Carbon Assimilation Strategies in Ultrabasic Groundwater: Clues from the Integrated Study of a Serpentinization-Influenced Aquifer

Lauren M. Seyler,[a,b] William J. Brazelton,[c] Craig McLean,[a,d] Lindsay I. Putman,[e] Alex Hyer,[c] Michael D. Y. Kubo,[b,f] Tori Hoehler,[f] Dawn Cardace,[g] Matthew O. Schrenk[e,h]

aMarine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA
bBlue Marble Space Institute of Science, Seattle, Washington, USA
cDepartment of Biology, University of Utah, Salt Lake City, Utah, USA
dDepartment of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
eDepartment of Earth and Environmental Sciences, Michigan State University, East Lansing, Michigan, USA
fSpace Sciences Division, NASA Ames Research Center, Mountain View, California, USA
gDepartment of Geosciences, University of Rhode Island, Kingston, Rhode Island, USA
hDepartment of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan, USA

ABSTRACT Serpentinization is a low-temperature metamorphic process by which ultramafic rock chemically reacts with water. Such reactions provide energy and materials that may be harnessed by chemosynthetic microbial communities at hydrothermal springs and in the subsurface. However, the biogeochemistry mediated by microbial populations that inhabit these environments is understudied and complicated by overlapping biotic and abiotic processes. We applied metagenomics, metatranscriptomics, and untargeted metabolomics techniques to environmental samples taken from the Coast Range Ophiolite Microbial Observatory (CROMO), a subsurface observatory consisting of 12 wells drilled into the ultramafic and serpentinite mélange of the Coast Range Ophiolite in California. Using a combination of DNA and RNA sequence data and mass spectrometry data, we found evidence for several carbon fixation and assimilation strategies, including the Calvin-Benson-Bassham cycle, the reverse tricarboxylic acid cycle, the reductive acetyl coenzyme A (acetyl-CoA) pathway, and methylotrophy, in the microbial communities inhabiting the serpentinite-hosted aquifer. Our data also suggest that the microbial inhabitants of CROMO use products of the serpentinization process, including methane and formate, as carbon sources in a hyperalkaline environment where dissolved inorganic carbon is unavailable.

IMPORTANCE This study describes the potential metabolic pathways by which microbial communities in a serpentinite-influenced aquifer may produce biomass from the products of serpentinization. Serpentinization is a widespread geochemical process, taking place over large regions of the seafloor and at continental margins, where ancient seafloor has accreted onto the continents. Because of the difficulty in delineating abiotic and biotic processes in these environments, major questions remain related to microbial contributions to the carbon cycle and physiological adaptation to serpentinite habitats. This research explores multiple mechanisms of carbon fixation and assimilation in serpentinite-hosted microbial communities.

KEYWORDS carbon assimilation, carbon fixation, formaldehyde, formate, methane, serpentinization

Serpentinization is the process by which ultramafic rock in the lower crust and upper mantle of the Earth is hydrated, leading to the oxidation of ferrous iron in the minerals olivine and pyroxene and the release of hydrogen (H_2) and hydroxyl ions (OH^-). The reducing power supplied by serpentinization, when mixed with oxidants from surface and subsurface fluids, creates chemical disequilibria that microbial populations can harness [1–3]. However, due to high pH and the release of cations during rock weathering, much of the dissolved inorganic carbon (DIC) in serpentinite systems is precipitated as calcite and aragonite minerals [4]. Consequently, small carbon-bearing compounds, including methane (CH_4), carbon monoxide (CO), and formate, may take on important roles in sustaining microbial ecosystems in serpentinites [5]. The high pH and limited availability of both DIC and electron acceptors represent fundamental challenges for microbial communities hosted in serpentinization-influenced environments [6].

Overlapping biogenic and abiogenic processes in serpentinizing rocks make it difficult to delineate carbon cycling pathways used in these environments (Fig. 1). Under highly reducing conditions, in the presence of specific mineral catalysts, hydrogen reacts with carbon dioxide (CO_2) or carbon monoxide to form methane and small-chain hydrocarbons via Fischer-Tropsch-type reactions [7–10]. These compounds are also formed by biological activity, or diagenesis of organic matter [11]. For example, high concentrations of dissolved organic material have been detected in close association with serpentine-hosted hydrogarnets recovered from the Mid-Atlantic Ridge. This material may have been produced by past microbiological activity [12] or by purely abiotic mechanisms [13]. Acetate in endmember fluids at the Lost City Hydrothermal Field is likely produced by biological activity [14], but other hydrocarbons in the rock-hosted fluids do not have a clear biotic or abiotic source [15], though formate and C_2^+ alkanes appear to be produced abiotically [9, 14]. Biotic versus abiotic sources of methane in serpentinizing systems are likewise difficult to resolve and vary from site to site [16–19].

In situ microbial communities may utilize the products of serpentinization—including methane, carbon monoxide, formate, and acetate—as sources of carbon [14, 20–25]. Microbial communities in serpentinite springs in the Voltri Massif, Italy, appear to perform both aerobic methanotrophy and methanogenesis using CO_2 liberated from acetate [26]. Methanogens in The Cedars, CA, are capable of producing methane from both inorganic carbon and a variety of organic carbon substrates. Communities that reside deeper in the springs at The Cedars also possess a reductive acetyl coenzyme A (acetyl-CoA) pathway for autotrophic acetogenesis [27]. Fermentation of organic matter may also lead to acetate production [19]. Carbon monoxide may be used as a substrate by microbial communities in the Tablelands, Canada, but it appears to be primarily used as a source of energy rather than carbon [28]. The carbonate chimneys of the Lost City Hydrothermal Field are dominated by a single species of archaea that may be capable of both production and consumption of methane in tandem [21]. Genes encoding NAD(+)-dependent formate dehydrogenase in metagenomes recovered from the Samail Ophiolite, Oman, aquifer suggest that formate is an important source of carbon [29]. In short, sources and sinks of carbon in these environments remain enigmatic to researchers, but pathways and strategies utilizing alternate sources of carbon supplied by serpentinization appear to be prevalent.

In this study, we examined the microbial metabolic potential and biogeochemistry of a serpentinization-influenced aquifer in the Coast Range Ophiolite Microbial Observatory (CROMO), CA. Wells have been drilled at various depths and locations at CROMO such that microbial processes can be explored across a wide range of physical and chemical conditions. Carbon assimilation pathways in the microbial communities at CROMO were investigated using a combination of metagenomic, metatranscriptomic, and metabolomic approaches. Metagenomic data sets from 9 of the 12 CROMO wells, and metatranscriptomic data sets from 4 wells, obtained over multiple sampling trips, were queried for complete pathways and individual genes associated with carbon fixation and methane cycling. We then applied environmental metabolomics to directly
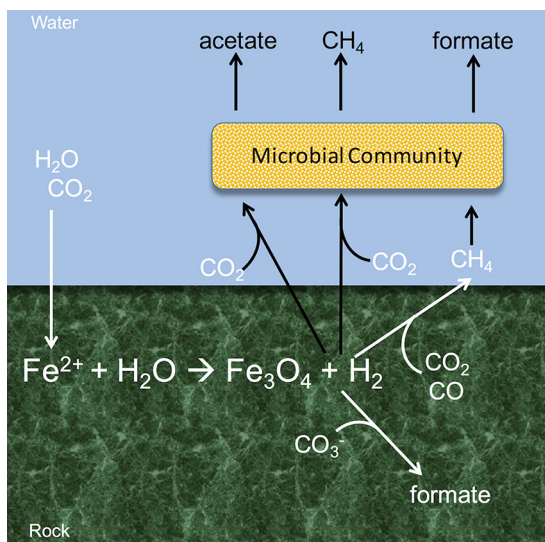
FIG 1 Diagram depicting abiotic (white) and biotic (black) sources of carbon compounds in serpentinizing environments. Adapted from the work of Preiner et al. (83).

assess the products of ecosystem activity in intra- and extracellular extracts from the two most alkaline wells (CSW1.1 and QV1.1). This study linked the genetic potential of serpentinite-hosted communities to the observed biogeochemistry of the groundwater in CROMO.

RESULTS

Presence and expression of carbon fixation and assimilation pathways. The metagenomic data collected from 9 of the 12 CROMO wells contain evidence for a range of carbon fixation and assimilation pathways (Fig. 2). Functional profiling of the metagenomes using HUMAnN2 identified multiple C_1 compound utilization and assimilation pathways, including the reductive acetyl-CoA (Wood-Ljungdahl) pathway, the reverse tricarboxylic acid (rTCA) cycle, formaldehyde oxidation to formate, and formaldehyde fixation to biomass. Whenever these complete pathways were identified in the metagenomic data, they were also detected in the corresponding metatranscriptomic data (Fig. 2). A complete Calvin-Benson-Bassham (CBB) cycle was also detected in all nine metagenomes and all four metatranscriptomes, at similar coverages across wells (see Fig. S1 in the supplemental material). Carbon fixation pathways found exclusively in archaea (i.e., hydroxypropionate-hydroxybutylate cycle and 3-hydroxypropionate bicycle) were not detected. Metagenome-assembled genome (MAG) bins contained the reductive pentose phosphate/Calvin-Benson-Bassham cycle, reductive TCA cycle, and reductive acetyl-CoA pathways (Fig. 3).

Several pathways associated with methanotrophy and methylotrophy were detected in the metagenomes and metatranscriptomes. The complete pathway for aerobic methane oxidation to formaldehyde was present in most of the metagenomes (Fig. 4) and transcribed in wells QV1.2 and N08B, two of the shallower, more oxic wells in the observatory. Formaldehyde uptake and assimilation pathways were detected in all of the sequenced wells (Fig. 2). Genes for the complete glutathione-dependent and tetrahydrofolate (THF) pathways for formaldehyde oxidation to formate were present in all of the metagenomes, expressed in all four metatranscriptomes (Fig. 2 and 4), and present in multiple MAGs (Fig. 3). The *gfa* gene, which codes for an enzyme that catalyzes a thermodynamically spontaneous step in the glutathione-dependent pathway (30), is found in only some formaldehyde-oxidizing bacteria and was absent from most of the metagenomes. Genes for formaldehyde fixation via the tetrahydromethanopterin (H4MPT) pathway had low coverage across the metagenomes but were expressed in the metatranscriptomes of QV1.2 and N08B. The capability for formalde-
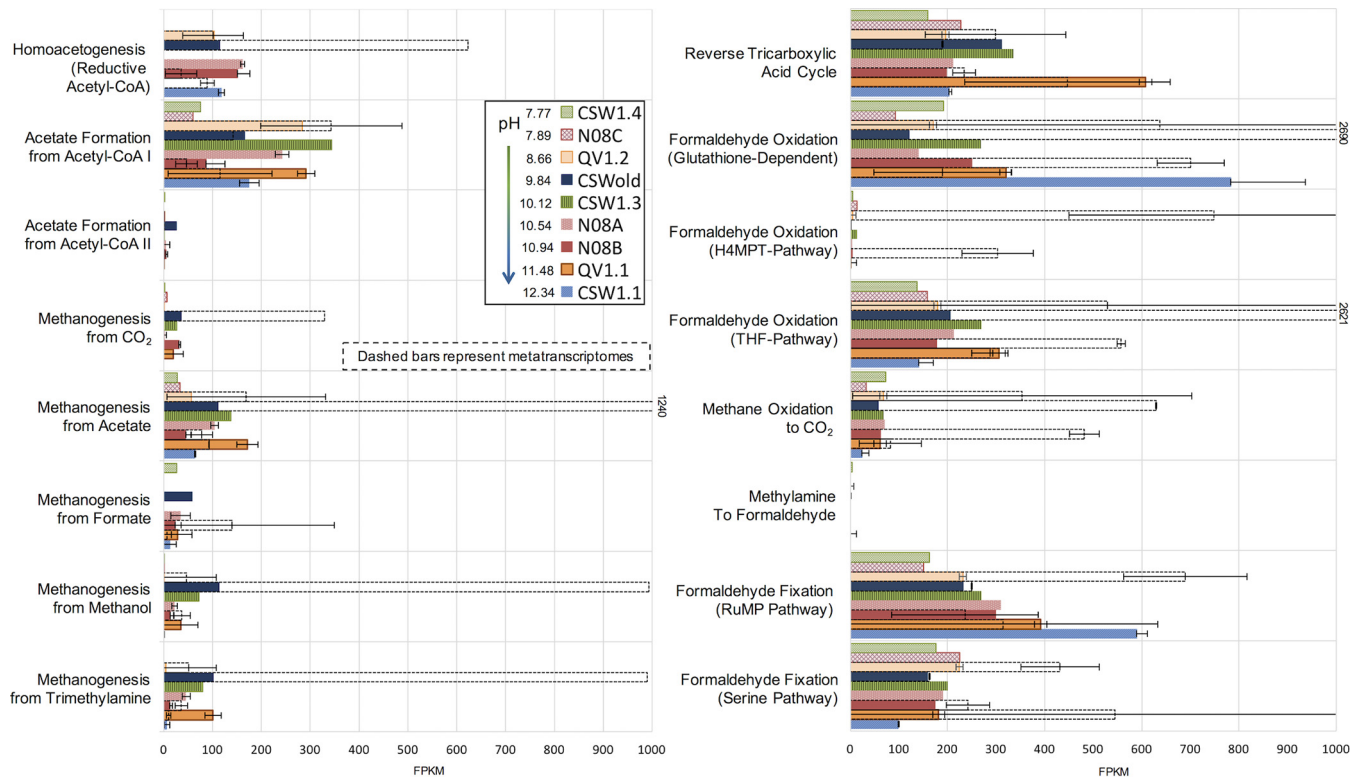
FIG 2 Metagenome coverage of Functional Ontology Assignments for Metagenomes (FOAM) pathways assigned using HUMAnN2. If a well was sampled for metagenomics more than once, the average percent coverage and standard error of the mean for that well are provided. Metatranscriptome coverage for four wells sampled for RNA is indicated using bars with dotted outlines.

hyde fixation to biomass by the ribulose monophosphate (RuMP) and serine pathways was likewise suggested in all nine metagenomes and all four metatranscriptomes (Fig. 2), and the RuMP pathway was partially complete in 18 MAGs and complete in 1 MAG (Fig. 3).

Pathway identification using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and Functional Ontology Assignments for Metagenomes (FOAM) pathways suggested that homoacetogenesis via the reductive acetyl-CoA pathway (also known as the Wood-Ljungdahl pathway) was represented in the metagenomes of only half of the wells sequenced (N08A, N08B, QV1.2, CSW1.1, and CSWold) (Fig. 2). However, a gene-by-gene search indicated that the pathway was also present, and nearly complete, in the QV1.1 and CSW1.3 wells (Fig. 5 and Table S1; see also Table S4 at https://figshare.com/articles/Supplemental_Table_4_xlsx/11873754). In the three least alkaline wells (CSW1.4, N08C, and QV1.2), most of the genes in the pathway displayed low abundance (<2 fragments per kilobase per million [FPKM]) in the metagenomic data. Metatranscriptomic data showed that the pathway was actively transcribed in all four wells that were sampled for mRNA (Fig. 2 and 5). However, one of the initial steps of the reductive acetyl-CoA pathway, the conversion of CO_2 to formate by NADP-dependent formate dehydrogenase (FDH), was rarely expressed or altogether missing in both the metagenomes and the metatranscriptomes (Fig. 5). Seven of the 66 MAG bins also contained a nearly complete reductive acetyl-CoA pathway that was missing the genes encoding FDH (Fig. 3, bottom). The type I pathway for the formation of acetate from acetyl-CoA via acetylphosphate (present in *Methanosarcina* and *Clostridiales*) was relatively abundant across all metagenomes, while the type II pathway (which does not utilize an acetylphosphate intermediate) was sparsely represented in the metagenomic data (Fig. 2).

Taxonomy. Taxonomic assignments of the metagenomic contigs using Phylo-PythiaS+ were consistent with previous 16S rRNA sequencing data from the CROMO
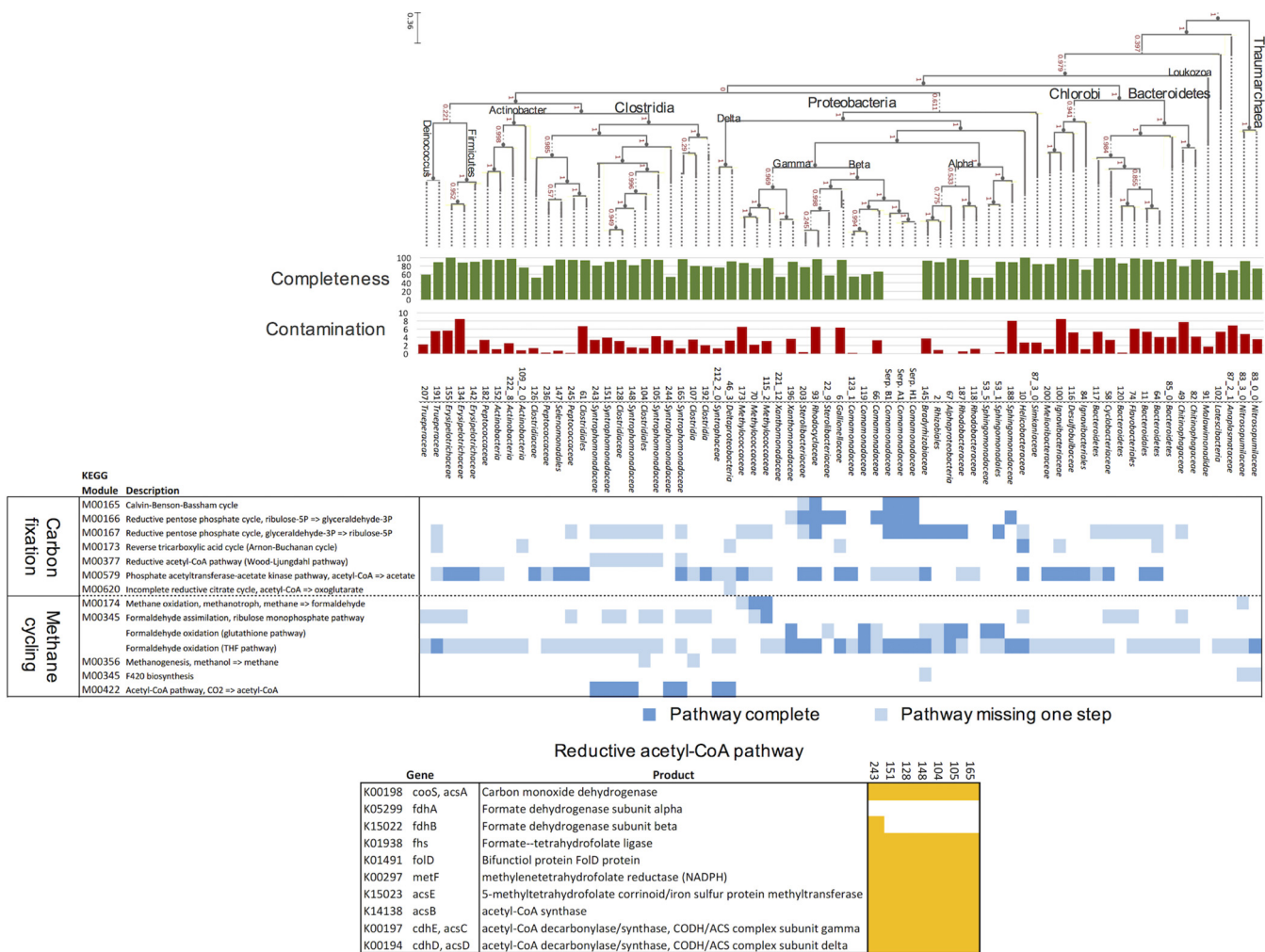
FIG 3 Selection of carbon cycling-associated KEGG modules detected in MAG bins, and *Serpentinomonas* isolate genomes (69), using BLASTKoala. Dark blue boxes indicate that the complete module is present; light blue boxes indicate that one step of the pathway is missing. The phylogenetic tree comparing the MAGs to the three *Serpentimonas* isolates was generated using SpeciesTree in KBase. Completeness and contamination scores for each MAG are also provided. (Bottom) Bins containing the reductive acetyl-CoA pathway; gene presence in the genome is indicated in yellow.

wells (6) and showed that the microbial communities in QV1.1 and CSW1.1 (the two most alkaline wells) were highly similar and dominated by keystone taxa of serpentinite-hosted microbial communities, including *Comamonadaceae* and *Clostridiales* (Fig. S3). Taxonomic classification of MAG bins, including bin completeness, contamination, and strain heterogeneity, is available in Table S2; a phylogenetic tree comparing the MAGs to the three *Serpentinomonas* isolates is included in Fig. 3. Three MAGs—66, 119, and 123_1—were closely related to the three *Serpentinomonas* isolates and were classified as *Comamonadaceae*, *Hydrogenophaga* sp. (Fig. 3; Table S2).

In the metagenomes, particulate methane monooxygenase (*pmoA*) and methanol dehydrogenase (*mxaF*) genes from the methane oxidation pathway were detected on contigs classified as *Methylophilales* (*Methylophilaceae*) and *Methylococcales* (*Methylococcaceae*); two MAG bins containing a complete methane oxidation pathway were classified as *Methylococcaceae* (Fig. 3; Table S2). Genes for methylenetetrahydrofolate dehydrogenase (*folD*) and formyltetrahydrofolate deformylase (*purU*) from the THF pathway for formaldehyde oxidation were associated with *Burkholderiales* (*Comamonadaceae*), *Deinococcales* (*Trueperaceae*), *Hydrogenophilales* (*Hydrogenophilaceae*), *Methylophilales* (*Methylophilaceae*), *Rhizobiales*, and *Rhodocyclales* (*Rhodocyclaceae*). The *S*-(hydroxymethyl)glutathione dehydrogenase (*frmA*) gene from the glutathione pathway for formaldehyde oxidation was detected on contigs identified as *Burkholderiales*
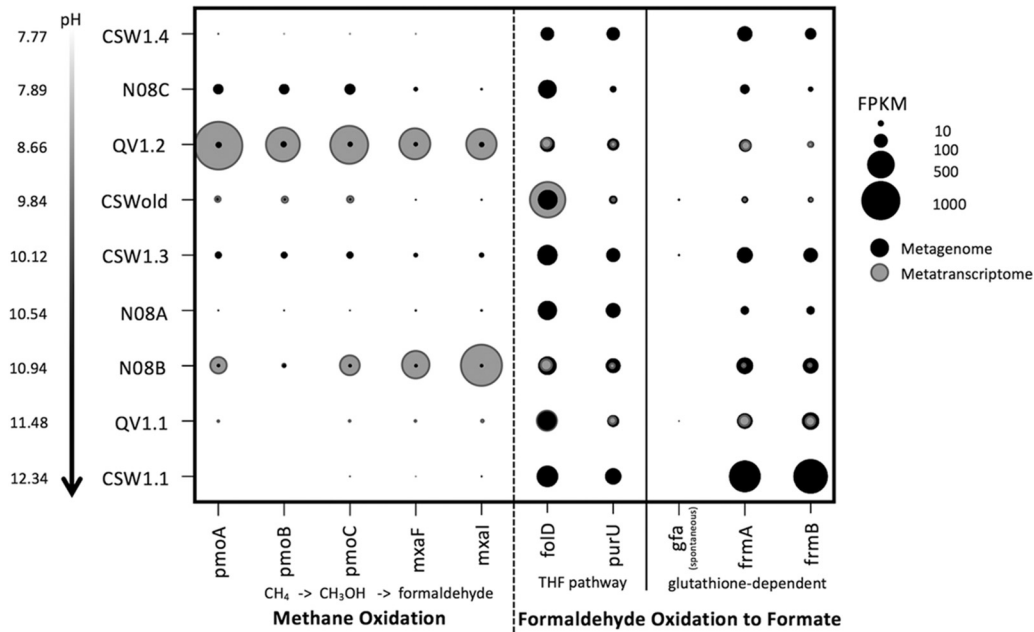
FIG 4 Fragments per kilobase of sequences per million mapped reads (FPKM) of genes in the homoacetogenic reductive acetyl-CoA pathway. Metagenomes are represented in black; metatranscriptomes are represented in gray. The genes for the two subunits of formate dehydrogenase (*fdhA* and *fdhB*) are missing or comparatively low in most wells. This enzyme is responsible for converting CO_2 to formate.

(*Comamonadaceae*), *Rhizobiales*, *Rhodobacterales*, and *Rhodocyclales* (*Rhodocyclaceae*). The *folD* gene is also present in the reductive acetyl-CoA pathway for acetogenesis; this gene, as well as the genes for acetyl-CoA synthase (*acsB*) and the catalytic subunit of anaerobic carbon-monoxide dehydrogenase (*cooS*), was found in *Clostridiales* (*Syntrophomodaceae*) and *Desulfuromodales*. All of the MAG bins containing the reductive acetyl-CoA pathway were classified as *Clostridiales* (Fig. 3; Table S2). The large and small subunits of RuBisCO (*rbcL* and *rbcS*) were detected on contigs classified as *Burkhold-*
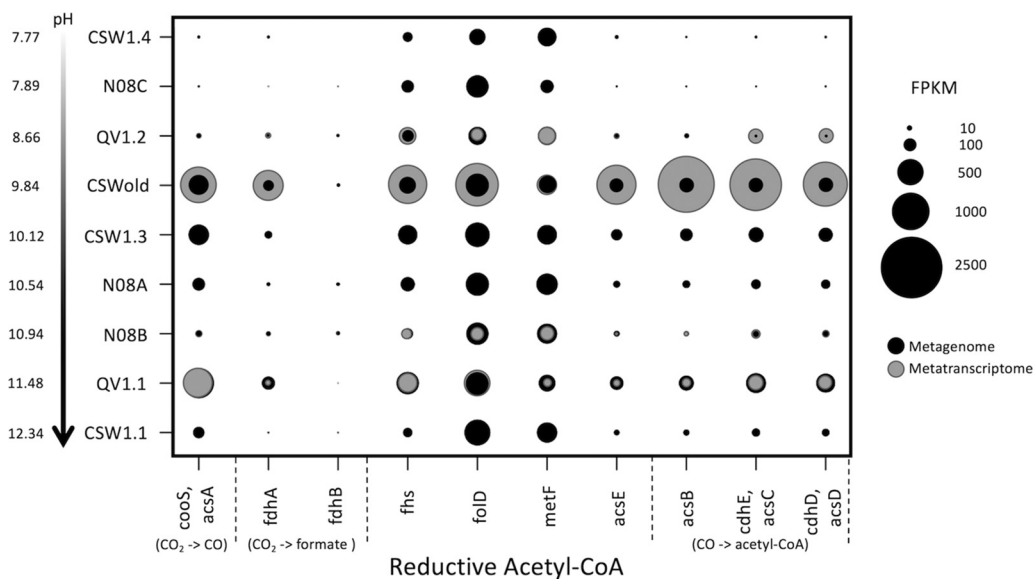
FIG 5 Fragments per kilobase of sequences per million mapped reads of genes in the methane oxidation, tetrahydrofolate (THF), and glutathione-dependent formaldehyde oxidation pathways. Metagenomes are represented in black; metatranscriptomes are represented in gray. The gene for *S*-(hydroxymethyl)glutathione synthase (*gfa*) is present at low abundance or missing in all wells; this enzyme catalyzes a spontaneous reaction in the formaldehyde oxidation pathway.
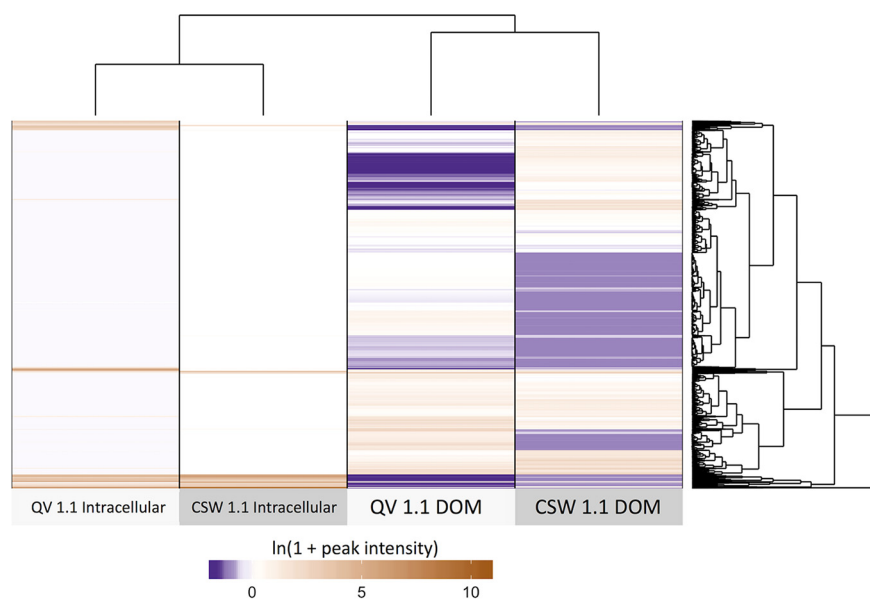
FIG 6 Heat map of untargeted metabolomics features across all four metabolome samples. The scale represents the ln(1 + peak intensity) [1 is added to avoid cases of ln(0)].

eriales (*Comamonadaceae*), *Hydrogenophilales* (*Hydrogenophilaceae*), *Rhodocyclales* (*Rhodocyclaceae*), and *Methylophilales* (*Methylophilaceae*). A complete CBB cycle was only identified in one MAG, however, classified as *Rhodocyclales* (*Rhodocyclaceae*). The ATP-citrate lyase alpha-subunit (*aclA*) gene from the reverse TCA cycle was found in *Clostridiales* (*Syntrophomodaceae*) in the metagenomes.

Metabolomics. Metabolite extracts were obtained from both cellular biomass on filters (intracellular metabolites) and dissolved organic carbon (DOC) in the filtrate (extracellular metabolites) from the two most alkaline wells, QV1.1 and CSW1.1. The metabolomes of QV1.1 and CSW1.1 displayed many unique features across both sample types. Metabolomes had more features in common between samples of the same type than samples from the same well. In particular, the DOC pools of QV1.1 and CSW1.1 were highly distinct from one another and from intracellular extracts, and these metabolomes contained the greatest number of features unique to the sample (Fig. 6; Fig. S4). Dissolved inorganic carbon concentration in all wells was inversely proportional to pH (Pearson correlation coefficient [r] = −0.798), while nonpurgeable organic carbon was positively correlated to dissolved oxygen (DO) concentration ($r = 0.587$) (Fig. 7). Most of the wells in the CROMO cluster are anoxic according to the U.S. Geological Survey definition (<0.5 mg/liter), which is consistent with the hydrogen generated through the serpentinization reaction (Table S3). The correlation between bulk dissolved organic matter (DOM) and dissolved oxygen concentration could be a consequence of surficial inputs to the shallower wells, as reported in an earlier study (25). However, recent hydrogeological studies at CROMO showed that the seasonal surface-connected aquifer extends only about 7 m into the subsurface before hitting an aquitard, beneath which is a confined serpentinite aquifer (31). Additionally, subsurface tritium isotope values are below detection limits in all the wells with the exception of N08C, indicating that the waters are at least 76 years old (L. I. Putman, unpublished data). It is therefore unlikely that seasonal recharge is reaching the fluids we sampled at a rate that would be relevant for replenishing DO in the wells. All wells were below the detection limits of the colorimetric formaldehyde test kit (<0.4 ppm).

DISCUSSION

Previous studies of microbial communities at CROMO have used marker gene analysis, shotgun metagenomics, microcosm experiments, and geochemical measure-
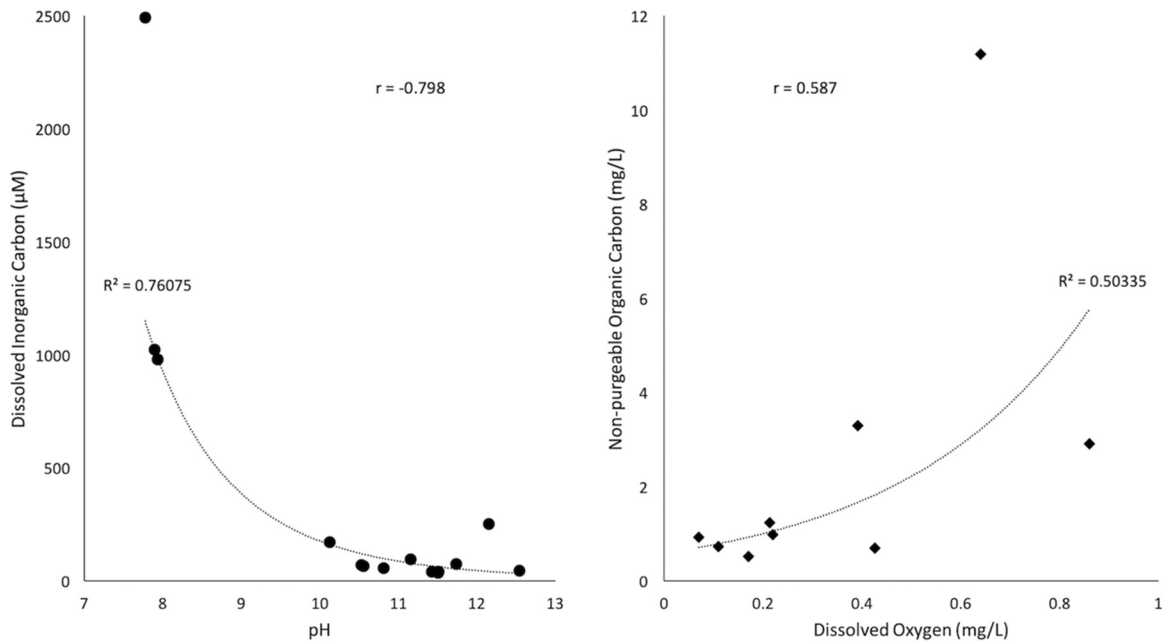
FIG 7 Scatterplots depicting dissolved inorganic carbon (DIC) versus pH (left) and nonpurgeable organic carbon (NPOC) versus dissolved oxygen (DO). The Pearson correlation coefficient (r) is indicated on each graph, as well as the R^2 of the trend lines.

ments to elucidate pathways of biogeochemical cycling and to describe the relationship between geochemical parameters and microbial diversity (6, 25). We expanded upon these efforts by combining metagenomics and metatranscriptomics with untargeted metabolomics to verify genomic expression and metabolic activity. We also analyzed MAGs produced by combining metagenomic data across all of the wells at CROMO to assess the completeness of key pathways. Using this integrated omics approach, we found evidence of multiple carbon fixation pathways, including the Calvin-Benson-Bassham cycle, the reverse TCA cycle, acetogenesis via the reductive acetyl-CoA pathway, and carbon assimilation via methylotrophy. We also determined that these pathways were expressed by the microbial communities in the serpentinite-hosted groundwater at CROMO. These pathways were associated with several key taxa associated with serpentinite-hosted microbial communities, including *Comamonadaceae*, *Clostridiales*, *Desulfuromodales*, and *Hydrogenophilales*.

Methane uptake and formaldehyde cycling. The addition of methane to microcosms from CROMO has been shown to stimulate microbial growth (25). In the previous metagenomic study of 4 samples by Twing et al. (6), no *pmoA* genes were detected; however, the gene *mxaF* was identified in the CSW1.3 metagenome. Our expanded search through nine metagenomes and four metatranscriptomes succeeded in finding the genes for particulate methane monooxygenase (PMO) and methanol dehydrogenase (MXA) in the CROMO microbial communities, though at low abundance in the higher-pH, more-reducing wells (Fig. 4), suggesting that the prior perceived absence of these genes could have been the result of a smaller data set. Aerobic methanotrophy therefore appears to be more likely to occur in lower-pH, higher-E_h wells. Two MAG bins contain the complete methane oxidation pathway, and another two contain a partial pathway (Fig. 3). No other known methane monooxygenases or methanol dehydrogenases were identified in the metagenomes or metatranscriptomes.

The prevalence of formaldehyde and formate assimilation pathways in CROMO reinforces the importance of methylotrophy as a carbon assimilation strategy in these microbial communities. There is a wealth of metagenomic and metatranscriptomic evidence that formaldehyde oxidation to formate and formaldehyde fixation to biomass can be performed by the microbial communities in CROMO (Fig. 2). Formaldehyde is an important intermediate in methylotrophic metabolic pathways, and though it is

highly toxic, methylotrophic microorganisms have developed a variety of functionally redundant modules for rapidly converting formaldehyde to less toxic compounds (32). Formaldehyde can also act as a carbon and electron shuttle between microbial subpopulations, in which it is produced by autotrophic primary producers as a by-product and then taken up by methylotrophic microbes and incorporated into biomass (33).

In methanotrophs, formate oxidation to CO_2 is carried out by the NADP-dependent formate dehydrogenase that converts CO_2 to formate in the Wood-Ljungdahl pathway, operating in reverse. As previously stated, the genes for this enzyme were not well represented (and in some cases missing) in the metagenomic data (Fig. 5). Formate can also be oxidized to CO_2 by an NAD-dependent formate dehydrogenase (FDH) as part of the degradation of oxalate; however, the complete suite of genes required to produce this enzyme was likewise at low coverage or partially missing, particularly in the most alkaline, most reducing wells (Fig. S2). This contrasts with subsurface wells in the Samail Ophiolite, where FDH-encoding genes were enriched in metagenomes recovered from alkaline and hyperalkaline groundwater (29). Therefore, while there is substantial metagenomic and metatranscriptomic evidence for the production of formate from formaldehyde by the microbial communities in CROMO, DNA/RNA evidence for the further oxidation of formate to CO_2 is lacking.

Bicarbonate fixation strategies in the metagenomes and metatranscriptomes. Despite the DIC concentrations being challengingly low for CO_2 fixation at the high pH of the deepest wells (5, 29, 34), we found complete Calvin-Benson-Bassham (CBB; Fig. S1) and reverse tricarboxylic acid (rTCA; Fig. 2) cycles in all nine of our metagenomes and all four metatranscriptomes. Sequences corresponding to the *rbcL* gene of the RuBisCo enzyme were previously detected in metagenomes from four wells (CSW1.1, CSW1.3, QV1.1, and QV1.2) in the earlier metagenomic study by Twing et al. (6). QV1.1 displays the most abundant coverage of the rTCA pathway in both its metagenome (607.71 ± 55.83 FPKM) and metatranscriptome (446.86 ± 149.59 FPKM) (Fig. 2), even though the concentration of dissolved inorganic carbon (DIC) in QV1.1 is lower than almost any other well (11.48 μM). Conversely, the presence/expression of the CBB cycle was roughly equal throughout all the wells (Fig. S1).

The prevalence of carbon-fixing populations in these wells presents a conundrum, as above pH 11 the carbonate anion, a molecule unable to be acted upon directly by known autotrophic pathways, is the dominant species of DIC (35). *Serpentinomonas* spp. belonging to the family *Comamonadaceae* are among the most abundant taxa in the highest-pH wells in CROMO (6). RuBisCO large and small subunits (encoded by *rbcL* and *rbcS*) were detected on *Comamonadaceae* contigs in our metagenomic data, suggesting the ability to fix carbon. A previous study found that *Serpentinomonas* require the addition of calcium carbonate (even when bicarbonate is also provided) for growth in culture and form aggregates on carbonate precipitates (34). Bicarbonate-fixing microbes may create a localized microenvironment of decreased pH where carbonate may be converted to bicarbonate, or they may have bicarbonate transporters that solubilize carbonate (for examples, see reference 36). We did not detect any ATP-binding cassette transporters for bicarbonate in our metagenomic data (see Table S4 at https://figshare.com/articles/Supplemental_Table_4_xlsx/ 11873754). Alternatively, formate may be used as a carbon source in the CBB cycle (37) or in the rTCA cycle (38). Formate could thus be produced from formaldehyde via formaldehyde oxidation and then used in assimilatory/anabolic pathways.

Evidence for the reductive acetyl-CoA pathway was also previously detected in several wells (6). In this study, we found a nearly complete homoacetogenic reductive acetyl-CoA pathway in the metagenomic and metatranscriptomic data from all but the two least alkaline wells, except for one gene. The gene that encodes the beta subunit of formate dehydrogenase (*fdhB*), which converts CO_2 to formate as the first step of the reductive acetyl-CoA pathway, represents <1 FPKM in every metagenome but two (N08A and N08B) (Fig. 5) and is present only in the metatranscriptomes of QV1.2 and N08B. Likewise, the genes for FDH are missing from every MAG bin that possesses the reductive acetyl-CoA pathway (Fig. 3, bottom). Yet this step of the pathway may be

unnecessary for the pathway to proceed if formate is being produced by other means, either through the oxidation of formaldehyde via the glutathione-dependent or tetrahydrofolate pathway or by abiotic production via Fischer-Tropsch-like reactions. Abiotically produced formate has been implicated as a possible carbon source in serpentinizing environments such as the Lost City (14, 23, 39) and the Samail Ophiolite (29). While formate-fixing reactions are not common, as formate has a relatively low reactivity compared to that of CO_2 or other carboxylic acids, formate can be used as the sole carbon source by some microorganisms (37). In the Samail Ophiolite, rates of assimilation or oxidation of formate were higher than rates of assimilation or dissimilation of bicarbonate or CO regardless of pH (29), suggesting that formate may be a preferred source of carbon and/or electrons. Formate concentrations in CROMO are similar to DIC concentrations or lower (Table S3; see also references 6 and 25), but the addition of formate stimulates microbial growth (25). We also detected two formate transporters in the metagenomic data: *fdhC*, which has been detected in putative formate-metabolizing bacteria in Lost City chimneys (36), was detected in two MAG bins (126 and 192, both *Clostridium*), and *focA* was detected in one MAG bin (222_8, *Actinobacter*) and across multiple metagenomes and all four metatranscriptomes (Table S1; see also Table S4 at https://figshare.com/articles/Supplemental_Table_4_xlsx/11873754).

As DIC becomes more limiting due to pH, bypassing the conversion of CO_2 to formate may be an important strategy for carbon fixation. Provided that sufficient reductant is available (for example, from the oxidation of H_2), assimilation of CO_2, CO, or formate can occur via the reductive acetyl-CoA pathway (40). The assimilation of formate via the acetogenic reductive acetyl-CoA pathway avoids ATP consumption, supports energy conservation through the use of multiple electron bifurcation mechanisms, and removes the need for an external electron acceptor apart from coassimilated CO_2 (37, 38). While the *cooS* gene for the conversion of CO_2 to CO is present (Fig. 5), this step could likely be bypassed as well in wells where the CO concentration is sufficiently high (Table S3). Evidence for carbon monoxide uptake (28) and the reductive acetyl-CoA pathway (27) as important carbon assimilation pathways was previously found in the Tablelands, Canada, and The Cedars, CA, respectively. In the Samail Ophiolite, it appears that carbon limitation outweighs energy limitation for the autotrophic members of microbial communities at high pH (29), making carbon compounds important as both a feedstock for biosynthesis and a source of electrons.

A putative metabolic pathway for the assimilation of methane into biomass is outlined in Fig. 8. Methane oxidation to formaldehyde is carried out by the *Methylococcales* and *Methylophilales*, and then formaldehyde is either fixed to biomass via the RuMP and serine pathways or oxidized to formate via the THF (*Burkholderiales*, *Deinococcales*, *Hydrogenophilales*, *Methylophilales*, *Rhizobiales*, and *Rhodocyclales*) and glutathione-dependent (*Burkholderiales*, *Rhizobiales*, *Rhodobacterales*, and *Rhodocyclales*) pathways. Formate can then be assimilated into biomass via the reductive acetyl-CoA pathway by the *Clostridiales* or *Desulfuromodales*, as shown, or through the CBB or rTCA cycle (37, 38).

Methanogenesis. While methanogenic archaea have not been detected in 16S rRNA data obtained from CROMO (6), methane isotopologue data suggest a thermogenic/biogenic source (18), and metagenomic and metatranscriptomic evidence suggests the possibility of methanogenesis by multiple pathways in at least some of the wells. However, the metagenomes do not include the essential backbone of the methanogenesis pathway, including the crucial methyl coenzyme M reductase alpha subunit (*mcrA*) gene, required for making methane. The *mcrA* gene was detected in the metatranscriptomes, but other key genes for methane production were absent (Table S1; see also Table S4 at https://figshare.com/articles/Supplemental_Table_4_xlsx/11873754). Low abundances of *mcr* genes versus transcripts may be the result of methanogenesis under stress, such as hydrogen limitation (41). The presence of some methanogenesis-related genes indicates that some populations at CROMO may metabolize small organic compounds such as acetate and methanol. Whether these
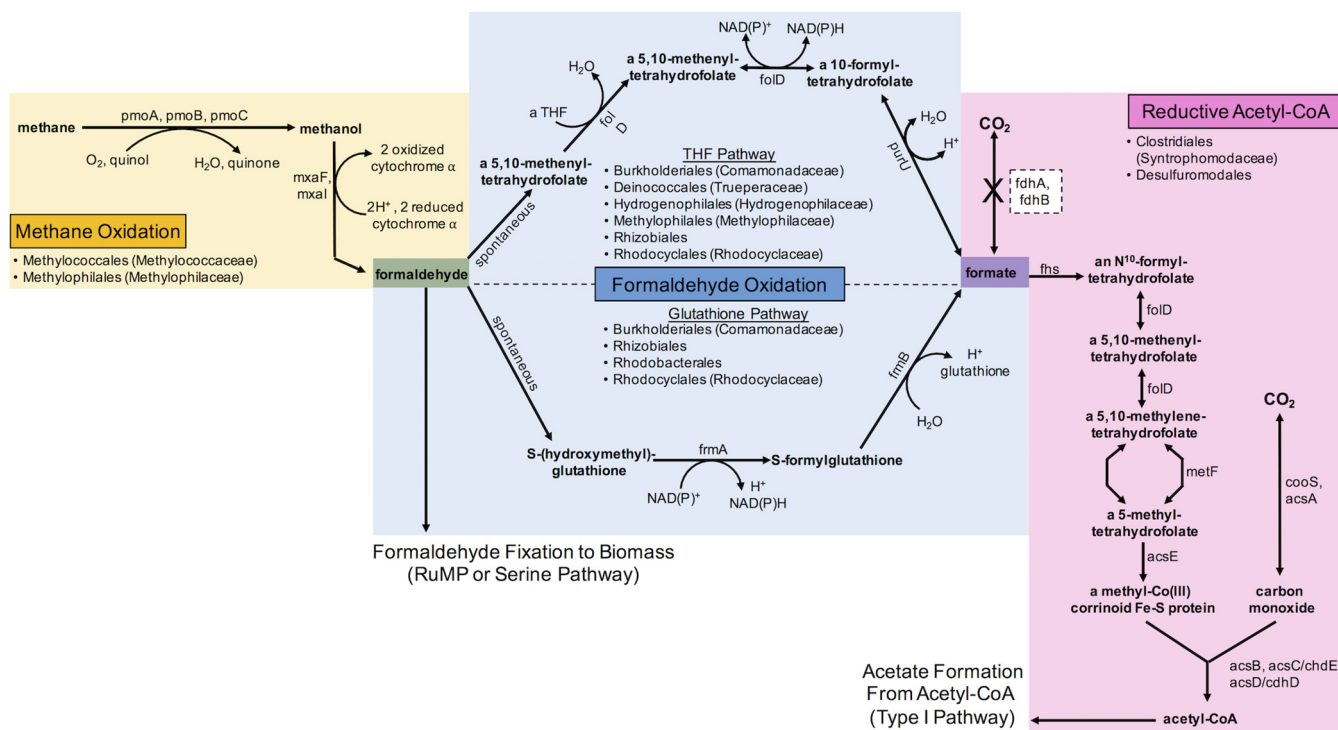
FIG 8 Putative pathway for carbon assimilation in the CROMO wells. Formaldehyde is either converted directly to biomass via the RuMP or serine pathway or oxidized to formate, which is then fed into the homoacetogenic reductive acetyl-CoA pathway. The pathway for methane conversion to formaldehyde is shown but appears to be rarely expressed in most of the wells.

substrates are converted to methane by methanogens remains unclear. Biological methanogenesis appears to be occurring in other continental serpentinizing sites, such as the Voltri Massif (26) and The Cedars (19). In the Tablelands, however, no evidence of biological methane production has been detected (28), and methane production genes in metagenomes recovered from the Samail Ophiolite are likewise sparse (29). Methane cycling in ophiolite-hosted microbial communities therefore seems to vary from site to site, and further work will be required to uncover the geochemical conditions that drive its production and uptake.

Searching the metabolome for intermediates. Intermediates of formaldehyde oxidation and formaldehyde assimilation pathways did not match the masses and retention times of any of the features in our metabolomics data. In negative ion mode, glutathione derivatives should have appeared in the data as deprotonated or double-negatively charged molecules, but we did not detect masses that matched either of those possibilities. This may be due to the relatively low concentrations of DOM (3 to 13 mg/liter) (42), as well as bias introduced by our extraction method (43). Tetrahydrofolate compounds would likely have been more easily detected in positive ion mode; however, we ran our samples in negative mode in the hopes of capturing the broadest array of metabolites possible with no preconceived notions of what compounds would be present in the data. Metabolites may also fragment or form noncovalent interactions with other compounds upon entering the mass spectrometer, making identification of particular intermediates challenging (44).

Additionally, low cell abundances in alkaline groundwater make it difficult to obtain enough biomass during filtration to accurately depict the entire intracellular metabolome of the microbial community. An ideal target of total carbon would be 0.3 mg (45). Assuming each cell contains 30 fg of carbon (46), and with an average cell abundance of 2×10^5 cells/ml (6; unpublished data), 50 liters of water would need to be filtered. Although CSW1.1 and QV1.1 have high fluid outputs compared to most of the other CROMO wells, filtration limitations such as clogging due to carbonate precipitates,

keeping the fluid cold to prevent further chemical alteration of the metabolites, and prevention of contamination all precluded our ability to filter such large amounts of water. We were also unable to detect formaldehyde using colorimetric methods in any of the wells. If formaldehyde and its intermediates are cycled rapidly within the cell, it may be difficult to identify and quantify these compounds with the techniques used in this study. Interestingly, while the two wells with the highest pH (QV1.1 and CSW1.1) share highly similar microbial communities (Fig. S3) with similar metabolic capabilities, the DOC pools observed in these wells appeared significantly different in character (Fig. 6). Putative annotations of unique metabolite features in the DOC pools include dipeptides, fatty acids, vitamins, phosphorylated metabolites, and secondary metabolites. This observation underscores the need for further characterization of the available carbon in serpentinite-hosted aquifers like CROMO through metabolomics techniques.

Conclusion. Through an in-depth analysis of metagenomic and metatranscriptomic data taken over several time points, we have determined that methanotrophy, the reductive acetyl-CoA pathway, the reverse TCA cycle, and the Calvin-Benson-Bassham cycle may all be important means of carbon fixation and assimilation for microbes living in the Coast Range Ophiolite groundwater. We also identified several mechanisms by which microbes in the hyperalkaline fluids of deep serpentinizing rock may overcome limiting concentrations of DIC, including the use of formate and formaldehyde as carbon sources for homoacetogenesis and the production of biomass. Further optimization of metabolomics techniques for this environment may be of use to track the fate of carbon and delineate between abiotic and biotic processes in serpentinizing systems such as CROMO.

MATERIALS AND METHODS

Site description. The Coast Range Ophiolite is a 155 million- to 170 million-year-old portion of oceanic crust tectonically emplaced on the North American continental margin in northern California. Trapped Cretaceous seawater and circulating meteoric water reacts with ultramafic rock (47, 48), as evidenced by numerous calcium hydroxide rich springs throughout the formation (49). Previous studies of continental serpentinite environments have sampled similar springs in the Tablelands, Newfoundland, Canada (22, 24) and The Cedars, CA (34), providing information on the surface expression of subsurface biogeochemical processes. In order to directly access the serpentinite subsurface environment and monitor groundwater biogeochemistry with minimal contamination, a microbial observatory was established in the Coast Range Ophiolite in 2011 at the University of California (UC)-Davis McLaughlin Natural Reserve (Lower Lake, CA) (50). The Coast Range Ophiolite Microbial Observatory (CROMO) consists of two sets of wells 1.4 km apart—the Quarry Valley (QV) and Core Shed Wells (CSW)—each with an uncased main borehole (designated "1.1") surrounded by two or four monitoring wells, respectively. In addition to the eight-well array drilled in 2011, four geographically close wells are also sampled for context: N08A, N08B, and N08C at Quarry Valley, drilled by the Homestake Mining Company, Inc., and CSWold, drilled by UC-Davis. All of the wells are cased with polyvinyl chloride (PVC), with the exception of the 1.1 wells, which are cased only to bedrock, leaving the bottom of the hole uncased. A detailed hydrogeological description of the wells is available in the work of Ortiz et al. (31).

Extraction of DNA and RNA. For each environmental sample from the CROMO wells, 4 liters of fluid was pumped via low-flow bladder pump (50) and immediately filtered through Sterivex 0.2-μm filter cartridges (Millipore, Billerica, MA) using a portable peristaltic pump (Masterflex, Cole Parmer, Vernon Hills, IL). Cartridges were kept on ice during filtration, immediately stored in liquid nitrogen upon completion, shipped to the home laboratory, and stored at $-80°C$ until processing. Total genomic DNA extractions were completed as previously described by Twing et al. (6) and briefly described here. Freeze-thaw cycles and lysozyme/proteinase K treatment were performed to lyse cells, followed by purification with phenol-chloroform, precipitation using ethanol, and purification using QiaAmp (Qiagen, Hilden, Germany) columns according to the manufacturer's instructions. A Qubit 2.0 fluorometer (Thermo Fisher, Waltham, MA) was used to quantify extracted DNA using a Qubit double-stranded DNA (dsDNA) high-sensitivity assay kit.

Extractions for RNA from CROMO wells were performed as described previously, with slight modifications (51, 52). Briefly, frozen 0.2-μm Sterivex filter cartridges were broken open, cut into four equal pieces, and divided into two screw-cap Eppendorf tubes containing phenol, 20% sodium dodecyl sulfate, 5$\times$ low-pH buffer, and 0.2 to 0.5 g of baked zirconium beads. Samples were bead-beaten for 3 min, heated in a 60°C water bath for 10 min, bead beaten again for 3 min, and centrifuged at 4°C and 18,407 $\times$ g to separate phases. Supernatant was transferred to a fresh Eppendorf tube and chilled. A 1$\times$ concentration of low-pH buffer was added to the remaining sample, and bead-beating was repeated. Supernatants were combined, and phenol, 1:1 phenol-chloroform, and chloroform were added in series with vortexing and centrifugation in between. Between steps, aqueous phases were transferred to clean Eppendorf tubes. The final aqueous phase was transferred to a clean Eppendorf tube with additions of ammonium acetate, isopropanol, and magnesium chloride before vortexing and incubation at $-20°C$

overnight. Samples were centrifuged for 30 min at 4°C, washed with ethanol, and dried under vacuum before suspension in RNase-free water and storage at −80°C until analyzed.

Metagenome and metatranscriptome analyses. Metagenome and metatranscriptome sequencing was conducted at the Joint Genome Institute (JGI) on an Illumina HiSeq2000 instrument. A Covaris LE220 ultrasonicator was used to shear DNA samples into 270-bp fragments, and size selection was performed using solid-phase reversible immobilization (SPRI) magnetic beads (53). DNA fragments were end-repaired, A-tailed, and ligated with Illumina-compatible adapters with barcodes unique for each library. KAPA Biosystem's next-generation sequencing library quantitative PCR (qPCR) kit and Roche LightCycler 280 reverse transcription-PCR (RT-PCR) instrument were used to quantify libraries. Ten-library pools were assembled and prepared for Illumina sequencing in one lane each. Clustered flow cells were produced using a TruSeq paired-end cluster kit (v.3) and Illumina's cBot instrument. The Illumina HiSeq2000 instrument was utilized with a TruSeq SBS sequencing kit (v.3) and a 2 × 150 indexed run recipe to sequence the samples. The raw sequence reads were trimmed by the JGI with a minimum quality score cutoff of 10 and to remove adapters. These trimmed reads from CROMO wells were previously reported by Twing et al. (6), but additional quality filtering and a new assembly, distinct from the JGI assembly reported by Twing et al., were used for this study.

The trimmed reads from the JGI were subjected to an additional quality screen to trim 3′ adapters with cutadapt v.1.15 (54), to remove replicate sequences, and to trim sequences again with a threshold of 20 along a sliding window of 6 bases with qtrim (part of the seq-qc package [https://github.com/Brazelton-Lab/seq-qc]). All CROMO metagenomes and metatranscriptomes were pooled together for a master CROMO assembly computed with Ray Meta v.2.3.1 (55). Phylogenetic affiliation of contigs was assigned using PhyloPythiaS+ (56). Metagenome and metatranscriptome short reads were mapped to the pooled assembly using Bowtie2 v.2.2.6 (57). The Prokka pipeline (58) was used for gene prediction and functional annotation of contigs. The arguments −metagenome and −proteins were used with Prokka v.1.12 to indicate that genes should be predicted with the implementation of Prodigal v.2.6.2 (59) optimized for metagenomes as described by Twing et al. (6). Metagenome-assembled genome (MAG) bins were constructed with ABAWACA (https://github.com/CK7/abawaca) using tetranucleotide frequencies and differential abundance as measured by Bowtie2-mapped read abundances. Bin quality was computed with CheckM (60), and only high-quality MAG bins are reported here (>50% completeness and <10% contamination, as defined by Bowers et al. [61]). The completeness and contamination of some bins were improved by refinement with Binsanity (62). Taxonomic classifications of MAGs were assigned using GTDB-Tk (https://github.com/Ecogenomics/GTDBTk) (60).

Predicted protein-coding sequences were annotated by searching the Kyoto Encyclopedia of Genes and Genomes (KEGG [63, 64]) release v. 83.2 within Prokka. HTSeq v.0.6.1 (65) and were used to calculate predicted protein abundances. The abundances of predicted protein functions in all CROMO metagenomes and metatranscriptomes were normalized to predicted protein size and metagenome size. Data reported here are in units of metagenome fragments per kilobase of predicted protein sequence per million mapped reads (FPKM). Abundances of metabolic pathways were obtained by mapping KEGG protein identifiers (IDs) and their normalized counts onto the FOAM ontology (66) with MinPath (67) as implemented in HUMAnN2 v.0.6.0 (68).

The complete translated genome assemblies of *Serpentinomonas* isolates from The Cedars (*Comamonadaceae* bacterium A1, B1, and H1 [69]) were obtained as amino acid FASTA files from RefSeq (accession numbers GCF_000828895.1, GCF_000828915.1, and GCF_000696225.1). Open reading frames (ORFs) were annotated as KEGG orthologs (KOs) using BlastKOALA (70). KEGG modules were annotated within genomes using the KEGG module reconstruction tool. MAGs were translated to protein sequences using Prodigal (59) and annotated in the same manner. Modules within the three cultured representative genomes and the MAGs were then directly compared in order to compare pathways identified in available cultivars from The Cedars to pathways in the uncultured communities hosted in CROMO's hyperalkaline groundwater. A phylogenetic tree comparing MAGs to the *Serpentinomonas* isolates was generated using SpeciesTree in KBase (71), in which relatedness is determined by alignment similarity to a selected subset of COG (Clusters of Orthologous Groups) domains and aligned using FastTree 2 (72) with 100 bootstrap replicates.

Metabolite extraction and analysis. Three 4-liter samples of groundwater from QV1.1 and CSW1.1 were taken in June 2016 and filtered using 0.22-μm by 47-mm polytetrafluoroethylene (PTFE) filters on a glass vacuum filtration tower. Filters were frozen in liquid nitrogen and stored at −80°C until extraction (42, 45). Approximately 30 ml of filtrate was set aside and frozen at −20°C for total organic carbon analysis on a Shimadzu total organic carbon analyzer. Formaldehyde was measured using a CHEMetrics Vacu-vials colorimetric kit (detection range, 0.4 to 8.0 ppm). The remaining filtrate was acidified to pH 2 to 3 using concentrated HCl and stored at 4°C for 5 days until further analysis by solid-phase extraction (SPE).

Frozen filters were cut into small pieces over a muffled piece of aluminum foil, using methanol-washed scissors, and then transferred into muffled amber glass vials. Two milliliters of cold extraction solvent (40:40:20 acetonitrile-methanol-0.1 M formic acid) was added to each vial, and the samples were sonicated for 10 min (73). Extracts were transferred to a centrifuge tube via pipette and centrifuged at 20,000 × g for 5 min at 4°C. The supernatant was then transferred to a new amber glass vial, neutralized with 6 M ammonium hydroxide, and vacuum centrifuged to dryness. Samples were dissolved in 495 μl of 95:5 water-acetonitrile with 5 μl of 5-μg/ml biotin-(ring-6,6-d$_2$) added as an injection standard (42, 74).

Dissolved organics were captured from the acidified filtrate on SPE Bond Elut-PPL cartridges (Agilent Technologies) and eluted with 100% methanol into muffled amber glass vials (71, 75). SPE-PPL cartridges were rinsed with methanol immediately before use. The supernatant was passed through the cartridge using 1/8-in. by 1/4-in. PTFE tubing to pull the supernatant into the cartridge to minimize the possibility

of contamination from plastic leaching, while Viton tubing was used to remove the discarded flow-through via peristaltic pump (flow rate not exceeding 40 ml/min). The cartridges were then rinsed with at least 2 cartridge volumes of 0.01 M HCl to remove salt, and the sorbent was air dried for 5 min. Dissolved organic matter (DOM) was eluted from the cartridge with 2 ml of methanol via gravity flow into a muffled amber glass vial. The eluate was then vacuum centrifuged to dryness. Precipitated samples were dissolved in 495 μl of 95:5 water-acetonitrile, with 5 μl of 5-μg/ml biotin-(ring-6,6-d$_2$).

The above-described protocol was repeated for 4 liters of Milli-Q water, and organics extracted from the filter and filtrate as described above were used as extraction blanks. A total of 495 μl of 95:5 water-acetonitrile plus 5 μl of 5-μg/ml biotin-(ring-6,6-d$_2$) was also run as a blank.

Samples were analyzed using tandem liquid chromatography-mass spectrometry (LC-MS/MS) at the Michigan State University Metabolomics Core Facility. Triplicate samples from each well were separated chromatographically in a Acquity ultraperformance liquid chromatography (UPLC) BEH C$_{18}$ column (1.7 μm; 2.1 mm by 50 mm) using a polar/nonpolar 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% solvent A/solvent B to 1.0%/99.0% solvent A/solvent B over 9 min, held an additional 3 min at 1.0%/99.0% solvent A/solvent B, and then reversed to 99.9%/0.1% solvent A/solvent B and held another 3 min. At a rate of 400 μl/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 to 1500 Da). DOM levels were too low to allow distinction between samples and blanks from the first run, so triplicate samples were pooled and run as before. After data was converted to centroid mode, raw data files were converted from proprietary Waters RAW format into XML-based mzML format using ProteoWizard (76). Data were processed using XCMS (77) parametrized by AutoTuner (78). Isotopes and adducts of each feature were identified using CAMERA (79). 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. The table of features was subject to quality assurance through blank correction by removing any feature from a sample that also appeared within the blank control (described above) (80). Features were first putatively annotated as KEGG compounds using Mummichog (81). Chemical standards of putative annotations were purchased whenever available to check the retention time to increase confidence of annotation from level 1 to 2 (82). Due to the great complexity of the secondary mass spectral data and insufficient coverage of mass spectral databases for natural organic matter, MS^2 information could not be readily used to increase confidence of annotation.

Data availability. CROMO metagenome and metatranscriptome sequences are publicly available in the NCBI SRA with the BioProject IDs PRJNA410019, PRJNA410020, PRJNA410022 to PRJNA410033, PRJNA410035 to PRJNA410037, PRJNA410054, PRJNA410057, PRJNA410286, PRJNA410403, PRJNA410404, PRJNA410553 to PRJNA410555, and PRJNA410557. Raw spectrometry data are available in the MetaboLights database under study identifier MTBLS1260.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.2 MB.
FIG S2, PDF file, 0.1 MB.
FIG S3, PDF file, 0.8 MB.
FIG S4, PDF file, 0.1 MB.
TABLE S1, XLSX file, 0.1 MB.
TABLE S2, DOCX file, 0.02 MB.
TABLE S3, CSV file, 0.01 MB.

REFERENCES

1. McCollom TM. 2007. Geochemical constraints on sources of metabolic energy for chemolithoautotrophy in ultramafic-hosted deep-sea hydrothermal systems. Astrobiology 7:933–950. https://doi.org/10.1089/ast.2006.0119.
2. Amend JP, McCollom TM, Hentscher M, 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. https://doi.org/10.1016/j.gca.2011.07.041.
3. Cardace D, Meyer Dombard DR, Woycheese KM, Arcilla CA. 2015. Feasible metabolisms in high pH springs of the Philippines. Front Microbiol 6:10. https://doi.org/10.3389/fmicb.2015.00010.
4. Barnes I, O'Neil JR, Trescases JJ. 1978. Present day serpentinization in New Caledonia, Oman and Yugoslavia. Geochim Cosmochim Acta 42:144–145. https://doi.org/10.1016/0016-7037(78)90225-9.
5. Schrenk MO, Brazelton WJ, Lang SQ. 2013. Serpentinization, carbon, and deep life. Rev Minerol Geochem 75:575–606. https://doi.org/10.2138/rmg.2013.75.18.
6. Twing KI, Brazelton WJ, Kubo MD, Hyer AJ, Cardace D, Hoehler TM,

McCollom TM, Schrenk MO. 2017. Serpentinization-influenced groundwater harbors extremely low diversity microbial communities adapted to high pH. Front Microbiol 8:308. https://doi.org/10.3389/fmicb.2017.00308.

7. McCollom TM, Seewald JS. 2001. A reassessment of the potential for reduction of dissolved CO_2 to hydrocarbons during serpentinization of olivine. Geochim Cosmochim Acta 65:3769–3778. https://doi.org/10.1016/S0016-7037(01)00655-X.

8. Charlou JL, Donval JP, Fouquet Y, Jean-Baptiste P, Holm N. 2002. Geochemistry of high H_2 and CH_4 vent fluids issuing from ultramafic rocks at the Rainbow hydrothermal field (36 Deg14′N, MAR). Chem Geol 191:345–359. https://doi.org/10.1016/S0009-2541(02)00134-1.

9. Proskurowski G, Lilley MD, Seewald JS, Früh-Green GL, Olson EJ, Lupton JE, Sylva SP, Kelley DS. 2008. Abiotic hydrocarbon production at Lost City Hydrothermal Field. Science 319:604–607. https://doi.org/10.1126/science.1151194.

10. McCollom TM. 2013. Laboratory simulations of abiotic hydrocarbon formation in Earth's deep subsurface. Rev Mineral Geochem 75:467–494. https://doi.org/10.2138/rmg.2013.75.15.

11. Sephton MA, Hazen RM. 2013. On the origins of deep hydrocarbons. Rev Mineral Geochem 75:449–465. https://doi.org/10.2138/rmg.2013.75.14.

12. Ménez B, Pasini V, Brunelli D. 2012. Life in the hydrated suboceanic mantle. Nature Geosci 5:133–137. https://doi.org/10.1038/ngeo1359.

13. Sforna MC, Brunelli D, Pisapia C, Pasini V, Malferrari D, Ménez B. 2018. Abiotic formation of condensed carbonaceous matter in the hydrating oceanic crust. Nat Commun 9:5049. https://doi.org/10.1038/s41467-018-07385-6.

14. Lang SQ, Butterfield DA, Schulte M, Kelley DS, Lilley MD. 2010. Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. Geochim Cosmochim Acta 74:941–952. https://doi.org/10.1016/j.gca.2009.10.045.

15. Delacour A, Früh-Green GL, Bernasconi SM, Schaeffer P, Kelley DS. 2008. Carbon geochemistry of serpentinites in the Lost City Hydrothermal System (30°N, MAR). Geochim Cosmochim Acta 72:3681–3702. https://doi.org/10.1016/j.gca.2008.04.039.

16. Bradley A, Summons RE. 2010. Multiple origins of methane at the Lost City Hydrothermal Field. Earth Planet Sci Lett 297:34–41. https://doi.org/10.1016/j.epsl.2010.05.034.

17. Etiope G. 2017. Abiotic methane in continental serpentinization sites: an overview. Procedia Earth Planet Sci 17:9–12. https://doi.org/10.1016/j.proeps.2016.12.006.

18. Wang DT, Gruen DS, Lollar BS, Hinrichs K-U, Stewart LC, Holden JF, Hristov AN, Pohlman JW, Morrill PL, Könneke M, Delwiche KB, Reeves EP, Sutcliffe CN, Ritter DJ, Seewald JS, McIntosh JC, Hemond HF, Kubo MD, Cardace D, Hoehler TM, Ono S. 2015. Nonequilibrium clumped isotope signals in microbial methane. Science 348:428–431. https://doi.org/10.1126/science.aaa4326.

19. Kohl L, Cumming E, Cox A, Rietze A, Morrisey L, Lang SQ, Richter A, Suzuki S, Nealson KH, Morrill PL. 2016. Exploring the metabolic potential of microbial communities in ultra-basic, reducing springs at The Cedars, CA, USA: experimental evidence of microbial methanogenesis and heterotrophic acetogenesis. J Geophys Res Biogeosci 121:1203–1220. https://doi.org/10.1002/2015JG003233.

20. Schrenk MO, Kelley DS, Bolton SA, Baross JA. 2004. Low archaeal diversity linked to subseafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. Environ Microbiol 6:1086–1095. https://doi.org/10.1111/j.1462-2920.2004.00650.x.

21. Brazelton WJ, Mehta MP, Kelley DS, Baross JA. 2011. Physiological differentiation within a single-species biofilm fueled by serpentinization. mBio 2:e00127-11. https://doi.org/10.1128/mBio.00127-11.

22. Brazelton WJ, Nelson B, Schrenk MO. 2012. Metagenomic evidence for H_2 oxidation and H_2 production by serpentinite-hosted subsurface microbial communities. Front Microbiol 2:268. https://doi.org/10.3389/fmicb.2011.00268.

23. Lang SQ, Früh-Green GL, Bernasconi SM, Lilley MD, Proskurowski G, Méhay S, Butterfield DA. 2012. Microbial utilization of abiogenic carbon and hydrogen in a serpentinite-hosted system. Geochim Cosmochim Acta 92:82–99. https://doi.org/10.1016/j.gca.2012.06.006.

24. Brazelton WJ, Morrill PL, Szponar N, Schrenk MO. 2013. Bacterial communities associated with subsurface geochemical processes in continental serpentinite springs. Appl Environ Microbiol 79:3906–3916. https://doi.org/10.1128/AEM.00330-13.

25. Crespo-Medina M, Twing KI, Kubo MD, Hoehler TM, Cardace D, McCollom TM, Schrenk MO. 2014. Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. Front Microbiol 5:604. https://doi.org/10.3389/fmicb.2014.00604.

26. Brazelton WJ, Thornton CN, Hyer A, Twing KI, Longino AA, Lang SQ, Lilley MD, Früh-Green GL, Schrenk MO. 2017. Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy. PeerJ 5:e2945. https://doi.org/10.7717/peerj.2945.

27. Suzuki S, Ishii S, Hoshino T, Rietze A, Tenney A, Morrill PL, Inagaki F, Kuenen JG, Nealson KH. 2017. Unusual metabolic diversity of hyperalkaliphilic microbial communities associated with subterranean serpentinization at The Cedars. ISME J 11:2584–2598. https://doi.org/10.1038/ismej.2017.111.

28. Morrill PL, Brazelton WJ, Kohl L, Rietze1 A, Miles SM, Kavanagh H, Schrenk MO, Ziegler SE, Lang SQ. 2014. Investigations of potential microbial methanogenic and carbon monoxide utilization pathways in ultra-basic reducing springs associated with present-day continental serpentinization: the Tablelands, NL, CAN. Front Microbiol 5:613. https://doi.org/10.3389/fmicb.2014.00613.

29. Fones EM, Colman DR, Kraus EA, Nothaft DB, Poudel S, Rempfert KR, Spear JS, Templeton AS, Boyd ES. 2019. Physiological adaptations to serpentinization in the Samail Ophiolite, Oman. ISME J 13:1750–1762. https://doi.org/10.1038/s41396-019-0391-2.

30. Goenrich M, Bartoschek S, Hagemeier CH, Griesinger C, Vorholt JA. 2002. A glutathione-dependent formaldehyde-activating enzyme (Gfa) from *Paracoccus denitrifcans* detected and purified via two-dimensional proton exchange NMR spectroscopy. J Biol Chem 277:3069–3072. https://doi.org/10.1074/jbc.C100579200.

31. Ortiz E, Tominaga M, Cardace D, Schrenk MO, Hoehler TM, Kubo MD, Rucker DF. 2018. Geophysical characterization of serpentinite hosted hydrogeology at the McLaughlin Natural Reserve, Coast Range Ophiolite. Geochem Geophys Geosyst 19:114–131. https://doi.org/10.1002/2017GC007001.

32. Marx CJ, Van Dien SJ, Lidstrom ME. 2005. Flux analysis uncovers key role of functional redundancy in formaldehyde metabolism. PLoS Biol 3:e16. https://doi.org/10.1371/journal.pbio.0030016.

33. Moran JJ, Whitmore LM, Isern NG, Romine MF, Riha KM, Inskeep WP, Kreuzer HW. 2016. Formaldehyde as a carbon and electron shuttle between autotrophic and heterotrophic populations in acidic hydrothermal vents of Norris Geyser Basin, Yellowstone National Park. Extremophiles 20:291–299. https://doi.org/10.1007/s00792-016-0821-2.

34. Suzuki S, Ishii S, Wu A, Cheung A, Tenney A, Wanger G, Kuenen JG, Nealson KH. 2013. Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinizing ecosystem. Proc Natl Acad Sci U S A 110:15336–15341. https://doi.org/10.1073/pnas.1302426110.

35. Berg IA. 2011. Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. Appl Environ Microbiol 77:1925–1936. https://doi.org/10.1128/AEM.02473-10.

36. Engel AS, Porter ML, Stern LA, Quinlan S, Bennett PC. 2004. Bacterial diversity and ecosystem function of filamentous microbial mats from aphotic (cave) sulfidic springs dominated by chemolithoautotrophic "*Epsilonproteobacteria*." FEMS Microbiol Ecol 51:31–53. https://doi.org/10.1016/j.femsec.2004.07.004.

37. Bar-Even A. 2016. Formate assimilation: the metabolic architecture of natural and synthetic pathways. Biochemistry 55:3851–3863. https://doi.org/10.1021/acs.biochem.6b00495.

38. Cotton CAR, Edlich-Muth C, Bar-Even A. 2018. Reinforcing carbon fixation: CO_2 reduction and supporting carboxylation. Curr Opin Biotechnol 49:49–56. https://doi.org/10.1016/j.copbio.2017.07.014.

39. Lang SQ, Früh-Green GL, Bernasconi SM, Brazelton WJ, Schrenk MO, McGonigle JM. 2018. Deeply-sourced formate fuels sulfate reducers but not methanogens at Lost City hydrothermal field. Sci Rep 8:755. https://doi.org/10.1038/s41598-017-19002-5.

40. Takami H, Noguchi H, Takaki Y, Uchiyama I, Toyoda A, Nishi S, Chee GJ, Arai W, Nunoura T, Itoh T, Hattori M, Takai K. 2012. A deeply branching thermophilic bacterium with an ancient acetyl-CoA pathway dominates a subsurface ecosystem. PLoS One 7:e30559. https://doi.org/10.1371/journal.pone.0030559.

41. Kato S, Kosaka T, Watanabe K. 2008. Comparative transcriptome analysis of Methanothermobacter thermautotrophicus to different environmental stimuli. Environ Microbiol 10:893–905. https://doi.org/10.1111/j.1462-2920.2007.01508.x.

42. Fiore CL, Longnecker K, Kido Soule MC, Kujawinski E. 2015. Release of ecologically relevant metabolites by the cyanobacterium *Synechococcus*

elongatus CCMP 1631. Environ Microbiol 17:3949–3963. https://doi.org/10.1111/1462-2920.12899.

43. Johnson WM, Kido Soule MC, Kujawinski EB. 2017. Interpreting the impact of matrix on extraction efficiency and instrument response in a targeted metabolomics method. Limnol Oceanogr Methods 15:417–428. https://doi.org/10.1002/lom3.10181.

44. Zamboni N, Saghatelian A, Patti GJ. 2015. Defining the metabolome: size, flux, and regulation. Mol Cell 58:699–706. https://doi.org/10.1016/j.molcel.2015.04.021.

45. Kido Soule MC, Longnecker K, Johnson WM, Kujawinksi EB. 2015. Environmental metabolomics: analytical strategies. Mar Chem 177:274–387.

46. Fukuda R, Ogawa H, Nagata T, Koike I. 1998. Direct determination of carbon and nitrogen contents of natural bacterial assemblages in marine environments. Appl Environ Microbiol 64:3352–3358. https://doi.org/10.1128/AEM.64.9.3352-3358.1998.

47. Peters EK. 1993. D-¹⁸O enriched waters of the Coast Range Mountains, northern California: connate and ore-forming fluids. Geochim Cosmochim Acta 57:1093–1104. https://doi.org/10.1016/0016-7037(93)90043-V.

48. Shervais JW, Murchey BL, Kimbrough DL, Renne PR, Hanan B. 2005. Radioisotopic and biostratigraphic age relations in the Coast Range Ophiolite, northern California: implications for the tectonic evolution of the Western Cordillera. Geol Soc Am Bull 177:633–653. https://doi.org/10.1130/B25443.1.

49. Barnes I, O'Neil JR. 1969. The relationship between fluids in some fresh Alpine-type ultramafics and possible modern serpentinization, Western United States. Geol Soc America Bull 80:1947–1960. https://doi.org/10.1130/0016-7606(1969)80[1947:TRBFIS]2.0.CO;2.

50. Cardace D, Hoehler T, McCollom T, Schrenk M, Carnevale D, Kubo MD, Twing K. 2013. Establishment of the Coast Range Ophiolite Microbial Observatory (CROMO): drilling objectives and preliminary outcomes. Sci Dril 16:45–55. https://doi.org/10.5194/sd-16-45-2013.

51. Lin C, Stahl DA. 1995. Taxon-specific probes for the cellulolytic genus Fibrobacter reveal abundant and novel equine-associated populations. Appl Environ Microbiol 61:1348–1351. https://doi.org/10.1128/AEM.61.4.1348-1351.1995.

52. MacGregor BJ, Moser DP, Alm E, Nealson KH, Stahl DA. 1997. Crenarchaeota in Lake Michigan sediment. Appl Environ Microbiol 63:1178–1181. https://doi.org/10.1128/AEM.63.3.1178-1181.1997.

53. Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res 22:939–946. https://doi.org/10.1101/gr.128124.111.

54. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.

55. Boisvert S, Raymond F, Godzaridis E, Laviolette F, Corbeil J. 2012. Ray Meta: scalable de novo metagenome assembly and profiling. Genome Biol 13:R122. https://doi.org/10.1186/gb-2012-13-12-r122.

56. Gregor I, Dröge J, Schirmer M, Quince C, McHardy AC. 2016. PhyloPythiaS+: a self-training method for the rapid reconstruction of low-ranking taxonomic bins from metagenomes. PeerJ 4:e1603. https://doi.org/10.7717/peerj.1603.

57. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

58. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

59. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.

60. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996–1004. https://doi.org/10.1038/nbt.4229.

61. Bowers RM, Genome Standards Consortium, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooseph S, Sutton G, Glöckner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu W-T, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Lapidus A, Meyer F, Yilmaz P, Parks DH, Eren AM, Schriml L, Banfield JF, Hugenholtz P, Woyke T. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of

bacteria and archaea. Nat Biotechnol 35:725–731. https://doi.org/10.1038/nbt.3893.

62. Graham ED, Heidelberg JF, Tully BJ. 2017. BinSanity: unsupervised clustering of environmental microbial assemblies using coverage and affinity propagation. PeerJ 5:e3035. https://doi.org/10.7717/peerj.3035.

63. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. 1999. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 27:29–34. https://doi.org/10.1093/nar/27.1.29.

64. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:4.

65. Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31:166–169. https://doi.org/10.1093/bioinformatics/btu638.

66. Prestat E, David MM, Hultman J, Taş N, Lamendella R, Dvornik J, Mackelprang R, Myrold DD, Jumpponen A, Tringe SG, Holman E, Mavromatis K, Jansson JK. 2014. FOAM (Functional ontology Assignments for Metagenomes): a Hidden Markov model (HMM) database with environmental focus. Nucleic Acids Res 42:e145. https://doi.org/10.1093/nar/gku702.

67. Ye Y, Doak TG. 2009. A parsimony approach to biological pathway reconstruction/inference for genomes and metagenomes. PLoS Comput Biol 5:e1000465. https://doi.org/10.1371/journal.pcbi.1000465.

68. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B, White O, Kelley ST, Methé B, Schloss PD, Gevers D, Mitreva M, Huttenhower C. 2012. Metabolic reconstruction for metagenomic data and its application to the human microbiome. PLoS Comput Biol 8:e1002358. https://doi.org/10.1371/journal.pcbi.1002358.

69. Suzuki S, Kuenen JG, Schipper K, van der Velde S, Ishii S, Wu A, Sorokin DY, Tenney A, Meng X, Morrill PL, Kamagata Y, Muyzer G, Nealson KH. 2014. Physiological and genomic features of highly alkaliphilic hydrogen-utilizing Betaproteobacteria from a continental serpentinizing site. Nat Commun 5:3900. https://doi.org/10.1038/ncomms4900.

70. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol 428:726–731. https://doi.org/10.1016/j.jmb.2015.11.006.

71. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia JM, Chia JM, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Rashmi J, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569. https://doi.org/10.1038/nbt.4163.

72. Price MN, Dehal PS, Arkin AP. 2010. FastTree2—approximately maximum-likelihood trees for large alignments. PLoS One 5:e9490. https://doi.org/10.1371/journal.pone.0009490.

73. Rabinowitz JD, Kimball E. 2007. Acidic acetonitrile for cellular metabolome extraction from *Escherichia coli*. Anal Chem 79:6167–6173. https://doi.org/10.1021/ac070470c.

74. Salek RM, Steinbeck C, Viant MR, Goodacre R, Dunn WB. 2013. The role of reporting standards for metabolite annotation and identification in metabolomic studies. GigaScience 2:13. https://doi.org/10.1186/2047-217X-2-13.

75. Dittmar T, Koch B, Hertkorn N, 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. https://doi.org/10.4319/lom.2008.6.230.

76. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak MY, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL, Mallick P. 2012. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol 30:918–920. https://doi.org/10.1038/nbt.2377.

77. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. 2006. XCMS:

processing mass spectrometry data for metabolite profiling using non-linear peak alignment, matching, and identification. Anal Chem 78: 779–787. https://doi.org/10.1021/ac051437y.

78. McLean C, Kujawinski EB. 2019. AutoTuner: high fidelity, robust, and rapid parameter selection for metabolomics data processing. bioRxiv https://doi.org/10.1101/812370.

79. Kuhl C, Tautenhahn R, Böttcher C, Larson TR, Neumann S. 2012. CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. Anal Chem 84:283–289. https://doi.org/10.1021/ac202450g.

80. Broadhurst D, Goodacre R, Reinke SN, Kuligowski J, Wilson ID, Lewis MR, Dunn WB. 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. https://doi.org/10.1007/s11306-018-1367-3.

81. Li S, Park Y, Duraisingham S, Strobel FH, Khan N, Soltow QA, Jones DP, Pulendran B. 2013. Predicting network activity from high throughput metabolomics. PLoS Comput Biol 9:e1003123. https://doi.org/10.1371/journal.pcbi.1003123.

82. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TWM, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriot P, Nicholls AW, Reily MD, Thaden JJ, Viant MR. 2007. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 3:211–221. https://doi.org/10.1007/s11306-007-0082-2.

83. Preiner M, Xavier JC, Sousa FL, Zimorski V, Neubeck A, Lang SQ, Greenwell HS, Kleinermanns K, Tüysüz H, McCollom TM, Holm NG, Martin WF. 2018. Serpentinization: connecting geochemistry, ancient metabolism and industrial hydrogenation. Life 8:41. https://doi.org/10.3390/life8040041.