Title

Metabolites Associated with Adaptation of Microorganisms to an Acidophilic, Metal-Rich Environment Identified by Stable-Isotope-Enabled Metabolomics

Journal

mBio, 4(2)

ISSN

2161-2129

Authors

Mosier, Annika C

Justice, Nicholas B

Bowen, Benjamin P

et al.

Publication Date

2013-05-01

DOI

10.1128/mbio.00484-12

Peer reviewed

Metabolites Associated with Adaptation of Microorganisms to an Acidophilic, Metal-Rich Environment Identified by Stable-Isotope-Enabled Metabolomics

Annika C. Mosier,[a] Nicholas B. Justice,[a] Benjamin P. Bowen,[b] Richard Baran,[b] Brian C. Thomas,[a] Trent R. Northen,[b] Jillian F. Banfield[a,c]

Department of Earth and Planetary Science, University of California, Berkeley, California, USA[a]; Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA[b]; Department of Environmental Science, Policy, and Management, University of California, Berkeley, California, USA[c]

A.C.M. and N.B.J. contributed equally to this work.

ABSTRACT **Microorganisms grow under a remarkable range of extreme conditions. Environmental transcriptomic and proteomic studies have highlighted metabolic pathways active in extremophilic communities. However, metabolites directly linked to their physiology are less well defined because metabolomics methods lag behind other *omics* technologies due to a wide range of experimental complexities often associated with the environmental matrix. We identified key metabolites associated with acidophilic and metal-tolerant microorganisms using stable isotope labeling coupled with untargeted, high-resolution mass spectrometry. We observed >3,500 metabolic features in biofilms growing in pH ~0.9 acid mine drainage solutions containing millimolar concentrations of iron, sulfate, zinc, copper, and arsenic. Stable isotope labeling improved chemical formula prediction by >50% for larger metabolites (>250 atomic mass units), many of which were unrepresented in metabolic databases and may represent novel compounds. Taurine and hydroxyectoine were identified and likely provide protection from osmotic stress in the biofilms. Community genomic, transcriptomic, and proteomic data implicate fungi in taurine metabolism. *Leptospirillum* group II bacteria decrease production of ectoine and hydroxyectoine as biofilms mature, suggesting that biofilm structure provides some resistance to high metal and proton concentrations. The combination of taurine, ectoine, and hydroxyectoine may also constitute a sulfur, nitrogen, and carbon currency in the communities.**

IMPORTANCE **Microbial communities are central to many critical global processes and yet remain enigmatic largely due to their complex and distributed metabolic interactions. Metabolomics has the possibility of providing mechanistic insights into the function and ecology of microbial communities. However, our limited knowledge of microbial metabolites, the difficulty of identifying metabolites from complex samples, and the inability to link metabolites directly to community members have proven to be major limitations in developing advances in systems interactions. Here, we show that combining stable-isotope-enabled metabolomics with genomics, transcriptomics, and proteomics can illuminate the ecology of microorganisms at the community scale.**

Received 29 October 2012 **Accepted** 11 February 2013 **Published** 12 March 2013

Citation Mosier AC, Justice NB, Bowen BP, Baran R, Thomas BC, Northen TR, Banfield JF. 2013. Metabolites associated with adaptation of microorganisms to an acidophilic, metal-rich environment identified by stable-isotope-enabled metabolomics. mBio 4(2):00484-12. doi:10.1128/mBio.00484-12.

Editor Mary Ann Moran, University of Georgia

Address correspondence to Trent R. Northen, TRNorthen@lbl.gov, or Jillian F. Banfield, jbanfield@berkeley.edu.

Over the past decade, metagenomic approaches have illuminated the metabolic potential of communities without the need for cultivation and isolation (1, 2). Leveraging the genomic context uncovered by this method, community proteomics and transcriptomics can provide insight into the potential function of coexisting microorganisms *in situ*. However, these analyses are blind to the flux of small-molecule metabolites that are foundational to the physiological or phenotypic state of an organism. Metabolomic measurements can bring to light the key intra- and extracellular metabolites involved in cellular processes such as ion homeostasis, redox status, nutrient cycling, energetics, and cell-cell signaling (e.g., see references 3 and 4). By capturing relative sizes of the metabolite pools, metabolomics is a reflection of the net expression of many genes, pathways, and processes.

Metabolomics studies may prove particularly useful in studying adaptation to extreme environments, since metabolites essential to survival in these environments may be relatively abundant (e.g., for ion homeostasis). Yet, studies of metabolites have typically targeted specific molecules of interest from isolated organisms. Few studies have leveraged untargeted, high-throughput metabolomics techniques to study microbial adaptation (e.g., adaptation of *Pyrococcus furiosus* to temperature stress [5] and *Streptomyces coelicolor* to salt stress [6]).

Traditional physiological studies of microbial isolates have defined many adaptation mechanisms to extreme conditions, including improved membrane selectivity and stability, detoxification, enhancement of cellular repair capabilities, and alteration of macromolecular structures. However, adaptation of organisms to their environments in part relates to behavior within a community context—a facet that is not captured in typical laboratory-

based studies of isolates but is captured by community metabolomics studies. The abundance of some compounds (e.g., sugars and amino acids) results from uptake, consumption, and excretion by many different organisms; thus, the overall concentration reflects the net metabolic state of the community. Untargeted metabolomics has tremendous potential for hypothesis generation in microbial communities since it provides a direct biochemical observation of the community metabolism. However, only a few studies (e.g., see reference 7) have used untargeted metabolomics to study adaptation to environmental challenges in a community context. This is in part because metabolomics methods are still challenging (relative to the more standardized *omics* approaches such as transcriptomics) due to a wide range of experimental complexities often associated with the environmental matrix (e.g., abundant salt can result in extensive experimental artifacts).

To investigate community-level adaptations to the simultaneous challenges of high proton and metal concentrations, we examined the metabolome of microbial biofilm communities in an acid mine drainage (AMD) environment (Richmond Mine, Iron Mountain, CA). Previous research in the system demonstrated that proteins and metabolite features exhibited correlative patterns reflective of functional differentiation of bacterial species (8), suggesting that combining *omics* approaches may prove useful in defining adaptation strategies specific to particular groups of organisms. Microbial biofilms found at Richmond Mine grow at low pH (typically 0.5 to 1.2) and elevated temperature (30 to 56°C) in solutions containing millimolar concentrations of sulfate, iron, zinc, copper, and arsenic (9)—conditions that together make untargeted metabolomics extremely challenging. Here, we used a combination of stable isotope probing and untargeted metabolomics to facilitate the identification of metabolites from *in situ* and laboratory-cultivated AMD microbial biofilms. We integrated metabolite identifications with previously acquired genomic and proteomic data to elucidate adaptation to acidophilic and metal-rich conditions on the metabolic level.

RESULTS AND DISCUSSION

Community compositions of AMD biofilms with different growth strategies. For metabolomic analyses presented here, biofilm samples were collected from the air-AMD solution interface of the AB-muck site within the Richmond Mine on 15 July 2011 (here referred to as the mine biofilm sample). The mine biofilm was categorized as late developmental stage, based primarily on biofilm thickness. AMD biofilms were also grown in laboratory bioreactors using the mine biofilm as inoculum for three different lengths of cultivation time: 26, 35, and 51 days (here identified as BR-26days, BR-35days, and BR-51days, indicating bioreactor growth and length of cultivation). In all cases, the cultivated biofilms were thick and well developed at the time of sampling. BR-26days and BR-35days were grown with ^{15}N-labeled ammonium sulfate (resulting in ^{15}N-labeled nitrogen atoms in the biofilm metabolites) in order to identify the number of nitrogen atoms in each metabolite and also to confirm biological origin.

The community composition of these biofilm samples was determined using fluorescence *in situ* hybridization (FISH) targeting identification of broad phylogenetic groups as well as individual species and strains (see Fig. S1 in the supplemental material). Insufficient biomass precluded FISH analyses on BR-35days. Prior studies showed that biofilms in the Richmond Mine are dominated by *Leptospirillum* group II, a chemoautotrophic iron-oxidizing bacterium (1, 10). Genomic reconstructions revealed two distinct strains of *Leptospirillum* group II, referred to as the 5-way and UBA genotypes (11, 12). In this study, the *Leptospirillum* group II 5-way genotype was more abundant than the UBA genotype in both bioreactors. Conversely, nearly all of the bacteria in the mine biofilm belonged to the *Leptospirillum* group II UBA genotype, which is consistent with prior studies showing that the UBA genotype typically predominates over the 5-way genotype in the mine bacteria (13).

Previous studies have shown that as the biofilms mature and thicken, they also diversify, with increasing proportions of *Leptospirillum* group III bacteria and *Archaea*, as well as other low-abundance taxa from the *Eukarya*, *Firmicutes*, and *Actinobacteria* lineages (1, 14). The bioreactor biofilms had a much higher percentage of *Leptospirillum* group III bacteria than that seen in the mine biofilm (22 to 32% compared to only 2% in the mine). Geochemical and FISH data collected previously in the mine suggest a positive correlation between the abundance of *Leptospirillum* group III and ammonium concentrations ($r = 0.96$, $n = 6$), which may suggest that ammonium concentrations in the bioreactors (2 mM compared to an average concentration of 177 μM in the mine) favor *Leptospirillum* group III. Recent studies also indicate that high ferric iron concentrations in the bioreactors relative to mine solutions may also select for *Leptospirillum* group III (S. Ma, S.E. Spaulding, B.C. Thomas, J.F. Banfield, submitted for publication).

The abundance of *Archaea* varied between the mine and bioreactor biofilms. *Archaea* made up 29 to 35% of the communities in the BR-26days and mine biofilms. Conversely, BR-51days had much higher percentages of *Archaea* (61%), which may be the result of longer cultivation time than that of BR-26days. The low-abundance community member *Sulfobacillus* was found in all samples (0.2 to 1.2%). Eukaryotes were not identified by FISH but are not necessarily absent from the biofilms; the inherent heterogeneity of the biofilms and patchy eukaryal distribution may prohibit visualization in microscopy with a random sampling design.

Detection and chemical formula prediction of metabolites found in AMD enabled by stable isotope labeling. Metabolites were extracted from natural and cultivated AMD biofilms (mine, BR-26days, BR-35days, and BR-51days) and analyzed using liquid chromatography coupled to high-resolution mass spectrometry (LC-MS). Both hydrophilic interaction liquid chromatography (HILIC) and C_{18} reverse-phase (RP) columns were used to separate polar and nonpolar organics, respectively. LC-ESI-MS (LC-MS with electrospray ionization) is one of the most widely used metabolomic platforms given its versatility in separation techniques coupled with the wide range of compounds that can be desorbed/ionized (15). Generally, LC-MS approaches fall into one of two categories: (i) targeted analyses, which aim to quantify changes within a defined set of known metabolites diagnostic of a given phenotype of interest, and (ii) untargeted analyses, which seek to discern both novel and previously characterized metabolites (16, 17). An untargeted LC-MS approach was used in the current study, rather than the commonly used gas chromatography-mass spectrometry (GC-MS) approach, to maximize the diversity of metabolites detected.

More than 3,500 raw features (ions with unique retention time and mass-to-charge ratio combinations) were identified in each of the RP and HILIC data sets. These uncurated data, however, included features associated with background chromatography

TABLE 1 Comparison of the number of features and metabolites resulting from various levels of curation

Analytical column	No. of features after three-way visualization analysis	No. of metabolites		
		After manual curation	Containing nitrogen	With chemical formula
Agilent Zorbax SB-C_{18} column	191	56	27	38
Acquity UPLC BEH HILIC column	50	24	19	18

noise and features also found in extraction blanks, as well as multiple adducts and fragment ions of the same compounds. In order to obtain the highest-quality data possible, we used a strict manual curation strategy to identify compounds of interest in our samples and prevent errors associated with automated peak detection, deconvolution, and alignment. Three-way visualization plots (18) of the ^{14}N-biofilm, ^{15}N-biofilm, and extraction blank data sets (available at http://geomicrobiology.berkeley.edu/pages/metabolites.html) were used to narrow down the large list of raw features to "pure spectra," that is, the spectra that would likely result from a pure compound within the biological matrix. From this visualization, as well as manual observation of very abundant peaks, we generated a list of 241 likely parent ion features (the highest-intensity feature from a pure spectrum) from the RP and HILIC analyses (Table 1).

Eighty parent compounds (56 in RP and 24 in HILIC) (Table 1) were confirmed after grouping coeluting features and identifying common adducts, fragment ions, and neutral losses, including those associated with esterification of carboxylic acids (presumably as a result of extraction buffer and residual acidic AMD). The number of metabolites found here is consistent with those in other untargeted LC-MS metabolomic studies from complex, natural biological matrices (19–22). Forty-eight percent of the reverse-phase (RP) metabolites and 79% of the HILIC metabolites contained at least one nitrogen atom (determined by stable isotope labeling), providing a strong level of certainty that the compounds are of biological origin. Mass spectra of all compounds were manually visualized to ensure absence in extraction blanks.

The use of stable isotopes greatly aids in determining the chemical formulae of unknown metabolites by constraining the possible elemental composition (19, 23, 24). Here, using the number of nitrogen atoms informed by stable isotope labeling, chemical formulae were determined for 38 of the RP features and 18 of the HILIC features (see Table S1 in the supplemental material). The nitrogen labeling method also informs the formulae of compounds without nitrogen, as their chemical formulae are constrained by the "zero" nitrogen count. Generally, high-mass-accuracy (<5-ppm) mass spectrometers can confidently assign unique chemical formulae for features under ~200 to 250 atomic mass units (amu) (23, 25). Indeed, we were able to confidently assign chemical formulae for most features of <250 amu without the assistance of nitrogen labeling. For those 37 features above 250 amu in Table S1, nitrogen labeling allowed us to confidently assign chemical formulae to 20 of them, twice as many formulae as were possible with spectral information alone. Features for which chemical formulae could not be determined either were of high *m/z* or low spectral quality or had an unidentified adduct.

Metabolite annotation and identification. We obtained MS/MS spectra at collision energies of 10, 20, and 40 eV on all features with sufficient peak heights (available at http://geomicrobiology.berkeley.edu/pages/metabolites.html). The chemical formulae and MS/MS data were matched with metabolites in online databases (MetaCyc, KEGG, MassBank, and METLIN) (see Table S1 in the supplemental material). MS/MS data from more than 90% of the features had no match to metabolites in MS/MS databases (MassBank and METLIN), and as per protocols established by the Chemical Analysis Working Group of the Metabolomics Standards Initiative, these metabolites are classified as "unknown" (26, 27). While these data may speak to the novelty of some AMD metabolites, it must also be noted that these databases rely on commercially available standards, which are estimated to represent only half of all biological metabolites (15).

From this analysis, we found three metabolites that were particularly interesting and warranted further investigation within the context of the biofilm community: (i) phosphatidylethanolamine (PE) lipids, (ii) taurine, and (iii) hydroxyectoine. Unusual *lyso* phosphatidylethanolamine lipids and methylated derivatives previously identified in the Richmond Mine (28) were also found in the mine and bioreactor biofilms presented here. Fischer et al. (28) suggested a link between these lipids and the *Leptospirillum* group II UBA genotype based on correlations of lipid and proteome abundance patterns. Interestingly, we found features consistent with some of these same lipids (same *m/z* and retention times for 454.294 and 480.309) in pure cultures of *Acidomyces richmondensis*, a fungus known to be abundant in the mine and often the dominant eukaryal species (29, 30). The *A. richmondensis* genome contained 8 genes predicted to be involved in the synthesis, methylation, and binding of PEs, some of which were expressed in community transcriptomic and proteomic data (C.S. Miller, A.C. Mosier, S.W. Singer, C. Pan, B.C. Thomas, J.F. Banfield, unpublished data). Together, these results show that fungi and bacteria may both be involved in the metabolism of these lipids. Fischer et al. (28) suggested that these lipids may prevent uptake of toxic levels of iron cations in AMD biofilms.

Taurine: metabolic interactions and abundance within the AMD community. Among the metabolites present in the AMD biofilms, we identified taurine (2-aminoethanesulfonic acid) by comparing MS/MS spectra with a taurine chemical standard (see Fig. S2 in the supplemental material). Taurine is a phylogenetically ancient compound (31) and is involved in numerous physiological functions across disparate forms of life, including membrane stabilization, stimulation of glycolysis and glycogenesis, regulation of phosphorylation, and antioxidation (references 31 and 32 and references therein). Some microbes can use taurine as an exclusive source of carbon, nitrogen and sulfur (references 33 and 34 and references therein). Taurine is a particularly effective osmoregulator and is used as a compatible solute by a variety of microorganisms (e.g., see references 31 and 35 and references therein). Compatible solutes (generally very soluble, low-

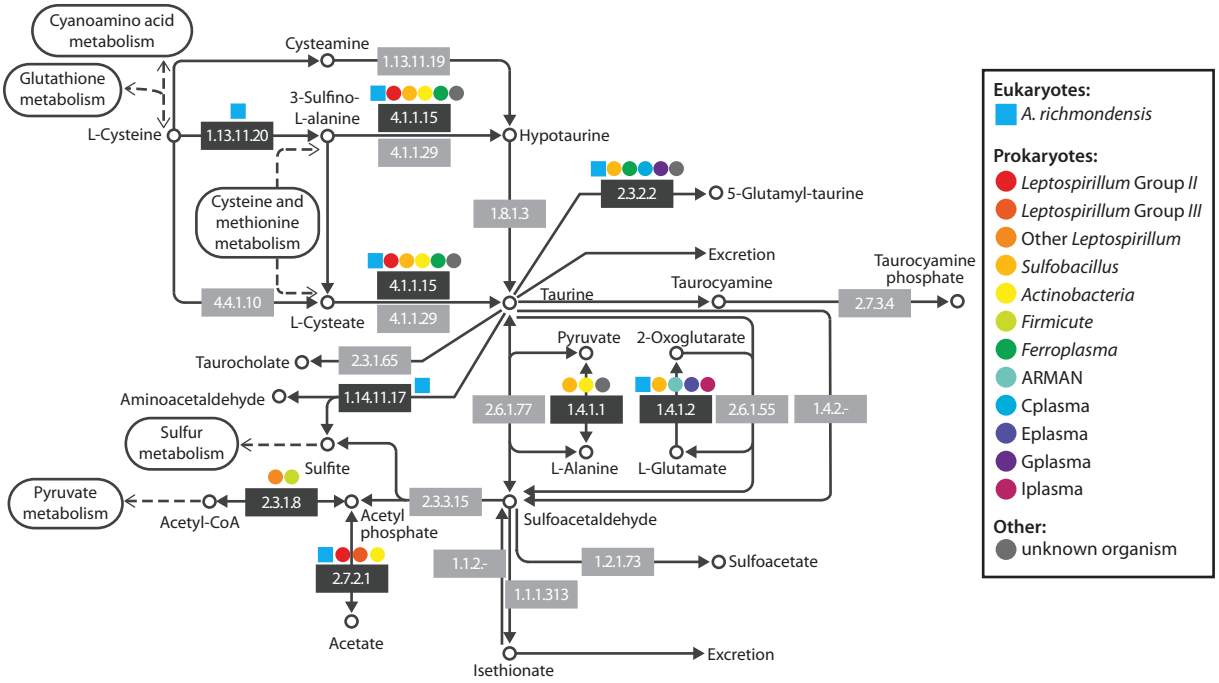

FIG 1 Taurine metabolism by prokaryotes and eukaryotes in the AMD biofilms.

molecular-weight organic molecules) can be accumulated in the cytoplasm as a mechanism for coping with hyperosmotic stress. Compatible solutes can also protect proteins, nucleic acids, and membranes from the harmful effects of heat, freezing, drying, and oxygen radicals (references 36 to 38 and references therein).

Although many of the potential roles for taurine are relevant, its properties as a compatible solute may be particularly useful in microbial adaptation to the high-ionic-strength waters within the Richmond Mine (9). We explored genomic sequences of ~20 AMD biofilm community members (including bacteria, archaea, and fungi) to determine which organisms may produce taurine (Fig. 1). This data set contains nearly 80,000 gene sequences and captures essentially all organisms representing more than a few percent of the community. Reciprocal BLAST searches against the KEGG database indicated that archaea and bacteria in the AMD communities are unable to generate taurine (no archaea or bacteria are known to synthesize taurine). The only archaeal or bacterial enzyme potentially involved in taurine biosynthesis was a glutamate decarboxylase (EC 4.1.1.15), which has broad functionality in several different metabolic pathways.

We evaluated the likelihood for eukaryotic taurine biosynthesis using the genome of the dominant fungus, *A. richmondensis* (Fig. 1). The *A. richmondensis* genome contains cysteine dioxygenase (EC 1.13.11.20) and glutamate decarboxylase (EC 4.1.1.15) genes involved in two routes of taurine metabolism. Enzymes mediating the oxidation of hypotaurine to taurine and 3-sulfino-L-alanine to L-cysteate were not evident; however, it has been shown that these reactions may occur nonenzymatically (references 39 and 40 and references therein). Ferric iron, found in high concentrations in the mine, may act as a chemical oxidant of these compounds, thereby completing the pathways of taurine production in *A. richmondensis*. Efforts to identify taurine in a pure culture of *A. richmondensis* failed, although culture conditions may not have favored taurine production.

We also assessed the potential for taurine degradation using community genomic sequence data (Fig. 1). While bacteria and archaea in the mine carry genes for some enzymes involved in taurine degradation pathways, they do not appear to have a definitive or complete mechanism for the breakdown of taurine. The most likely candidate route of degradation is via gammaglutamyltranspeptidase (EC 2.3.2.2) found in *Sulfobacillus* and three archaeal species (*Ferroplasma*, Cplasma, and Gplasma); however, this enzyme has broad activity and is also involved in cyanoamino acid, glutathione, and arachidonic acid metabolism. *Sulfobacillus* also has genes encoding two taurine transport proteins (TauA and TauC), suggesting that taurine may indeed have a biological role in this organism. Taurine diffuses slowly through cell membranes, and taurine biotransformation enzymes are usually soluble and intracellular, so transport of taurine into the bacterial cell is required for utilization of the compound (31, 41). Cells responding to hyperosmolar conditions can increase intracellular taurine content via active transport of taurine into the cell.

Genomic evidence suggests that the fungal species *A. richmondensis* is capable of degrading taurine via taurine catabolism dioxygenase TauD/TfdA enzymes. TauD is a dioxygenase that converts taurine to sulfite and aminoacetaldehyde, with reaction requirements of oxygen, Fe^{2+}, and α-ketoglutarate (42). In *Escherichia coli*, the *tauD* gene is expressed only under conditions of sulfate starvation (42, 43). There are 12 copies of the taurine catabolism dioxygenase *tauD/tfdA* genes in the *A. richmondensis* genome. Interestingly, transcripts of all 12 *tauD/tfdA* genes were detected in a fungal streamer biofilm community from the mine (Miller et al., unpublished). Two of these transcripts were relatively abundant in the transcriptome (ranked in the top 1,500 transcripts out of 10,305 total genes), and their proteins were also detected in a community metaproteome (Miller et al., unpublished). Other Dothidiomycetes (the fungal class including *A. richmondensis*) genomes contain between 1 and 10 copies of

taurine catabolism dioxygenase genes per genome (based on BLAST searches).

Given the genomic potential for taurine biosynthesis and likelihood for degradation in the AMD biofilms, we evaluated the abundance of taurine across the different growth conditions (natural mine biofilms and biofilms grown in bioreactors for 26, 35, and 51 days) based on peak heights in the MS spectra (see Fig. S3 in the supplemental material). Taurine concentrations were an order of magnitude higher in all three bioreactor biofilms than in the mine biofilm. This discrepancy is likely explained by different biogeochemical conditions in the mine and bioreactor biofilms. It is possible that the bioreactor communities generate more taurine relative to the mine communities or, equally possible, that more taurine is consumed in the mine. Metagenomic evidence implicates *A. richmondensis* as the dominant organism involved in taurine biosynthesis and degradation. The low-abundance community member *Sulfobacillus* may also have the potential for taurine consumption; while FISH shows higher numbers of *Sulfobacillus* in the mine biofilms, the percentage of the total community is only on the order of 1%.

Hydroxyectoine: metabolic interactions within the AMD community. We identified hydroxyectoine (confirmed with chemical standards and MS/MS spectra; see Fig. S4 in the supplemental material) and possibly ectoine (correct mass and chemical formula but incomplete MS/MS data) in natural and cultivated AMD biofilms and were interested in their role as compatible solutes used in adaptation to hyperosmotic stress. Ectoine biosynthesis occurs in three enzymatic steps: (i) L-diaminobutyric acid transaminase (EctB or ThpB) converts L-aspartate-beta-semialdehyde into L-diaminobutyric acid, (ii) acetylation to *N*-γ-acetyldiaminobutyric acid occurs via L-diaminobutyric acid acetyltransferase (EctA or ThpA), and (iii) cyclic condensation then leads to the formation of ectoine through ectoine synthase (EctC or ThpC) (44–46). Hydroxyectoine is primarily generated through the hydroxylation of ectoine by ectoine hydroxylase (EctD or ThpD) (47, 48). Compared to ectoine, hydroxyectoine can confer additional protective properties against heat stress (47) and freeze-drying (49).

Complete ectoine and hydroxyectoine biosynthesis pathways have been identified previously in one archaeal genome, *Nitrosopumilus maritimus* (50), and in over 50 bacterial genomes with particular representation among the alpha- and gammaproteobacteria and *Actinobacteria* (reference 38 and references therein). In the AMD biofilms, both the 5-way and UBA genotypes of *Leptospirillum* group II have all of the genes necessary for ectoine and hydroxyectoine biosynthesis (*ectABCD*), and their protein products have been identified by community proteomics (51). Previous studies showed high numbers of ectoine synthase proteins in *Leptospirillum* group II grown under nonoptimized culture conditions; however, ectoine synthases were found in levels similar to those in the natural biofilm upon optimization of the culturing media (52). *Leptospirillum* group II EctB proteins were significantly more abundant in low-pH cultures (pH 0.85 versus pH 1.45), suggesting greater osmotic stress during growth in more acidic solutions (53). Interestingly, *Leptospirillum* group II was the first acidophilic bacterium described with a complete pathway for biosynthesis of ectoine and hydroxyectoine (51). Other AMD community members also have genes in the ectoine biosynthesis pathway. *ectB* genes were found in some archaea (Cplasma, Eplasma, and *Ferroplasma*), and *Sulfobacillus* has both *ectA* and *ectB* genes; however, ectoine biosynthesis by these organisms cannot be confirmed since the complete Ect operon was not found.

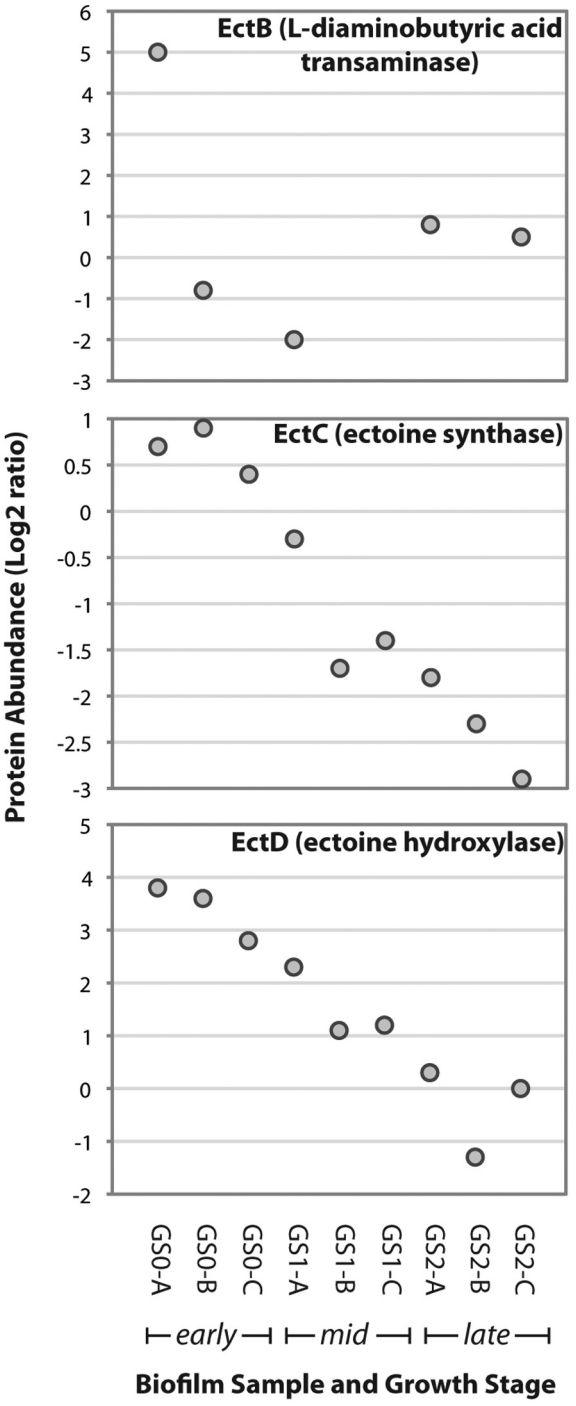

FIG 2 Abundance of ectoine and hydroxyectoine biosynthesis proteins across different growth stages of AMD biofilms. Data are based on previously acquired quantitative proteomics data (54).

We evaluated the abundance of ectoine and hydroxyectoine biosynthesis proteins in *Leptospirillum* group II bacteria across different growth stages of AMD biofilms (early, mid-, and late growth stages) (Fig. 2), using previously acquired quantitative

proteomics data (54). EctA proteins were not found in any of the samples, which may suggest low abundance. Protein abundance of EctB, EctC, and EctD generally decreased with biofilm development, suggesting that more ectoine and hydroxyectoine are produced in the early growth stages. In early biofilm development, the organisms may have greater exposure to the AMD solution because the biofilm is still thin and friable and contains less extracellular polymeric substance (EPS) (55). EPS has been reported to provide protection from a variety of environmental stresses, including osmotic shock (e.g., see reference 56 and references therein). Thus, with less protection in early biofilm development, *Leptospirillum* group II bacteria may produce more ectoine and hydroxyectoine in order to cope with greater exposure to the high-ionic-strength AMD solution.

Some organisms are capable of concurrently using ectoine as an osmoprotectant and as an energy and carbon substrate (57–59). Genes involved in ectoine metabolism have been identified in *Sinorhizobium meliloti* (*eutABCD*) (57) and *Halomonas elongata* (*doeABCD*) (58). In the AMD biofilms, some of these genes involved in ectoine utilization have been identified in *Sulfobacillus* and *Firmicutes* genomes, but not the complete operons.

Conclusion. We used an untargeted approach based on stable isotope labeling coupled with high-resolution mass spectrometry to uncover metabolites within natural and cultivated AMD microbial communities. We confirmed the identification of taurine and hydroxyectoine and used community proteogenomic data to determine which organisms are capable of producing or consuming these osmolytes. From this, we suggest that to mitigate less protection from EPS in early biofilm development, *Leptospirillum* group II bacteria may produce more ectoine and hydroxyectoine in order to cope with greater exposure to the high-ionic-strength AMD solution. We determined the specific chemical formulae of many other metabolites that were not identified because they are not present in current metabolic databases and suitable pure compound standards for a defined set of candidate molecules were not available. The set of abundant but not identified compounds could include novel metabolites, with potentially high biological and chemical relevance.

MATERIALS AND METHODS

Sample collection. AMD biofilms from Iron Mountain Mine (near Redding, CA) were flash-frozen on site in a dry ice-ethanol bath and then transferred to −80°C upon return to the laboratory. Laboratory-grown biofilms were cultured in bioreactors as previously described (52). In addition to the traditional media, AMD biofilms were also grown with [^{15}N]ammonium sulfate. The flow rate of the bioreactors was approximately 225 μl per min. Biofilms were harvested after ~4 to 7 weeks of growth and frozen at −80°C for use in metabolite analysis.

FISH. Fluorescence *in situ* hybridization (FISH) was carried out on fixed (4% paraformaldehyde) AMD biofilm samples as described previously (60, 61). For estimation of abundance, cells were counted in three replicate fields of view (average of 888 total cells counted per probe) and converted to a percentage of the total cell count found using the general nucleic acid stain 4′,6-diamidino-2-phenylindole (DAPI). A complete list of probes is available in Text S1 in the supplemental material.

Metabolite extraction. For metabolomics, three replicates with approximately 75 mg (wet weight) of biomass were used from each sample (natural biofilm, cultured biofilm, and ^{15}N-labeled cultured biofilm) and were extracted in 700 μl of methanol-isopropanol-water (3:3:2 [vol/vol/vol]). The full extraction methodology is described in Text S1 in the supplemental material.

Liquid chromatography-mass spectrometry analysis. Samples were analyzed using an Agilent 1200 capillary LC system with an Agilent 6520 dual-ESI-quadrupole time of flight (Q-TOF) mass spectrometer (±2-ppm MS mass accuracy and mass resolution of 20,000). Full descriptions of chromatography, MS/MS, and data analysis can be found in Text S1 in the supplemental material.

Community genomic analysis. Bacterial, archaeal, and eukaryal genes involved in taurine and ectoine metabolism were identified from a well-curated metagenomic and genomic data set containing nearly 80,000 gene sequences and capturing essentially all organisms representing more than a few percent of the community. MycoCosm (62) was used to evaluate gene annotations for the *Acidomyces richmondensis* fungal genome. Reciprocal BLAST searches against the KEGG database were conducted using the KAAS server (63). Transcript and protein abundances of *Acidomyces richmondensis* were determined by Miller et al. (unpublished). Abundances of proteins involved in ectoine and hydroxyectoine synthesis were determined by Mueller et al. (54).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00484-12/-/DCSupplemental.

Text S1, DOCX file, 0.1 MB.
Figure S1, TIF file, 4.6 MB.
Figure S2, TIF file, 3.4 MB.
Figure S3, TIF file, 2.5 MB.
Figure S4, TIF file, 4.5 MB.
Table S1, DOCX file, 0.1 MB.

ACKNOWLEDGMENTS

T. W. Arman (President, Iron Mountain Mines) and R. Sugarek (Environmental Protection Agency) are thanked for providing site access to the Richmond Mine, and R. Carver and M. Jones are thanked for onsite assistance.

Funding was provided by grants from the U.S. Department of Energy, DE-FG02-10ER64996, DE-FG02-10ER65028, and DE-FOA-0000143.

REFERENCES

1. **Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF.** 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature **428**:37–43.
2. **Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall AM, Mead D, Azam F, Rohwer F.** 2002. Genomic analysis of uncultured marine viral communities. Proc. Natl. Acad. Sci. U. S. A. **99**:14250–14255.
3. **Wasser JS, Lawler RG, Jackson DC.** 1996. Nuclear magnetic resonance spectroscopy and its applications in comparative physiology. Physiol. Zool. **69**:1–34.
4. **Kristal BS, Vigneau-Callahan KE, Matson WR.** 1998. Simultaneous analysis of the majority of low-molecular-weight, redox-active compounds from mitochondria. Anal. Biochem. **263**:18–25.
5. **Trauger SA, Kalisak E, Kalisiak J, Morita H, Weinberg MV, Menon AL, Poole FL, Adams MW, Siuzdak G.** 2008. Correlating the transcriptome, proteome, and metabolome in the environmental adaptation of a hyperthermophile. J. Proteome Res. **7**:1027–1035.
6. **Kol S, Merlo ME, Scheltema RA, de Vries M, Vonk RJ, Kikkert NA, Dijkhuizen L, Breitling R, Takano E.** 2010. Metabolomic characterization of the salt stress response in Streptomyces coelicolor. Appl. Environ. Microbiol. **76**:2574–2581.
7. **Halter D, Goulhen-Chollet F, Gallien S, Casiot C, Hamelin J, Gilard F, Heintz D, Schaeffer C, Carapito C, Van Dorsselaer A, Tcherkez G, Arsène-Ploetze F, Bertin PN.** 2012. In situ proteo-metabolomics reveals metabolite secretion by the acid mine drainage bio-indicator, euglena mutabilis. ISME J. **6**:1391–1402.
8. **Wilmes P, Bowen BP, Thomas BC, Mueller RS, Denef VJ, VerBerkmoes NC, Hettich RL, Northen TR, Banfield JF.** 2010. Metabolome-proteome differentiation coupled to microbial divergence. mBio **1**(5):e00246-10. http://dx.doi.org/10.1128/mBio.00246-10.
9. **Druschel G, Baker B, Gihring T, Banfield J.** 2004. Acid mine drainage biogeochemistry at Iron Mountain, California. Geochem. Trans. **5**:13–32.

10. **Bond PL, Druschel GK, Banfield JF.** 2000. Comparison of acid mine drainage microbial communities in physically and geochemically distinct ecosystems. Appl. Environ. Microbiol. **66**:4962–4971.
11. **Lo I, Denef VJ, Verberkmoes NC, Shah MB, Goltsman D, DiBartolo G, Tyson GW, Allen EE, Ram RJ, Detter JC, Richardson P, Thelen MP, Hettich RL, Banfield JF.** 2007. Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria. Nature **446**: 537–541.
12. **Simmons SL, Dibartolo G, Denef VJ, Goltsman DS, Thelen MP, Banfield JF.** 2008. Population genomic analysis of strain variation in Leptospirillum group II bacteria involved in acid mine drainage formation. PLoS Biol. **6**:e177. http://dx.doi.org/10.1371/journal.pbio.0060177.
13. **Denef VJ, Kalnejais LH, Mueller RS, Wilmes P, Baker BJ, Thomas BC, VerBerkmoes NC, Hettich RL, Banfield JF.** 2010. Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. Proc. Natl. Acad. Sci. U. S. A. **107**:2383–2390.
14. **Dick GJ, Andersson AF, Baker BJ, Simmons SL, Thomas BC, Yelton AP, Banfield JF.** 2009. Community-wide analysis of microbial genome sequence signatures. Genome Biol. **10**:R85. http://dx.doi.org/10.1186/gb-2009-10-8-r85.
15. **Garcia DE, Baidoo EE, Benke PI, Pingitore F, Tang YJ, Villa S, Keasling JD.** 2008. Separation and mass spectrometry in microbial metabolomics. Curr. Opin. Microbiol. **11**:233–239.
16. **Oldiges M, Lütz S, Pflug S, Schroer K, Stein N, Wiendahl C.** 2007. Metabolomics: current state and evolving methodologies and tools. Appl. Microbiol. Biotechnol. **76**:495–511.
17. **Fiehn O.** 2001. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. Comp. Funct. Genomics **2**:155–168.
18. **Baran R, Robert M, Suematsu M, Soga T, Tomita M.** 2007. Visualization of three-way comparisons of omics data. BMC Bioinformatics **8**:72. http://dx.doi.org/10.1186/1471-2105-8-72.
19. **Baran R, Bowen BP, Bouskill NJ, Brodie EL, Yannone SM, Northen TR.** 2010. Metabolite identification in Synechococcus sp. PCC 7002 using untargeted stable isotope assisted metabolite profiling. Anal. Chem. **82**: 9034–9042.
20. **Baran R, Bowen BP, Northen TR.** 2011. Untargeted metabolic footprinting reveals a surprising breadth of metabolite uptake and release by Synechococcus sp. PCC 7002. Mol. Biosyst. **7**:3200–3206.
21. **van der Werf MJ, Overkamp KM, Muilwijk B, Coulier L, Hankemeier T.** 2007. Microbial metabolomics: toward a platform with full metabolome coverage. Anal. Biochem. **370**:17–25.
22. **Brauer MJ, Yuan J, Bennett BD, Lu W, Kimball E, Botstein D, Rabinowitz JD.** 2006. Conservation of the metabolomic response to starvation across two divergent microbes. Proc. Natl. Acad. Sci. U. S. A. **103**: 19302–19307.
23. **Hegeman AD, Schulte CF, Cui Q, Lewis IA, Huttlin EL, Eghbalnia H, Harms AC, Ulrich EL, Markley JL, Sussman MR.** 2007. Stable isotope assisted assignment of elemental compositions for metabolomics. Anal. Chem. **79**:6912–6921.
24. **Rodgers RP, Blumer EN, Hendrickson CL, Marshall AG.** 2000. Stable isotope incorporation triples the upper mass limit for determination of elemental composition by accurate mass measurement. J. Am. Soc. Mass Spectrom. **11**:835–840.
25. **Bowen BP, Fischer CR, Baran R, Banfield JF, Northen T.** 2011. Improved genome annotation through untargeted detection of pathway-specific metabolites. BMC Genomics **12**(Suppl 1):S6. http://dx.doi.org/10.1186/1471-2164-12-S3-S6.
26. **Dunn WB, Erban A, Weber RJM, Creek DJ, Brown M, Breitling R, Hankemeier T, Goodacre R, Neumann S, Kopka J, Viant MR.** 2012. Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics. Metabolomics **8**:1–23.
27. **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, Marriott P, Nicholls AW, Reily MD, Thaden JJ, Viant MR.** 2007. Proposed minimum reporting standards for chemical analysis. Metabolomics **3**:211–221.
28. **Fischer CR, Wilmes P, Bowen BP, Northen TR, Banfield JF.** 2011. Deuterium-exchange metabolomics identifies N-methyl lyso phosphatidylethanolamines as abundant lipids in acidophilic mixed microbial communities. Metabolomics **8**:566–578.
29. **Baker BJ, Lutz MA, Dawson SC, Bond PL, Banfield JF.** 2004. Metabolically active eukaryotic communities in extremely acidic mine drainage. Appl. Environ. Microbiol. **70**:6264–6271.
30. **Baker BJ, Tyson GW, Goosherst L, Banfield JF.** 2009. Insights into the diversity of eukaryotes in acid mine drainage biofilm communities. Appl. Environ. Microbiol. **75**:2192–2199.
31. **Huxtable RJ.** 1992. Physiological actions of taurine. Physiol. Rev. **72**: 101–163.
32. **Schuller-Levis GB, Park E.** 2003. Taurine: new implications for an old amino acid. FEMS Microbiol. Lett. **226**:195–202.
33. **Stapley EO, Starkey RL.** 1970. Decomposition of cysteic acid and taurine by soil micro-organisms. J. Gen. Microbiol. **64**:-77–84.
34. **Denger K, Smits TH, Cook AM.** 2006. Genome-enabled analysis of the utilization of taurine as sole source of carbon or of nitrogen by Rhodobacter sphaeroides 2.4.1. Microbiology **152**:3197–3206.
35. **Cayley S, Lewis BA, Guttman HJ, Record MT.** 1991. Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity. Implications for protein-DNA interactions *in vivo*. J. Mol. Biol. **222**: 281–300.
36. **da Costa MS, Santos H, Galinski EA.** 1998. An overview of the role and diversity of compatible solutes in Bacteria and Archaea. Adv. Biochem. Eng. Biotechnol. **61**:117–153.
37. **Lentzen G, Schwarz T.** 2006. Extremolytes: natural compounds from extremophiles for versatile applications. Appl. Microbiol. Biotechnol. **72**: 623–634.
38. **Pastor JM, Salvador M, Argandoña M, Bernal V, Reina-Bueno M, Csonka LN, Iborra JL, Vargas C, Nieto JJ, Cánovas M.** 2010. Ectoines in cell stress protection: uses and biotechnological production. Biotechnol. Adv. **28**:782–801.
39. **Fontana M, Duprè S, Pecci L.** 2006. The reactivity of hypotaurine and cysteine sulfinic acid with peroxynitrite. Adv. Exp. Med. Biol. **583**:15–24.
40. **Coloso RM, Hirschberger LL, Dominy JE, Lee J-I, Stipanuk MH.** 2006. Cysteamine dioxygenase: evidence for the physiological conversion of cysteamine to hypotaurine in rat and mouse tissues. Adv. Exp. Med. Biol. **583**:25–36.
41. **Cook AM, Denger K.** 2006. Metabolism of taurine in microorganisms: a primer in molecular biodiversity? Adv. Exp. Med. Biol. **583**:3–13.
42. **Eichhorn E, van der Ploeg JR, Kertesz MA, Leisinger T.** 1997. Characterization of alpha-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. J. Biol. Chem. **272**:23031–23036.
43. **van der Ploeg JR, Weiss MA, Saller E, Nashimoto H, Saito N, Kertesz MA, Leisinger T.** 1996. Identification of sulfate starvation-regulated genes in Escherichia coli: a gene cluster involved in the utilization of taurine as a sulfur source. J. Bacteriol. **178**:5438–5446.
44. **Caspi R, Altman T, Dale JM, Dreher K, Fulcher CA, Gilham F, Kaipa P, Karthikeyan AS, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Paley S, Popescu L, Pujar A, Shearer AG, Zhang P, Karp PD.** 2010. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. **38**:D473–D479. http://dx.doi.org/10.1093/nar/gkp875.
45. **Louis P, Galinski EA.** 1997. Characterization of genes for the biosynthesis of the compatible solute ectoine from Marinococcus halophilus and osmoregulated expression in Escherichia coli. Microbiology **143**:1141–1149.
46. **Cánovas D, Vargas C, Iglesias-Guerra F, Csonka LN, Rhodes D, Ventosa A, Nieto JJ.** 1997. Isolation and characterization of salt-sensitive mutants of the moderate halophile Halomonas elongata and cloning of the ectoine synthesis genes. J. Biol. Chem. **272**:25794–25801.
47. **García-Estepa R, Argandoña M, Reina-Bueno M, Capote N, Iglesias-Guerra F, Nieto JJ, Vargas C.** 2006. The ectD gene, which is involved in the synthesis of the compatible solute hydroxyectoine, is essential for thermoprotection of the halophilic bacterium *Chromohalobacter salexigens*. J. Bacteriol. **188**:3774–3784.
48. **Prabhu J, Schauwecker F, Grammel N, Keller U, Bernhard M.** 2004. Functional expression of the ectoine hydroxylase gene (thpD) from Streptomyces chrysomallus in Halomonas elongata. Appl. Environ. Microbiol. **70**:3130–3132.
49. **Lippert K, Galinski EA.** 1992. Enzyme stabilization be ectoine-type compatible solutes: protection against heating, freezing and drying. Appl. Microbiol. Biotechnol. **37**:61–65.
50. **Walker CB, de la Torre JR, Klotz MG, Urakawa H, Pinel N, Arp DJ, Brochier-Armanet C, Chain PS, Chan PP, Gollabgir A, Hemp J, Hügler M, Karr EA, Könneke M, Shin M, Lawton TJ, Lowe T, Martens-Habbena W, Sayavedra-Soto LA, Lang D, Sievert SM, Rosenzweig AC, Manning G, Stahl DA.** 2010. Nitrosopumilus maritimus genome reveals

unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. Proc. Natl. Acad. Sci. U. S. A. **107**:8818–8823.
51. **Goltsman DS, Denef VJ, Singer SW, VerBerkmoes NC, Lefsrud M, Mueller RS, Dick GJ, Sun CL, Wheeler KE, Zemla A, Baker BJ, Hauser L, Land M, Shah MB, Thelen MP, Hettich RL, Banfield JF.** 2009. Community genomic and proteomic analyses of chemoautotrophic iron-oxidizing "Leptospirillum rubarum" (group II) and "Leptospirillum ferrodiazotrophum" (group III) bacteria in acid mine drainage biofilms. Appl. Environ. Microbiol. **75**:4599–4615.
52. **Belnap CP, Pan C, VerBerkmoes NC, Power ME, Samatova NF, Carver RL, Hettich RL, Banfield JF.** 2010. Cultivation and quantitative proteomic analyses of acidophilic microbial communities. ISME J. **4**:520–530.
53. **Belnap CP, Pan C, Denef VJ, Samatova NF, Hettich RL, Banfield JF.** 2011. Quantitative proteomic analyses of the response of acidophilic microbial communities to different pH conditions. ISME J. **5**:1152–1161.
54. **Mueller RS, Dill BD, Pan C, Belnap CP, Thomas BC, VerBerkmoes NC, Hettich RL, Banfield JF.** 2011. Proteome changes in the initial bacterial colonist during ecological succession in an acid mine drainage biofilm community. Environ. Microbiol. **13**:2279–2292.
55. **Jiao Y, Cody GD, Harding AK, Wilmes P, Schrenk M, Wheeler KE, Banfield JF, Thelen MP.** 2010. Characterization of extracellular polymeric substances from acidophilic microbial biofilms. Appl. Environ. Microbiol. **76**:2916–2922.
56. **Davey ME, O'Toole GA.** 2000. Microbial biofilms: from ecology to molecular genetics. Microbiol. Mol. Biol. Rev. **64**:847–867.
57. **Jebbar M, Sohn-Bösser L, Bremer E, Bernard T, Blanco C.** 2005. Ectoine-induced proteins in Sinorhizobium meliloti include an ectoine ABC-type transporter involved in osmoprotection and ectoine catabolism. J. Bacteriol. **187**:1293–1304.
58. **Schwibbert K, Marin-Sanguino A, Bagyan I, Heidrich G, Lentzen G, Seitz H, Rampp M, Schuster SC, Klenk H-P, Pfeiffer F, Oesterhelt D, Kunte HJ.** 2011. A blueprint of ectoine metabolism from the genome of the industrial producer Halomonas elongata DSM 2581(T). Environ. Microbiol. **13**:1973–1994.
59. **Vargas C, Jebbar M, Carrasco R, Blanco C, Calderón MI, Iglesias-Guerra F, Nieto JJ.** 2006. Ectoines as compatible solutes and carbon and energy sources for the halophilic bacterium Chromohalobacter salexigens. J. Appl. Microbiol. **100**:98–107.
60. **Amann RI, Ludwig W, Schleifer KH.** 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. **59**:143–169.
61. **Bond PL, Banfield JF.** 2001. Design and performance of rRNA targeted oligonucleotide probes for in situ detection and phylogenetic identification of microorganisms inhabiting acid mine drainage environments. Microb. Ecol. **41**:149–161.
62. **Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, Kuo A, Minovitsky S, Nikitin R, Ohm RA, Otillar R, Poliakov A, Ratnere I, Riley R, Smirnova T, Rokhsar D, Dubchak I.** 2012. The genome portal of the department of energy joint genome institute. Nucleic Acids Res. **40**:D26–D32. http://dx.doi.org/10.1093/nar/gkr947.
63. **Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M.** 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. **35**:W182–W185. http://dx.doi.org/10.1093/nar/gkm321.