648

Mandelli, F., Miranda, V. S., and Rodrigues, E. Z. (2012). Identification of carotenoids with high antioxidant capacity produced by extremophile microorganisms. *World J. Microbiol. Biotechnol.* 28, 1781–1790. doi: 10.1007/s11274-011-0993-y

McCollom, T. M. (2007). Geochemical constraints on sources of metabolic energy for chemolithoautotrophy in ultramafic-hosted deep-sea hydrothermal systems. *Astrobiology* 7, 933–950. doi: 10.1089/ast.2006.0119

McCollom, T. M. (2013). Laboratory simulations of abiotic hydrocarbon formation in Earth's deep subsurface. *Rev. Mineral. Geochem.* 75, 467–494.

McCollom, T. M., Klein, F., and Ramba, M. (2022). Hydrogen generation from serpentinization of iron-rich olivine on Mars, icy moons, and other planetary bodies. *Icarus* 372:114754. doi: 10.1016/j.icarus.2021.114754

McCollom, T. M., and Seewald, J. S. (2001). A reassessment of the potential for reduction of dissolved CO_2 to hydrocarbons during serpentinization of olivine. *Geochim. Cosmochim. Acta* 65, 3769–3778. doi: 10.1016/S0016-7037(01)00655-X

McCollom, T. M., and Seewald, J. S. (2006). Carbon isotope composition of organic compounds produced by abiotic synthesis under hydrothermal conditions. *Earth Planet. Sci. Lett.* 243, 74–84. doi: 10.1016/j.epsl.2006.01.027

Miller, H. M., Matter, J. M., Kelemen, P., Ellison, E. T., Conrad, M. E., Fierer, N., et al. (2016). Modern water/rock reactions in Oman hyperalkaline peridotite aquifers and implications for microbial habitability. *Geochim. Cosmochim. Acta* 179, 217–241. doi: 10.1016/j.gca.2016.01.033

Miller, D., O'Brien, K., Xu, H., and White, R. H. (2014). Identification of a 5'-deoxyadenosine deaminase in Methanocaldococcus janaschii and its possible role in recycling the radical S-adenosylmethionine enzyme reaction product 5'-deoxyadenosine. *J. Bacteriol.* 196, 1064–1072. doi: 10.1128/JB.01308-13

Neal, C., and Stanger, G. (1983). Hydrogen generation from mantle source rocks in Oman. *Earth Planet. Sci. Lett.* 66, 315–320. doi: 10.1016/0012-821X(83)90144-9

Neal, C., and Stanger, G. (1985). "Past and present serpentinisation of ultramafic rocks; an example from the Semail Ophiolite nappe of northern Oman," in *The Chemistry of Weathering*, ed. J. I. Drever (Dordrecht: Springer Netherlands), 249–275.

Neely, C. J., Graham, E. D., and Tully, B. J. (2020). MetaSanity: an integrated microbial genome evaluation and annotation pipeline. *Bioinformatics* 36, 4341–4344. doi: 10.1093/bioinformatics/btaa512

Newman, S. A., Lincoln, S. A., O'Reilly, S., Liu, X., Shock, E. L., Kelemen, P. B., et al. (2020). Lipid biomarker record of the serpentinite-hosted ecosystem of the Samail Ophiolite, Oman and implications for the search for biosignatures on Mars. *Astrobiology* 20, 830–845. doi: 10.1089/ast.2019.2066

Nicolas, A., Boudier, F., Ildefonse, B., and Ball, E. (2000). Accretion of Oman and United Arab Emirates ophiolite- discussion of a new structural map. *Mar. Geophys. Res.* 21, 147–180. doi: 10.1023/A:1026769727917

Nishida, T., Hori, R., Morita, N., and Okuyama, H. (2010). Membrane eicosapentaenoic acid is involved in the hydrophobicity of bacterial cells and affects the entry of hydrophilic and hydrophobic compounds. *FEMS Microbiol. Lett.* 306, 91–96. doi: 10.1111/j.1574-6968.2010.01943.x

Nishida, T., Morita, N., Yano, Y., Orikasa, Y., and Okuyama, H. (2007). The antioxidative function of eicosapentaenoic acid in a marine bacterium, *Shewanella marinintestina* IK-1. *FEBS Lett.* 581, 4212–4216. doi: 10.1016/j.febslet.2007.07.065

Nishida, T., Orikasa, Y., Ito, Y., Yu, R., Yamada, A., Watanabe, K., et al. (2006a). *Escherichia coli* engineered to produce eicosapentaenoic acid becomes resistant against oxidative damages. *FEBS Lett.* 580, 2731–2735. doi: 10.1016/j.febslet.2006.04.032

Nishida, T., Orikasa, Y., Watanabe, K., and Okuyama, H. (2006b). The cell membrane-shielding function of eicosapentaenoic acid for *Escherichia coli* against exogenously added hydrogen peroxide. *FEBS Lett.* 580, 6690–6694. doi: 10.1016/j.febslet.2006.11.030

Nothaft, D. B., Templeton, A. S., Boyd, E. S., Matter, J. M., Stute, M., Paukert Vankeuren, A. N., et al. (2021). Aqueous geochemical and microbial variation across discrete depth intervals in a peridotite aquifer assessed using a packer system in the Samail Ophiolite, Oman. *J. Geophys. Res. Biogeosci.* 126:e2021JG006319. doi: 10.1029/2021JG006319

Okuyama, H., Orikasa, Y., and Nishida, T. (2008). Significance of antioxidative functions of eicosapentaenoic and docosahexaenoic acids in marine microorganisms. *Appl. Environ. Microbiol.* 74, 570–574. doi: 10.1128/AEM.02256-07

Okuyama, H., Orikasa, Y., Nishida, T., Watanabe, K., and Morita, N. (2007). Bacterial genes responsible for the biosynthesis of eicosapentaenoic acid and docosahexanoic acids and their heterologous expression. *Appl. Environ. Microbiol.* 73, 665–670. doi: 10.1128/AEM.02270-06

Oyedeji, A. B., Green, E., Adebiyi, J. A., Ogundele, O. M., Gbashi, S., Adefisoye, M. J., et al. (2021). Metabolomic approaches for the determination of metabolites from pathogenic microorganisms: a review. *Food Res. Int.* 140:110042. doi: 10.1016/j.foodres.2020.110042

Parkes, R. J., Cragg, B. A., Bale, S. J., Getliff, J. M., Goodman, K., Rochelle, P. A., et al. (1994). Deep bacterial biosphere in Pacific Ocean sediments. *Nature* 371, 410–413. doi: 10.1038/371410a0

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Patti, G. J., Yanes, O., and Siuzdak, G. (2012). Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell. Biol.* 13, 263–269. doi: 10.1038/nrm3314

Paukert, A. N., Matter, J. M., Kelemen, P. B., Shock, E. L., and Havig, J. R. (2012). Reaction path modeling of enhanced *in situ* CO_2 mineralization for carbon sequestration in the peridotite of the Samail Ophiolite, Sultanate of Oman. *Chem. Geol.* 330-331, 86–100. doi: 10.1016/j.chemgeo.2012.08.013

Paukert Vankeuren, A. N., Matter, J. M., Stute, M., and Kelemen, P. B. (2019). Multitracer determination of apparent groundwater ages in peridotite aquifers within the Samail ophiolite, Sultanate of Oman. *Earth Planet. Sci. Lett.* 516, 37–48. doi: 10.1016/j.epsl.2019.03.007

Peregrín-Alvarez, J. M., Sanford, C., and Parkinson, J. (2009). The conservation and evolutionary modularity of metabolism. *Genome Biol.* 10:R63. doi: 10.1186/gb-2009-10-6-r63

Peyraud, R., Kiefer, P., Christen, P., and Vorholt, J. A. (2009). Demonstration of the ethylmalonyl-CoA pathway by using ^{13}C metabolomics. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4846–4851. doi: 10.1073/pnas.0810932106

Porcar, M., Louie, K. B., Kosina, S. M., Van Goethem, M. W., Bowen, B. P., Tanner, K., et al. (2018). Microbial ecology on solar panels in Berkeley, CA, United States. *Front. Microbiol.* 9:3043. doi: 10.3389/fmicb.2018.03043

Proskurowski, G., Lilley, M. D., Seewald, J. S., Früh-Green, G., Olson, E. J., Lupton, J. E., et al. (2008). Abiotic hydrocarbon production at lost city hydrothermal field. *Science* 319, 604–607. doi: 10.1126/science.1151194

Rabinowitz, J. D., and Kimball, E. (2007). Acidic acetonitrile for cellular metabolome extraction from *Escherichia coli*. *Anal. Chem.* 79, 6167–6173. doi: 10.1021/ac070470c

Rempfert, K. R., Miller, H. M., Bompard, N., Nothaft, D., Matter, J. M., Kelemen, P., et al. (2017). Geological and geochemical controls on subsurface microbial life in the Samail Ophiolite, Oman. *Front. Microbiol.* 8:56. doi: 10.3389/fmicb.2017.00056

Sabuda, M. C., Brazelton, W. J., Putman, L. I., McCollom, T. M., Hoehler, T. M., Kubo, M. D. Y., et al. (2020). A dynamic microbial sulfur cycle in a serpentinizing continental ophiolite. *Environ. Microbiol.* 22, 2329–2345. doi: 10.1111/1462-2920.15006

Schrenk, M. O., Brazelton, W. J., and Lang, S. Q. (2013). Serpentinization, carbon, and deep life. *Rev. Mineral. Geochem.* 75, 575–606. doi: 10.2138/rmg.2013.75.18

Schrenk, M. O., Kelley, D. S., Bolton, S. A., and Baroos, J. A. (2004). Low archaeal diversity linked to subseafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ. Microbiol.* 6, 1086–1095. doi: 10.1111/j.1462-2920.2004.00650.x

Schulte, M., Blake, D., Hoehler, T., and McCollom, T. (2006). Serpentinization and its implications for life on the early Earth and Mars. *Astrobiology* 6, 364–376. doi: 10.1089/ast.2006.6.364

Schwieterman, E. W., Cockell, C. S., and Meadows, V. S. (2015). Nonphotosynthetic pigments as potential biosignatures. *Astrobiology* 15, 341–361. doi: 10.1089/ast.2014.1178

Seyler, L., Kujawinski, E. B., Azua-Bustos, A., Lee, M. D., Marlow, J., Perl, S. M., et al. (2020). Metabolomics as an emerging tool in the search for astrobiologically relevant biomarkers. *Astrobiology* 20, 1251–1261. doi: 10.1089/ast.2019.2135

Seyler, L. M., Brazelton, W. L., McLean, C., Putman, L. I., Hyer, A., Kubo, M. D. Y., et al. (2020). Carbon assimilation strategies in ultrabasic groundwater: clues from the integrated study of a serpentinization-influenced aquifer. *mSystems* 5:e00607-19. doi: 10.1128/mSystems.00607-19

Shock, E. L., and Schulte, M. D. (1998). Organic synthesis during fluid mixing in hydrothermal systems. *J. Geophys. Res.* 103, 28513–28527. doi: 10.1029/98JE02142

Shulse, C. N., and Allen, E. E. (2011). Widespread occurrence of secondary lipid biosynthesis potential in microbial lineages. *PLoS One* 6:e20146. doi: 10.1371/journal.pone.0020146

Sleep, N. H., Meibom, A., Fridriksson, T., and Bord, D. K. (2004). H_2-rich fluids from serpentinization: geochemical and biotic implications. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12818–12823. doi: 10.1073/pnas.0405289101

Smith, C. A., O'Maille, G. O., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., et al. (2005). METLIN: a metabolite mass spectral database. *Ther. Drug Monit.* 27, 747–751. doi: 10.1097/01.ftd.0000179845.53213.39

Sogin, E. M., Puskás, E., Dubilier, N., and Liebeke, M. (2019). Marine metabolomics: a method for nontargeted measurement of metabolites in seawater by gas chromatography-mass spectrometry. *mSystems* 4:e00638-19. doi: 10.1128/mSystems.00638-19

Soldatou, S., Eldjarn, G. H., Huerta-Uribe, A., Rogers, S., and Duncan, K. R. (2019). Linking biosynthetic and chemical space to accelerate microbial secondary metabolite discovery. *FEMS Microbiol. Lett.* 366:fnz142. doi: 10.1093/femsle/fnz142

Streit, E., Kelemen, P., and Eiler, J. (2012). Coexisting serpentine and quartz from carbonate-bearing serpentinized peridotite in the Samail Ophiolite, Oman. *Contrib. Mineral. Petrol.* 164, 821–837. doi: 10.1007/s00410-012-0775-z

Suzuki, S., Nealson, N. H., and Ishii, S. (2018). Genomic and in-situ transcriptomic characterization of the candidate phylum NPL-UPL2 from highly alkaline highly reducing serpentinized groundwater. *Front. Microbiol.* 9:3141. doi: 10.3389/fmicb.2018.03141

Tang, Y. J., Yi, S., Zhuang, W.-Q., Zinder, S. H., Keasling, J. D., and Alvarez-Cohen, L. (2009). Investigation of carbon metabolism in "*Dehalococcoides ethenogenes*" strain 195 by use of isotopomer and transcriptomic analyses. *J. Bacteriol.* 191, 5224–5231. doi: 10.1128/JB.00085-09

Templeton, A. S., Ellison, E. T., Glombitza, C., Morono, Y., Rempfert, K. R., Hoehler, T. M., et al. (2021). Accessing the subsurface biosphere within rocks undergoing active

low-temperature serpentinization in the Samail Ophiolite (Oman Drilling Project). *J. Geophys. Res. Biogeosci.* 126:e2021JG006315.

Ternieten, L., Früh-Green, G. L., and Bernasconi, S. M. (2021). Carbon geochemistry of the active serpentinization site at the Wadi Tayin Massif: insights from the ICDP Oman Drilling Project: phase II. *J. Geophys. Res.Solid Earth* 126:e2021JB022712. doi: 10.1029/2021JB022712

Tian, B., and Hua, Y. (2010). Carotenoid biosynthesis in extremophilic Deinococcus-Thermus bacteria. *Trends Microbiol.* 18, 512–520. doi: 10.1016/j.tim.2010.07.007

Tilay, A., and Annapure, U. (2012). Novel simplified and rapid method for screening and isolation of polyunsaturated fatty acids producing marine bacteria. *Biotechnol. Res. Int.* 2012:542721. doi: 10.1155/2012/542721

Tilton, G. R., Hopson, C. A., and Wright, J. E. (1981). Uranium-lead isotopic ages of the Samail ophiolite, Oman, with applications to Tethyan Ocean ridge tectonics. *J. Geophys. Res.* 86, 2763–2775. doi: 10.1029/JB086iB04p02763

Turnbaugh, P. J., and Gordon, J. I. (2008). An invitation to the marriage of metagenomics and metabolomics. *Cell* 134, 708–713. doi: 10.1016/j.cell.2008.08.025

UniProt Consortium (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D489. doi: 10.1093/nar/gkaa1100

Usui, K., Hiraki, T., Kawamoto, J., Kurihara, T., Nogi, Y., Kato, C., et al. (2012). Eicosapentaenoic acid plays a role in stabilizing dynamic membrane structure in the deep-sea piezophile *Shewanella violacea*: a study employing high-pressure time-resolved fluorescence anisotropy measurement. *Biochim. Biophys. Acta* 1818, 574–583. doi: 10.1016/j.bbamem.2011.10.010

van Santen, J. A., Kautsar, S. A., Medema, M. H., and Linington, R. G. (2021). Microbial natural product databases: moving forward in the multi-omics era. *Nat. Prod. Rep.* 38, 264–278. doi: 10.1039/d0np00053a

Vance, S., Harnmeijer, J., Kimura, J., Hussman, H., Demartin, B., and Brown, J. M. (2007). Hydrothermal systems in small ocean planets. *Astrobiology* 7, 987–1005. doi: 10.1089/ast.2007.0075

Wang, F., Wang, P., Chen, M., and Xiao, X. (2004). Isolation of extremophiles with the detection and retrieval of *Shewanella* strains in deep-sea sediments from the west Pacific. *Extremophiles* 8, 165–168. doi: 10.1007/s00792-003-0365-0

Yoshida, K., Hashimoto, M., Hori, R., Adachi, T., Okuyama, H., Orikasa, Y., et al. (2016). Bacterial long-chain polyunsaturated fatty acids: their biosynthetic genes, functions, and practical use. *Mar. Drugs* 14, 94. doi: 10.3390/md14050094