

ARTICLE



Microbial carbon, sulfur, iron, and nitrogen cycling linked to the potential remediation of a meromictic acidic pit lake

Diana Ayala-Muñoz¹ , Jennifer L. Macalady² , Javier Sánchez-España³ , Carmen Falagán⁴ , Estelle Couradeau⁵ and William D. Burgos¹

© The Author(s), under exclusive licence to International Society for Microbial Ecology 2022

Cueva de la Mora is a permanently stratified acidic pit lake and a model system for extreme acid mine drainage (AMD) studies. Using a combination of amplicon sequencing, metagenomics and metatranscriptomics we performed a taxonomically resolved analysis of microbial contributions to carbon, sulfur, iron, and nitrogen cycling. We found that active green alga *Coccomyxa onubensis* dominated the upper layer and chemocline. The chemocline had activity for iron(II) oxidation carried out by populations of *Ca. Acidulodesulfobacterium*, *Ferroplasma*, *Leptospirillum*, and *Armatimonadetes*. Predicted activity for iron(III) reduction was only detected in the deep layer affiliated with *Proteobacteria*. Activity for dissimilatory nitrogen cycling including nitrogen fixation and nitrate reduction was primarily predicted in the chemocline. Heterotrophic archaeal populations with predicted activity for sulfide oxidation related to uncultured *Thermoplasmatales* dominated in the deep layer. Abundant sulfate-reducing *Desulfomonile* and *Ca. Acidulodesulfobacterium* populations were active in the chemocline. In the deep layer, uncultured populations from the bacterial phyla *Actinobacteria*, *Chloroflexi*, and *Nitrospirae* contributed to both sulfate reduction and sulfide oxidation. Based on this information we evaluated the potential for sulfide mineral precipitation in the deep layer as a tool for remediation. We argue that sulfide precipitation is not limited by microbial genetic potential but rather by the quantity and quality of organic carbon reaching the deep layer as well as by oxygen additions to the groundwater enabling sulfur oxidation. Addition of organic carbon and elemental sulfur should stimulate sulfate reduction and limit reoxidation of sulfide minerals.

The ISME Journal; <https://doi.org/10.1038/s41396-022-01320-w>

INTRODUCTION

Acidic pit lakes are common throughout the world and develop in flooded former open-pit mines. They are strongly affected by input of acid mine drainage (AMD) generated by oxidation of sulfide minerals from the surrounding bedrock. As a consequence, they are characterized by low pH and high concentrations of sulfate, iron, and toxic metal(loid)s [1]. As in other AMD sites, their environmental impacts are long-lasting and represent a challenge for management and remediation [2, 3]. Acidic pit lakes host unique and poorly understood microbial communities shaping their geochemistry. A better understanding of the complex connections between microbiological and geochemical processes in acidic pit lakes can therefore contribute (i) to improve environmental management and remediation plans as well as (ii) to the discovery of novel taxa and enzymes with potential applications in biotechnology.

Acidic pit lakes often develop stratified layers due to their steep walls and narrow geometry, typically with increasing pH and higher concentrations of sulfate and dissolved iron and metals with depth [4]. Each layer represents a habitat with different light, nutrient, and chemical energy availability. The upper oxygenated layer is frequently colonized by photoautotrophic green algae [1],

iron oxidizers such as *Acidithrix ferrooxidans* [5], and heterotrophic acidophiles such as *Acidiphilium* spp. [6]. *Leptospirillum*, *Acidithiobacillus*, and *Desulfomonile* species involved in iron and/or sulfur cycling have been reported in chemoclines [7]. Unlike in neutral-pH stratified lakes, phototrophic sulfide oxidizers are rare or absent in acidic pit lakes (Schultze et al., 2017). In deep anoxic layers, the reported presence of heterotrophic archaea related to *Thermoplasmata* [6], sulfate-reducing bacteria (*Desulfomonile* spp.), and sulfur-disproportionating bacteria (*Desulfocapsa* spp.) provide evidence that biosulfidogenesis under acidic conditions is possible and even common [8].

Among acidic pit lakes, Cueva de la Mora (CM) has emerged as a valuable model system to study biogeochemical processes because of its permanent stratification, well characterized and dramatic vertical physicochemical gradients, and microbial activity [9, 10]. CM is over 50-years old and one of more than 25 pit lakes located in the Iberian Pyrite Belt (IPB) [9]. The upper layer is oxic and 10-m thick with minor seasonal fluctuations due to evaporation in summer and dilution in winter. The deep layer is anoxic and 30-m thick with no major seasonal or interannual fluctuations. The upper and deep layers are separated by a 2 m-thick chemocline [11].

¹Department of Civil and Environmental Engineering, The Pennsylvania State University, 212 Sackett Building, University Park, PA 16802, USA. ²Department of Geosciences, The Pennsylvania State University, 211 Deike Building University Park, University Park, PA 16802, USA. ³Centro Nacional Instituto Geológico Minero de España (IGME), CSIC, Calera 1, 28760 Tres Cantos, Madrid, Spain. ⁴School of Biological Sciences, University of Portsmouth, King Henry Building, King Henry 1st St., Portsmouth PO1 2DY, UK. ⁵Department of Ecosystem Science and Management, The Pennsylvania State University, 50 ASI University Park, University Park, PA 16802, USA. ✉email: diaayala@gmail.com; wdb3@psu.edu

Received: 11 January 2022 Revised: 2 September 2022 Accepted: 7 September 2022

Published online: 19 September 2022

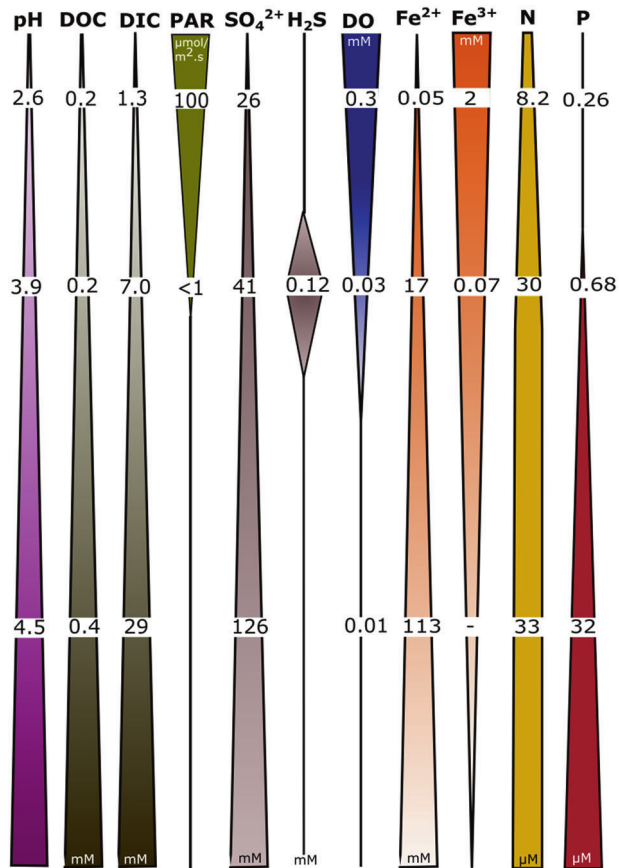


Fig. 1 Geochemical gradients along depth of the acidic pit lake Cueva de la Mora. Values represent the upper layer (3 m), chemocline (11 m), and deep layer (35 m) of the lake, respectively.

Due to oxidation of massive metal sulfides and the scarcity of carbonate minerals in the bedrock, the CM lake water is acidic with high sulfate and metal content [12]. The upper layer has a pH between 2.2 and 3.1, sulfate concentration of 26 mM, and dissolved total iron of 2 mM (mainly in the form of ferric iron). In contrast, the deep layer has a pH that increases with depth from 3.0 to 4.5, dissolved sulfate from 41 to 126 mM, and dissolved iron from 17 to 113 mM (mainly in the form of ferrous iron) (Fig. 1) [9]. Dissolved ferric iron precipitates as schwertmannite [$\text{Fe}_8\text{O}_8(\text{OH})_6(\text{SO}_4) \cdot n \text{H}_2\text{O}$] and jarosite [$\text{KFe}_3(\text{SO}_4)_2(\text{OH})_6$] in the upper layer [1]. The formation of schwertmannite is kinetically favored over other minerals in acid sulfate waters and buffers the pH of the system to around 2.6–3.1 [13]. At lower pH values (2.2–2.5), schwertmannite is transformed into jarosite in a proton-consuming process that represents an additional pH control in acidic pit lakes [14]. The deep layer and chemocline, on the other hand, are undersaturated with respect to Fe(III) mineral phases, allowing chemical redissolution or microbial reduction of Fe(III) particles as they settle in the water column. Metal(loid)s adsorbed on Fe(III) precipitates would then be released into solution by dissolution of the Fe(III) minerals. Long-term water-rock interactions also favored the high dissolved concentrations of toxic metal(loid)s in the isolated deep layer. Dissolved arsenic concentrations are ~ 0.01 mM in the upper layer and ~ 0.2 mM in the deep layer; similarly, dissolved zinc concentrations are ~ 0.02 mM in the upper layer and ~ 1.5 mM in the deep layer [15].

Despite its extreme geochemistry, CM has a relatively high nutrient content compared to other acidic pit lakes [10]. Organic and inorganic carbon, phosphorus, and nitrogen concentrations all increase with depth due to permanent stratification. Organo-mineral interactions facilitate transport of C, N, and P to the deep

water via adsorption to Fe(III) minerals [16, 17], physical settling, and release via reductive dissolution of the Fe(III) minerals [1]. Phosphate concentrations increase from $0.26 \mu\text{M}$ in the upper layer to $32 \mu\text{M}$ in the deep layer and dissolved organic carbon concentrations increase from $10 \mu\text{M}$ in the upper layer to $29 \mu\text{M}$ (C as glucose) (Fig. 1). Another nutrient transport mechanism occurs via Fe storage in phytoplankton and heterotrophs [18] followed by cell death, further encrustation by Fe(III) minerals, and settling into the deep layer. Total nitrogen increases with depth from $8.2 \mu\text{M}$ in the upper layer to $33 \mu\text{M}$ in the deep layer (Fig. 1) [19]. Carbon dioxide is also present at higher concentrations in the deep layer (29 mM) due to decomposition of organic debris and dissolution of trace carbonates present in the host rock (Fig. 1) [1, 10]. Despite the continuous export of particles and organic matter to the deep waters, there are sufficient nutrients to support the growth of green algae and heterotrophic bacteria in the upper layer as well as diverse microbial populations involved in iron and sulfur cycling in the chemocline [7, 20].

Although previous work has been conducted on the microbiology of CM, the deep layer has remained poorly characterized, and microbial activity under in situ conditions has never been assessed in any layer. As a path to a holistic understanding of the CM ecosystem, we used metagenomics and metatranscriptomics to compare microbial populations and metabolisms at three depths representative of the upper layer, chemocline, and deep layer. Our main questions were: [1] which are the predominant populations in each layer of the lake? and [2] what are their contributions to carbon, sulfur, iron, and nitrogen cycling? We found that algal primary production in the upper layer fueled both oxidative and reductive redox reactions in the chemocline. The deep layer microbial community carried out an unexpected array of oxidative and reductive processes and included many highly novel taxa that were both numerous and active.

RESULTS AND DISCUSSION

Microbial community composition with depth

Each layer in CM had a distinct microbial community composition reflecting strong differences in geochemistry and light availability. Based on 16S rRNA and 18S rRNA genes reconstructed with EMIRGE [21] from the metagenomic datasets (MG barplots in Fig. 2), within the 2 to $0.2 \mu\text{m}$ size fraction, Eukaryotes dominated the upper layer, while Archaea dominated the deep layer, and a combination of Bacteria and Eukaryotes were abundant at the chemocline. Based on 16S rRNA amplicon sequencing for bacteria and archaea, the upper layer was the least diverse as evidenced by the low number of amplicon sequence variants (ASVs) (72 and 46 from replicate samples). The two replicate samples from the chemocline had 119 and 167 ASVs, and the replicates from the deep layer had 192 and 206 ASVs (Table S1). A principal component analysis (PCoA) of the Bray Curtis pairwise dissimilarity metric based on the 16S rRNA gene count showed that the replicates per layer clustered together and separated from those of the other layers (Fig. 2B).

Based on 16S rRNA gene sequencing, Proteobacteria were dominant in the upper layer ($76\% \pm 8.9\%$) and chemocline ($68\% \pm 3.5\%$) and less abundant ($4.6\% \pm 0.41\%$) in the deep layer (Fig. 2). The most abundant ASV (57%, Table S1) in the upper layer was affiliated with the genus *Acidiphilium*, consistent with data reported by Falagán et al. [7] in CM and in other acidic pit lakes such as Lake 111 in Germany [4] and La Zarza and Filon Centro in Spain [8]. In the chemocline, the most abundant ASVs were affiliated with the genera *Desulfomonile* (43%) and the order *Candidatus Acidulodesulfobacterales* (formerly Sva0485 clade, 8.1%). *Desulfomonile* was previously reported in the chemocline of CM [7] and was also reported in other acidic pit lakes in the IPB [8]. These populations are expected to contribute to sulfate reduction [22, 23]. Other abundant ($>2\%$, Table S1) Proteobacteria

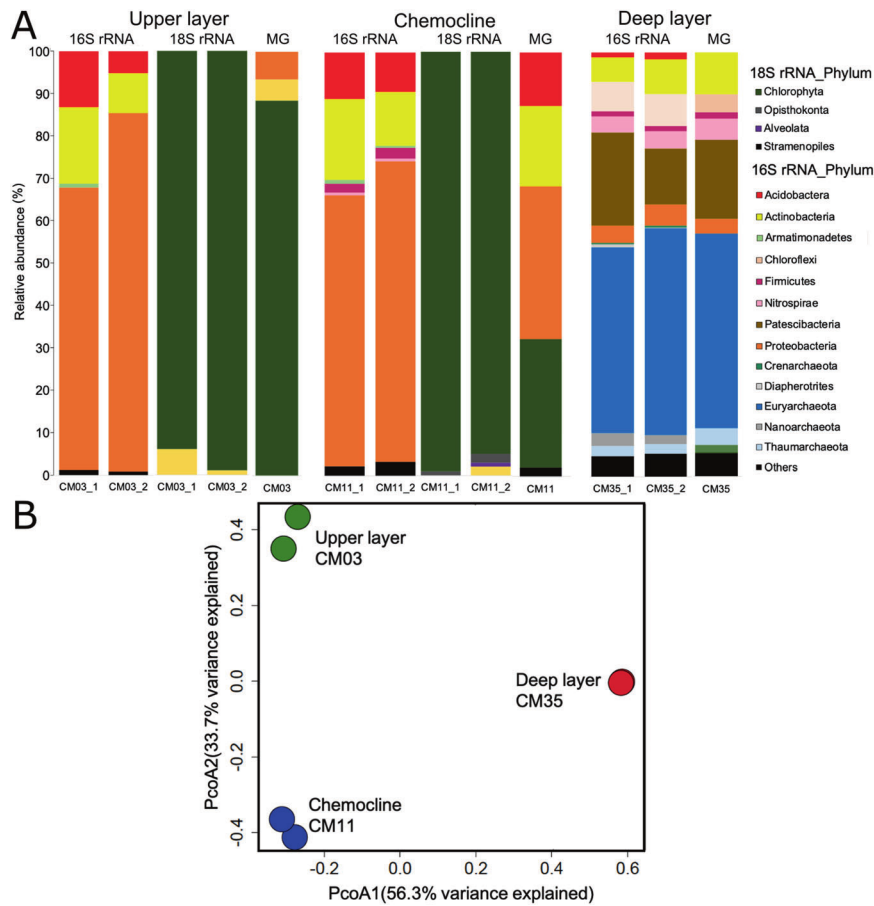


Fig. 2 Microbial diversity at the phylum level based on amplicon (16S and 18S rRNA genes) sequencing of microbial communities from Cueva de la Mora at three different depths representing the upper layer (CM03, 3-m), chemocline (CM11, 11-m), and deep layer (CM35, 35 m). **A** Barplots represent the relative abundances of Amplicon Sequence Variants (ASVs) grouped at the phyla level recovered from two biological replicates (1 and 2) per layer. The most abundant ASVs and their closest taxonomic affiliations are displayed in Table S1. 18S rRNA gene sequencing analysis was not performed for the deep layer. Barplots with the heading 'MG' represent the relative abundance of 16S/18S rRNA gene sequences reconstructed with EMIRGE from metagenomes of the respective layers as described in Ayala-Munoz et al. (2020). * Phylum not reconstructed by EMIRGE. **B** Principal Component Analysis (PCoA) of the Bray Curtis pairwise dissimilarity metric based on the ASVs table (16S rRNA gene count) from replicated biological samples collected from the upper layer (CM03), chemocline (CM11), and deep layer (CM35) of Cueva de la Mora.

ASVs were affiliated with iron oxidizers in the genus *Ferroplasma*, iron-reducers related to *Metallibacterium* and *Acidiphilium* species, and sulfur-oxidizers in the genus *Sulfuriferula*. *Ferroplasma* and *Acidiphilium* have been found forming iron-rich microscopic aggregates in the redoxcline of acidic pit lakes [24]. In the deep layer, most Proteobacteria ASVs were affiliated with an unclassified Deltaproteobacteria (2.5%). Other less abundant Proteobacteria ASVs (< 1%) included *Desulfomonile*, *Ca. Acidulosulfobacterium*, *Desulfobacca*, *Desulfatiferula*, and the iron-reducing genera *Geobacter* and *Acidiphilium*.

Actinobacteria populations made up to $14\% \pm 4.4\%$ of the ASVs in the upper layer, $16\% \pm 3.2\%$ in the chemocline, and $6.9\% \pm 1.2\%$ in the deep layer (Fig. 2). The most abundant ASVs (Table S1) were affiliated with iron oxidizers in the genus *Ferrithrix* (8.8%) in the upper layer, uncultured Acidimicrobiia (8.1%) in the chemocline, and uncultured Gaiellales (5.7%) in the deep layer. *Ferrithrix* and *Gaiellales* were also abundant in the chemocline (2.5 and 4.2%, respectively).

Acidobacteria was the third most abundant phylum in the upper layer ($9.2\% \pm 4.0\%$) and chemocline ($10\% \pm 0.85\%$) but less abundant in the deep layer ($1.4\% \pm 0.26\%$) (Fig. 2). ASVs related to the genus *Acidicapsa* were abundant in the upper layer and chemocline (7.3 and 5.5%, respectively, Table S1), where they are potential contributors to ferric iron reduction [7].

Armatimonadetes populations were in very low abundance in the upper layer ($0.44\% \pm 0.41\%$) and chemocline ($0.68\% \pm 0.22\%$) and not detected in the deep layer (Fig. 2). An ASV affiliated with the genus *Chthonomonas* was found in the upper layer and chemocline (Table S1). Armatimonadetes have been reported in other acidic systems, usually in low abundance [25, 26] and previous work suggests that they may be organoheterotrophic scavengers [27].

Firmicutes were more abundant in the chemocline ($2.3\% \pm 0.30\%$) than in the upper layer ($0.04\% \pm 0.02\%$) or the deep layer ($1.2\% \pm 0.05\%$) (Fig. 2). ASVs affiliated with the genus *Acidibacillus*, which include sulfur and iron oxidizers [28], were the most abundant in the chemocline (1.1%, Table S1). ASVs related to family Peptococcaceae and genera *Desulfitobacterium* and *Desulfosporosinus* were found in low abundance (<1%) in the deep layer (Table S1). Some *Desulfosporosinus* species studied to date are acid-tolerant sulfate-reducers [29, 30].

The Patescibacteria superphylum was the most abundant bacterial phylum in the deep layer ($18\% \pm 4.3$) followed by Chloroflexi ($7.3\% \pm 0.28\%$) (Fig. 2). The most abundant Patescibacteria ASVs were affiliated with the phylum Parcubacteria (11%) and other candidatus taxa (<2%) such as *Ca. Adlerbacteria*, *Ca. Colwellbacteria*, and *Ca. Wolfbacteria* (Table S1). The most abundant Chloroflexi ASV was affiliated with order

Table 1. Integrated evaluation of thermodynamics and gene expression for potentially operative metabolic processes in the lake layers. DeltaG Gibbs free energy in kJ mol⁻¹. MT Metatranscriptomic data. CM03 upper layer. CM11 chemocline. CM35 deep layer.

RXN#	Process/Metabolism	Layer	DeltaG	MT	RXN#	Process/Metabolism	Layer	DeltaG	MT
1	sulfur comproportionation $3\text{H}_2\text{S} + \text{SO}_4^{2-} + 2\text{H}^+ = 4\text{S}_{(0)} + 4\text{H}_2\text{O}$	CM03	-0.6	-	11	sulfide oxidation to sulfate with ferric iron reduction $\text{H}_2\text{S} + 8\text{Fe}^{3+} + 4\text{H}_2\text{O} = \text{SO}_4^{2-} + 8\text{Fe}^{2+} + 10\text{H}^+$	CM03	-604.2	×
		CM11	-2.1	-			CM11	-444.1	×
		CM35	-0.3	-			CM35	-504.2	✓
2	sulfur disproportionation $4\text{S}_{(0)} + 4\text{H}_2\text{O} = \text{SO}_4^{2-} + 3\text{H}_2\text{S} + 2\text{H}^+$	CM03	0.5	×	12	sulfur oxidation to sulfate with acetogenesis $6\text{CO}_2 + 4\text{S}_{(0)} + 10\text{H}_2\text{O} = 3\text{C}_2\text{H}_3\text{O}_2^- + 4\text{SO}_4^{2-} + 8\text{H}^+$	CM03	99.7	✓
		CM11	1.6	✓			CM11	95.6	✓
		CM35	0.2	✓			CM35	87.2	✓
3	sulfate reduction to sulfide with acetate oxidation $\text{C}_2\text{H}_3\text{O}_2^- + \text{SO}_4^{2-} + 3\text{H}^+ = 2\text{CO}_2 + \text{H}_2\text{S} + 2\text{H}_2\text{O}$	CM03	-132.3	×	13	sulfur oxidation to sulfate with nitrate reduction $\text{S}_{(0)} + 0.75\text{NO}_3^- + 1.75\text{H}_2\text{O} = \text{SO}_4^{2-} + 0.75\text{NH}_4^+ + 0.5\text{H}^+$	CM03	-332.0	✓
		CM11	-125.3	✓			CM11	-328.0	✓
		CM35	-116.0	✓			CM35	-331.0	✓
4	sulfate reduction to sulfide with methanotrophy $\text{CH}_4 + \text{SO}_4^{2-} + 2\text{H}^+ = \text{CO}_2 + \text{H}_2\text{S} + 2\text{H}_2\text{O}$	CM03	-37.2	×	14	sulfur oxidation to sulfate with ferric iron reduction $\text{S}_{(0)} + 6\text{Fe}^{3+} + 4\text{H}_2\text{O} = \text{SO}_4^{2-} + 6\text{Fe}^{2+} + 8\text{H}^+$	CM03	-452.6	×
		CM11	-30.9	×			CM11	-331.5	×
		CM35	-52.1	×			CM35	-377.9	✓
5	sulfide oxidation to sulfur with oxygen $\text{H}_2\text{S} + 0.5\text{O}_2 = \text{S}_{(0)} + \text{H}_2\text{O}$	CM03	-184.8	✓	15	sulfur oxidation to sulfate with oxygen $\text{S}_{(0)} + 1.5\text{O}_2 + \text{H}_2\text{O} = \text{SO}_4^{2-} + 2\text{H}^+$	CM03	-552.4	✓
		CM11	-99.7	✓			CM11	-292.9	✓
		CM35	-87.0	✓			CM35	-260.1	✓
6	sulfide oxidation to sulfate with oxygen $\text{H}_2\text{S} + 2\text{O}_2 = \text{SO}_4^{2-} + \text{H}^+$	CM03	-737.1	✓	16	sulfur reduction to sulfide with acetate oxidation $\text{C}_2\text{H}_3\text{O}_2^- + \text{S}_{(0)} + 0.5\text{H}_2\text{O} + 0.25\text{H}^+ = 0.5\text{CO}_2 + \text{H}_2\text{S}$	CM03	-32.6	×
		CM11	-392.6	✓			CM11	-29.7	×
		CM35	-347.1	✓			CM35	-28.8	✓
7	sulfide oxidation to sulfur with nitrate reduction $\text{H}_2\text{S} + 0.25\text{NO}_3^- + 0.5\text{H}^+ = \text{S}_{(0)} + 0.25\text{NH}_4^+ + 0.75\text{H}_2\text{O}$	CM03	-111.3	✓	17	sulfur reduction to sulfide with methanotrophy $0.25\text{CH}_4 + \text{S}_{(0)} + 0.5\text{H}_2\text{O} = 0.25\text{CO}_2 + \text{H}_2\text{S}$	CM03	-8.8	×
		CM11	-111.5	✓			CM11	-6.1	×
		CM35	-110.6	✓			CM35	-12.8	×
8	sulfide oxidation to sulfate with nitrate reduction $\text{H}_2\text{S} + \text{NO}_3^- + \text{H}_2\text{O} = \text{SO}_4^{2-} + \text{NH}_4^+$	CM03	-443.3	✓	18	ferric iron reduction to ferrous iron with methanotrophy $0.125\text{CH}_4 + \text{Fe}^{3+} + 0.25\text{H}_2\text{O} = 0.125\text{CO}_2 + \text{Fe}^{2+} + \text{H}^+$	CM03	-80.2	×
		CM11	-439.5	✓			CM11	-59.4	×
		CM35	-441.6	✓			CM35	-69.5	×
9	sulfide oxidation to sulfate with acetogenesis $2\text{CO}_2 + \text{H}_2\text{S} + 2\text{H}_2\text{O} = \text{C}_2\text{H}_3\text{O}_2^- + \text{SO}_4^{2-} + 3\text{H}^+$	CM03	132.3	✓	19	ferric iron reduction to ferrous iron with acetate oxidation $0.125\text{C}_2\text{H}_3\text{O}_2^- + \text{Fe}^{3+} + 0.25\text{H}_2\text{O} = 0.25\text{CO}_2 + \text{Fe}^{2+} + 0.875\text{H}^+$	CM03	-92.1	×
		CM11	125.3	✓			CM11	-71.2	×
		CM35	116.0	✓			CM35	-77.5	✓
10	sulfide oxidation to sulfate with methanogenesis $\text{CO}_2 + \text{H}_2\text{S} + 2\text{H}_2\text{O} = \text{CH}_4 + \text{SO}_4^{2-} + 2\text{H}^+$	CM03	37.2	×	20	ferrous iron oxidation to ferric iron with oxygen $\text{Fe}^{2+} + 0.25\text{O}_2 + \text{H}^+ = \text{Fe}^{3+} + 0.5\text{H}_2\text{O}$	CM03	-16.6	✓
		CM11	30.9	×			CM11	6.4	✓
		CM35	52.1	×			CM35	19.6	×

✓ Genes expressed for both half-cell reactions
 × Genes expressed for only one half cell or neither reactions

thermodynamically favorable + gene expression
 thermodynamically favorable + NO gene expression
 thermodynamically UN-favorable + gene expression
 thermodynamically UN-favorable + NO gene expression

Dehalococcoidales (1.8%, Table S1). Parcupacteria have been reported in other acidic systems but in low abundance [31]. In contrast, Chloroflexi have been found in high abundance in acidic environments [32], including acidic pit lakes [6]. Neither phylum has members with a clearly defined role in mine-impacted environments [6, 33].

Nitrospirae were more abundant in the deep layer (3.9% ± 0.15%) as compared to the upper layer (0.02% ± 0.02%) and chemocline (0.60% ± 0.04%) (Fig. 2). The most abundant ASVs in the chemocline were affiliated with the iron-oxidizing genus *Leptospirillum* (0.58%), and the most abundant ASVs in the deep layer were affiliated with the sulfate-reducing class *Thermodesulfobacteria* (3.5%) (Table S1) [34].

Archaea were dominant in the deep layer (Fig. 2). Euryarchaeota were the most abundant (46% ± 2.5%), followed by much smaller populations of Nanoarchaeota and Thaumarchaeota (~2% each), and Diapherotrites and Crenarchaeota (<0.4% each). The most abundant archaeal ASVs were affiliated with the family Thermoplasmatales and the class Methanomassiliicoccales (Euryarchaeota) (Table S1). Other less abundant archaeal ASVs were affiliated with the class Aigarchaeales (Crenarchaeota) and with group 1.1c and the family Nitrosotaleaceae (Thaumarchaeota) (Table S1). Thermoplasmata and Thaumarchaeota-like populations were previously reported in the deep layer of CM with a hypothesized contribution to carbon and nitrogen cycling [7].

Based on 18S rRNA gene sequencing, Chlorophyta (green algae) were abundant in the upper layer and chemocline (Fig. 2) and the dominant population was closely related to *Coccomyxa* sp. KN-2011-T3 later described as *Coccomyxa onubensis* (95% ± 2.5% of ASVs, Table S1) as previously reported [7, 20, 35].

At the phylum level, microbial community compositions based on 16S and 18S rRNA genes were consistent with the abundances of 16S/18S rRNA sequences obtained from the metagenomes using the bioinformatic tool EMIRGE [18, 21]. The most abundant phyla were captured by both methods (Fig. 2). For instance, Chlorophyta and Proteobacteria were the dominant phyla in the upper layer and chemocline and Euryarchaeota and Patiscibacteria were the dominant phyla in the deep layer. However, EMIRGE-reconstruction of 16S/18S rRNA sequences from metagenomes did not detect rare phyla. The dominance of eukaryotic phyla impeded the reconstruction of less abundant bacterial and archaeal phyla from the metagenomic sequencing dataset.

Integrated omic and thermodynamic evaluation of potentially operative metabolic processes in Cueva de la Mora

The bioenergetic favorability of various metabolic processes occurring under the measured conditions in each layer of the lake were evaluated by thermodynamic calculations (described below). Thermodynamic predictions were compared to corresponding gene expression data to establish potentially operative metabolic processes in each layer of Cueva de la Mora (Table 1). A color scheme was used in Table 1 to group reactions based on thermodynamic favorability (negative ΔG_{rxn}) and the expression of genes required for the corresponding metabolic process (detected in metatranscriptomes). Reactions in green were thermodynamically favorable with corresponding gene expression. Reactions in yellow were thermodynamically favorable but with no corresponding gene expression. Reactions in orange were thermodynamically unfavorable but with corresponding gene expression. Reactions in red were thermodynamically unfavorable with no

corresponding gene expression. Reactions in green were consistent with the expectation that genes would be expressed for thermodynamically favorable reactions. Reactions in red were consistent with the expectation that genes would not be expressed for thermodynamically unfavorable reactions. Reactions in yellow and orange showed inconsistencies between thermodynamic predictions and gene expression.

Element cycling in Cueva de la Mora

In the following sections, we integrate our discussion of microbial elemental cycling in terms of thermodynamic favorability and gene expression on an element-by-element basis. We focus on the cycling of carbon (Fig. 3), sulfur (Fig. 4), iron (Fig. 5), and nitrogen (Fig. 5) in each layer of the lake. We note when thermodynamic favorability and gene expression are theoretically consistent and attempt to explain situations when they are not.

Note that our analysis is limited by the current databases for annotation. We cannot rule out that some functions that we did not identify could in fact be occurring while being performed by enzyme in pathways yet to be discovered. We also acknowledge that our annotation pipeline did not pick up Horizontal Gene Transfers (HGT) which is an important mechanism of microbial ecology and evolution. We can't exclude that there might exist some discrepancies between the taxonomy reported and the function as some of these genes might exist in various genomic chassis if they are subjected to HGT.

Metagenomic and metatranscriptomic sequencing revealed important microbial contributions to element cycling in each layer. A list of marker genes with KEGG-Orthology identifiers (KO) was utilized in the analyses of carbon, sulfur, and nitrogen metabolisms (Table S2). Genes involved in iron metabolisms were identified by FeGenie [36]. Six metagenomes and five metatranscriptomes were obtained in total. The metagenomes ranged from 32 to 89 million reads with a percentage of conserved reads from 48 to 71% after quality filtering (Table S3). The

metatranscriptomes ranged from 13 to 74 million reads with a percentage of conserved reads ca. 70% after quality filtering (Table S3). Duplicate samples yielded similar results with respect to carbon, sulfur, iron, nitrogen (also oxygen and phosphorus) metabolisms (Figs. S1, S2, and S3).

Carbon cycling

Populations of the green microalga *Coccomyxa onubensis* were the most abundant and active cells in the upper layer. Members of the genus *Coccomyxa* from the Chlorophyta phylum fix inorganic carbon via the Calvin-Benson-Bassham cycle (CBB, a.k.a. reductive pentose phosphate cycle) [37]. Consistent with this, CBB transcripts (e.g., Form I RubisCO, *cbbL*) were highly abundant (transcripts per million, TPM~10⁴, Fig. 3) and taxonomically affiliated with *Coccomyxa* (Table S4). Despite low levels of photosynthetically active radiation (PAR; < 1 $\mu\text{mol m}^{-2} \text{s}^{-1}$) in the chemocline [38], *Coccomyxa* expressed autotrophy (CBB) and other genes (TCA and pentose phosphate) even at 11 meters depth (Fig. 3, Table S5). The dominance of *Coccomyxa* in both layers might be explained by the lack of predation and interspecies competition for nutrients [39]. Inorganic carbon concentrations in the upper layer and chemocline are between 1.3 to 7 mM of carbon, with nitrogen concentrations of 8 to 30 μM as nitrate and ammonium, and phosphorus of up to 0.7 μM as phosphate [10, 19, 20].

Based on our omics dataset, both eukaryotic algae and bacteria fix carbon in the chemocline (Fig. 3). Actinobacteria in the chemocline also expressed Form I RubisCO (*cbbL*, CBB) (Table S5) as did acidophilic Actinobacteria from the class Acidimicrobiia [40]. Proteobacteria expressed genes from three autotrophic pathways: CBB (Ferroplasma), Wood-Ljungdahl (Desulfotomaculum), and reductive citrate cycle (rTCA, Ca. Acidulodesulfobacterales) (Table S5). Among the marker genes found for each pathway, a Form I RubisCO (*cbbL*) was found taxonomically related to *Ferroplasma*, CODH/ACS complex subunits (ABCE) were found related to

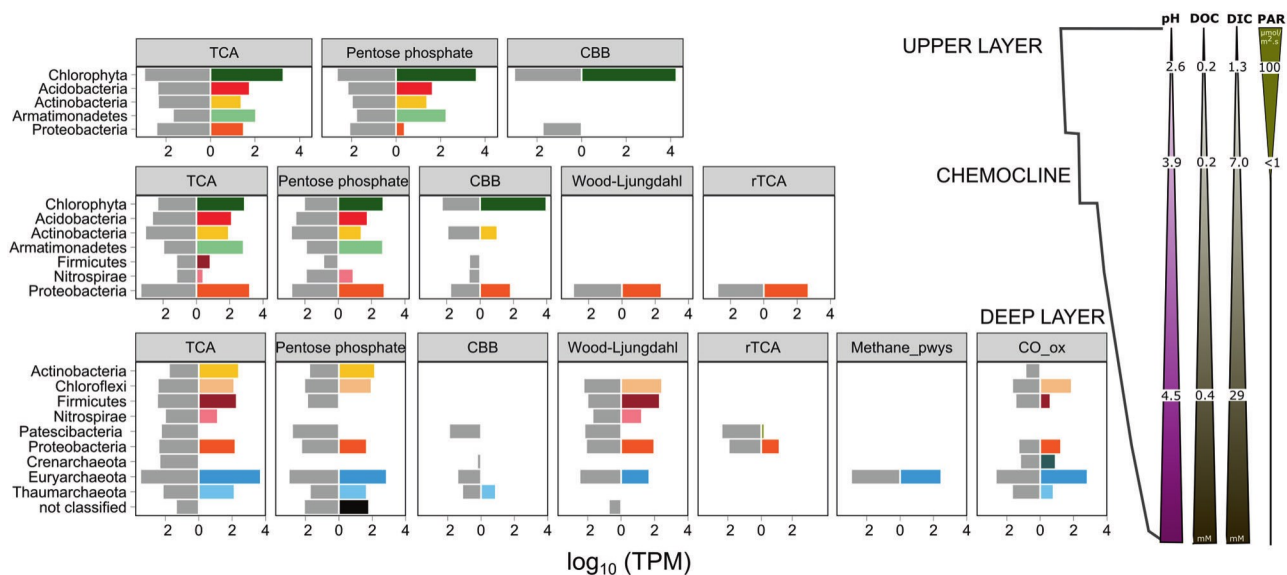


Fig. 3 Carbon metabolisms in the upper layer (CM03, 3-m), chemocline (CM11, 11-m), and deep layer (CM35, 35-m) of Cueva de la Mora and contributions per phyla (left-side). Physicochemical profiles displayed on right-side. Marker genes (Table S2) for specific reactions and/or pathways were annotated in the co-assembled metagenomes. Quality-filtered DNA and mRNA reads from each biological replicate were mapped to all genes predicted from co-assembled metagenomes. TPM (transcripts per million) values per predicted gene were calculated as a proxy for gene and transcript abundance (TPM number of DNA or mRNA reads mapped to one gene ($\times 10^6$) divided by its length/sum of number of DNA or mRNA reads mapped to all genes divided by respective lengths). Bars represent the sum of TPM values of all genes in a pathway. Gray bars represent the relative abundance of genes (DNA-TPM values, metagenome). Colored bars represent the relative abundance of transcripts (mRNA-TPM values, metatranscriptome) and correspond to colors of phyla used in Fig. 2. Results are shown for samples CM03_1, CM11_1, and CM35_1. Results from all replicates are shown in Figs. S1-S3. TCA Citric acid cycle, CBB Calvin-Benson-Bassham cycle or reductive pentose phosphate cycle, Wood-Ljungdahl reductive acetyl-CoA pathway, rTCA reductive citrate cycle, _pwys pathways; CO_ox carbon monoxide oxidation, DOC dissolved organic carbon, DIC dissolved inorganic carbon, PAR photosynthetically active radiation.

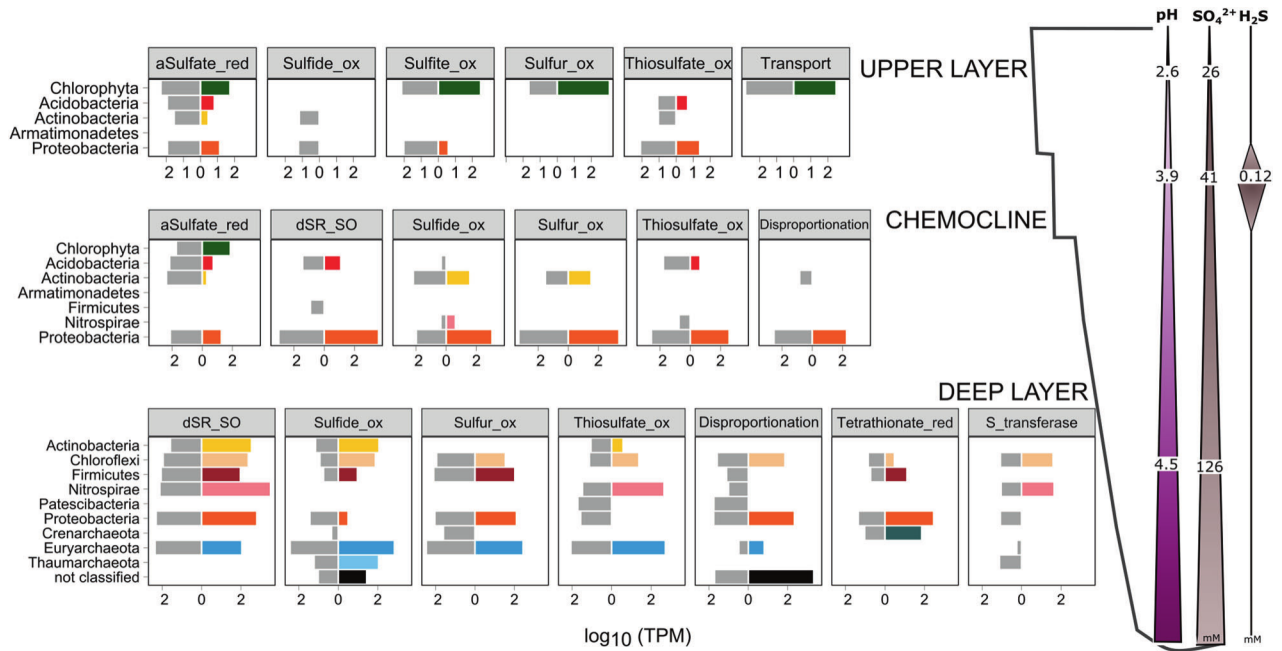


Fig. 4 Sulfur metabolisms in the upper layer (CM03, 3 m), chemocline (CM11, 11 m), and the deep layer (CM35, 35 m) of Cueva de la Mora and contributions per phyla (left-side). Physicochemical profiles displayed on right-side. Marker genes (Table S2) for specific reactions and/or pathways were annotated in the co-assembled metagenomes. Quality-filtered DNA and mRNA reads from each biological replicate were mapped to all genes predicted from the co-assembled metagenomes. TPM (transcripts per million) values per predicted gene were calculated as a proxy for gene and transcript abundance (TPM number of DNA or mRNA reads mapped to one gene ($\times 10^9$) divided by its length/sum of number of DNA or mRNA reads mapped to all genes divided by respective lengths). Bars represent the sum of TPM values of all genes in a pathway. Gray bars represent the relative abundance of genes (DNA-TPM values, metagenome). Colored bars represent the relative abundance of transcripts (mRNA-TPM values, metatranscriptome) and correspond to colors of phyla used in Fig. 2. Results are shown for samples CM03_1, CM11_1, and CM35_1. Results from both replicates are shown in Figs. S1-S3. dSR_SO dissimilatory sulfate reduction and/or sulfide oxidation, _ox oxidation; _red reduction.

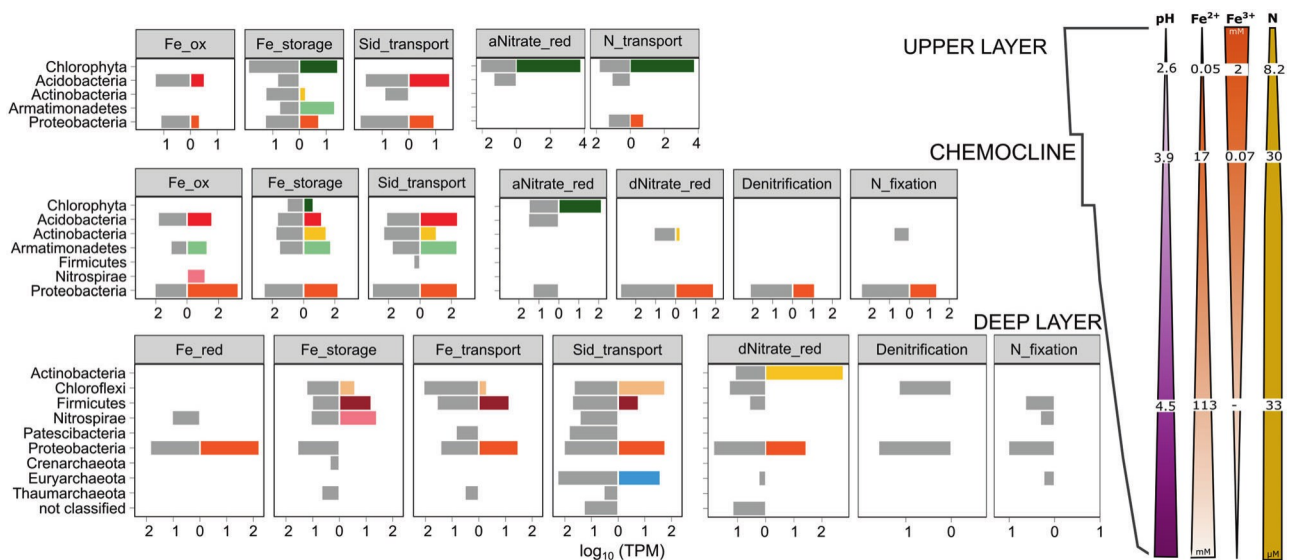


Fig. 5 Iron and nitrogen metabolisms in the upper layer (CM03, 3-m), chemocline (CM11, 11-m), and the deep layer (CM35, 35-m) of Cueva de la Mora and contributions per phyla (left-side). Physicochemical profiles displayed on right-side. Marker genes (Table S2) for specific reactions and/or pathways were annotated in the co-assembled metagenomes. Quality-filtered DNA and mRNA reads from each biological replicate were mapped to all genes predicted from the co-assembled metagenomes. TPM (transcripts per million) values per predicted gene were calculated as a proxy for gene and transcript abundance (TPM number of DNA or mRNA reads mapped to one gene ($\times 10^9$) divided by its length/sum of number of DNA or mRNA reads mapped to all genes divided by respective lengths). Bars represent the sum of TPM values of all genes part of a pathway. Gray bars represent the relative abundance of genes (DNA-TPM values, metagenome). Colored bars represent the relative abundance of transcripts (mRNA-TPM values, metatranscriptome) and correspond to colors of phyla used in Fig. 2. Results are shown for samples CM03_1, CM11_1, and CM35_1. Results from both replicates are shown in Figs. S1-S3. Fe iron, _ox oxidation, _red reduction, Sid siderophore, N Nitrogen, aNitrate_red assimilatory nitrate reduction, dNitrate_red dissimilatory nitrate reduction.

Desulfomonile, and a citryl-CoA lyase was found related to *Ca. Acidulodesulfobacterales* (Table S5). Taxonomic and functional assignments of these genes are consistent with previous work on the metabolisms of these three genera [22, 23, 41].

Organic carbon fixed by eukaryotic algae in the upper layer and chemocline support the growth of organoheterotrophic bacteria affiliated with *Acidobacteria*, *Actinobacteria*, *Armatimonadetes*, and *Proteobacteria*, as well as more diverse taxa in the deep layer. Bacteria in the upper layer had active citrate (TCA) and reductive pentose phosphate cycles (Fig. 3). Among the TCA cycle transcripts, those related to citrate synthase were phylogenetically affiliated with *Acidimicrobiales* (*Actinobacteria*), and *Acidiphilium* (*Proteobacteria*) (Table S4 and S5). *Acidobacteria* and *Armatimonadetes* in the chemocline lacked autotrophy genes and expressed genes involved in the TCA and reductive pentose phosphate cycles (Fig. 3), consistent with heterotrophic lifestyles [42, 43]. Some members of these bacterial phyla such as *Chthonomonas* and *Acidiphilium* have been described as organoheterotrophs [42]. In the upper layer, none of these putative organoheterotrophic bacteria were highly involved in sulfur, iron, or nitrogen cycling (see following sections), consistent with the proposal by Falagán et al. [7] that the dominant electron donor used by bacteria in this layer is algal organic carbon rather than ferrous iron or reduced sulfur species.

Both heterotrophic and autotrophic activity via gene expression were detected in the deep layer (Fig. 3). Many of the transcripts in the TCA and pentose phosphate cycles (TPM > 2000) were affiliated with *Thermoplasmatales* (Table S6) which contains taxa known for carbon scavenging [44–46]. Phyla with predicted autotrophic activity (Wood-Ljungdahl pathway) in the deep layer included *Chloroflexi*, *Firmicutes*, *Nitrospirae*, *Proteobacteria*, and *Euryarchaeota* (Fig. 3). The *Chloroflexi* transcripts were among the most abundant (TPM ~ 1000) and related to *Dehalococcoidia* (Table S6), a class that includes species known to have genes for this pathway [47]. *Firmicutes* transcripts involved in autotrophy were also abundant (TPM ~ 1000) and related to *Desulfallus* (Table S6), family *Peptococcaceae*, known for being lithoautotrophs [48]. Other less abundant transcripts (TPM ~ 100) were affiliated with *Thermodesulfovibrio* (*Nitrospirae*), *Desulfomonile* (*Proteobacteria*), and methanogenic archaeal genera such as *Methanomasilloccoccus* and *Methanoxithrix* (*Euryarchaeota*) (Table S6).

Heterotrophic and autotrophic activity in the deep layer was expected based on the dissolved organic and inorganic carbon concentrations at depth (Fig. 1). The concentration of dissolved organic carbon in the deep layer (0.44 mM) was higher than in the upper layer and chemocline (0.15–0.24 mM of carbon) [19]. Chemical and isotopic analyses performed in benthic sediments confirmed that one source of organic carbon in the deep layer was from sedimented phytoplankton [1]. The deep layer also had the highest concentration of dissolved inorganic carbon (29 mM versus 1–7 mM of C as carbon dioxide (CO₂) in the upper layer and chemocline [20], respectively). Carbon dioxide could be of biogenic origin from organic matter decomposition [10] or geogenic origin from the dissolution of trace carbonates present in the massive sulfide deposits [1].

Methanogenesis activity via gene expression was detected for *Euryarchaeota* (Fig. 3) affiliated with methanogenic genera such as *Methanosilicoccus*, *Methanosarcina*, and *Methanoxithrix* (Table S6). Among the transcripts involved in methanogenesis were *hdrA* (K03388), *mer* (K00320), and *fdhA* (K22516). However, no genes or transcripts considered as diagnostic indicators of methanogenesis (e.g., methyl-coenzyme M reductase, *mcr* [49]) were found in our dataset. Both methane concentrations (80–90 µM) [10] and putative methanogen populations from the order *Methanomassilicoccales* were relatively low (1.4% relative abundance based on 16S rRNA gene data, Table S1). Therefore, methane cycling is expected to play a minor role in carbon cycling in CM.

Activity for carbon monoxide (CO) oxidation was mostly predicted for the phyla *Euryarchaeota* (TPM~1000) (Fig. 3). Carboxydrotrophs, acetogens, sulfate-reducers, among others oxidize CO as a source of carbon or energy [50]. There is not geochemical evidence of the presence of carbon monoxide in CM, but according to our genomic evidence, CM deep layer populations have the potential to oxidize CO. Previous studies showed that some representatives of the *Thermoplasmatales* order, the so-called alphabet-plasmas, were predicted to oxidize CO [44].

Thermodynamic favorability for most carbon metabolisms did not vary as a function of depth (Table 1). For example, sulfate reduction coupled to acetate oxidation (reaction 5, R3) was always favorable while sulfide oxidation coupled to methanogenesis (R10) was always unfavorable. While sulfate reduction coupled to acetate oxidation (R3) was always favorable, genes were not expressed for this process in the upper layer (yellow color code) most likely because our thermodynamic predictions do not account for competitive processes. In the oxic upper layer of the lake, acetate oxidation would more likely be coupled to oxygen reduction versus sulfate reduction.

Theoretical consistency between thermodynamic favorability and gene expression differed for a number of carbon cycling reactions. Sulfate reduction coupled to acetate oxidation (R3) in the chemocline and deep layer was thermodynamically favorable with corresponding gene expression (green color code). Sulfide oxidation coupled to methanogenesis (R10) was always unfavorable in all layers with no corresponding gene expression (red color code). Sulfur(0) oxidation coupled with acetogenesis (R12) was thermodynamically unfavorable in all layers yet corresponding genes were detected (orange color code). In this case, the two half-reactions could be decoupled such that some microbes effectively complete half of the process. Ferric reduction coupled with methanotrophy (R18) was thermodynamically favorable in all layers yet no corresponding genes were detected (yellow color code). In this case, genes for this process may not have been expressed because of competitive utilization of the reactants (as noted above), or the genes for some of these reactions are not yet completely characterized (i.e., limited by target gene selection).

Sulfur cycling

Part of the sulfur cycling in the upper layer was characterized by gene transcripts related to sulfur assimilation and detoxification. *Chlorophyta* expressed genes related to assimilatory sulfate reduction, sulfite and sulfur oxidation, and sulfate/thiosulfate transport (Fig. 4). In particular, the *suox* gene (K00387) involved in sulfite oxidation and the *sdo* gene (K17725) involved in sulfur oxidation were highly expressed (TPM > 1000) (Table S4). Sulfite oxidase (*Suox*) is known for its role in scavenging sulfite molecules to prevent sulfite toxicity and maintain optimal metabolism of carbon, nitrogen, and sulfur [51]. Sulfur dioxygenase (*Sdo*) oxidizes sulfane sulfur of glutathione persulfide (GSS[•]) to sulfite and, along with sulfur-transferases, is also involved in sulfide detoxification [52].

Sulfur cycling in the chemocline, including both oxidative and reductive redox transformations, was likely carried out mainly by *Proteobacteria* (Fig. 4). The most abundant transcripts (TPM > 1000) involved in dissimilatory sulfate reduction and/or sulfide oxidation (*dSR_SO*, *aprAB*, *sat*, *dsrAB*, *qmoABC*) were affiliated with *Ca. Acidulodesulfobacterium* (Table S5). Other transcripts affiliated with this genus were *sqr* (sulfide oxidation), *hdrAC* (sulfur oxidation), and *doxDA* (thiosulfate oxidation) (Table S5). Populations for the order *Ca. Acidulodesulfobacterales* have been reported in other acidic systems and are thought to contribute to both dissimilatory sulfate reduction and sulfide oxidation [22]. Transcripts involved in dissimilatory sulfate reduction and/or sulfide oxidation were also found in *Desulfomonile*-related populations along with genes involved in sulfur oxidation and a

gene involved in thiosulfate disproportionation/sulfur reduction (*phsA/psrA*) (Table S5). Sulfide production in the CM chemocline has been reported previously, including the presence of metal sulfides and a peak in dissolved sulfide (120 μM) at 11-m depth [53]. Specifically, nano-sized particles of copper sulfide (covellite), zinc sulfide (wurtzite), and arsenic sulfide (realgar) seemed to precipitate selectively at different depths according to their respective pH-dependent solubility products [53]. This metal sulfide precipitation removes all of dissolved copper from the deep water and illustrates the potential for biosulfidogenesis to remediate extremely polluted pit lakes [53].

In the deep layer, organisms related to Actinobacteria, Chloroflexi, Firmicutes, Nitrospirae, and Proteobacteria were predicted to actively contribute to reductive sulfur processes such as sulfate and tetrathionate reduction (Fig. 4). Most genes from these phyla belonged to sufficiently novel taxa that they were not classified at lower taxonomic levels. Actinobacteria transcripts were affiliated with the order Gaiellales (Table S6), consistent with MAGs from Crystal Geyser (Utah, USA) described to have metabolic potential for sulfate/sulfite reduction [54]. Uncultured Chloroflexi found in the deep layer were likely involved in sulfate reduction. Other Chloroflexi populations with sulfate reduction potential were reported in the Black Sea [55]. Firmicutes transcripts were affiliated with the genera *Desulfotomaculum*, *Desulfallus*, and *Desulfofundulus* (Table S6), all known as strict anaerobes capable of using sulfate, sulfite, and thiosulfate as electron acceptors [48]. Nitrospirae had the highest abundance of transcripts (TPM > 1000) associated with dissimilatory sulfate reduction (Fig. 4). Some of the transcripts were affiliated with the genus *Thermodesulfovibrio* (Table S6). This genus contains obligate anaerobic, thermophilic bacteria able to reduce sulfate and other sulfur compounds [56]. Proteobacteria transcripts were affiliated with genera such as *Desulfobacca*, *Desulfoglaeba*, *Desulfomonile*, and *Sulfurifustis* (Table S6). Cultured members of all of these genera are known to reduce sulfate [23, 57, 58] with the exception of *Sulfurifustis* which has been described as an autotrophic sulfur-oxidizer [59]. Given the high concentration of sulfate in the deep layer (>120 mM), the predicted activity of these taxa for dissimilatory sulfate reduction could be important for attenuation of metal toxicity through metal sulfide precipitation.

Sulfur cycling could be important in the deep layer of CM as described in other meromictic lakes (e.g., Lake Kivu [60]) where interactions between sulfate-reducing bacteria and sulfur-oxidizing bacteria may occur. Unexpectedly, the microbial community in the deep layer included bacterial and archaeal transcripts for sulfur oxidation genes. Actinobacteria, Chloroflexi, Nitrospirae, and Euryarchaeota transcripts involved in sulfur oxidation processes (*Sulfide_ox*, *Sulfur_ox*, *Thiosulfate_ox*, Fig. 4) were abundant (TPM > 100) in the deep layer. It is notable that the most abundant taxa based on both 16S rRNA gene and metagenomic data in the deep layer, Thermoplasmatales (Euryarchaeota), contained transcripts annotated as putative sulfide quinone reductase (*sqr*) involved in sulfide oxidation to elemental sulfur (Table S6). Cultured Thermoplasmatales including *Thermoplasma acidophilum* are known to reduce elemental sulfur [61], but *T. acidophilum* also contained a *sqr* homologue potentially involved in sulfide oxidation [62]. Historically, high concentrations of free sulfide have not been detected in the deep layer of CM [53]. This lack of free sulfide could be caused by the precipitation of metal sulfides and/or cryptic sulfur cycling via biological sulfide oxidation in the deep layer.

If biological sulfide oxidation occurs in the deep layer, it is important to understand what electron acceptors are paired with sulfide/sulfur oxidation. Our findings include evidence for active oxygen respiration in populations related to Actinobacteria, Nitrospirae, and Euryarchaeota (Figure S3). Euryarchaeota-related transcripts annotated as cytochrome c oxidase *coxA* and cytochrome bd-l ubiquinol oxidase *cydA* were affiliated with the

order Thermoplasmatales (Table S6). We considered whether exposure of the samples to oxygen during filtration immediately prior to cryo-preservation could be responsible for the observed oxygen respiration activity. However, a strong argument against this explanation is that the oxygen respiration activity was associated with major populations adapted to the deep layer environment rather than rare, and poorly-adapted taxa. From a hydrological perspective, CM is a “recharge lake” where groundwater dynamics that could contribute to oxygen input to the deep layer are less evident than in other “flow-through” pit lakes in the IPB such as Herrerías or San Telmo [11]. Dissolved oxygen concentrations in the CM deep layer have consistently been below detection limits ($\sim 10 \mu\text{M}$). However, it is clear that even sub-nanomolar oxygen concentrations can support the activity of facultative anaerobes like *E. coli* [63]. Therefore, we consider it likely that oxygen is consistently or intermittently present at low/sub-nanomolar levels. Thermodynamic calculations demonstrate that sulfide (reactions 5 and 6, Table 1) and sulfur oxidation (reaction 15, Table 1) is favorable with nanomolar concentrations of oxygen. Under these conditions, it would not be surprising if novel or unexpected taxa (e.g., Thermoplasmatales) are respiring oxygen at this depth coupled to sulfur or sulfide oxidation.

Theoretical consistency between thermodynamic favorability and gene expression differed for several sulfur cycling reactions (Table 1). Sulfide oxidation coupled to oxygen reduction (R5, R6) or nitrate reduction (R7, R8) were all thermodynamically favorable with corresponding gene expression (green color code). In contrast, sulfide oxidation coupled with acetogenesis (R9) was thermodynamically unfavorable in all layers yet corresponding genes were detected (orange color code). In this case, the two half-reactions were likely decoupled such that some microbes effectively complete half of the process. Sulfide oxidation coupled to ferric reduction (R11) was thermodynamically favorable in all layers yet corresponding genes were only detected in the deep layer (green color code in the deep layer compared to yellow color code in the upper layer and chemocline). In the upper layers, sulfide oxidation coupled to oxygen or nitrate reduction may limit sulfide oxidation coupled to ferric reduction such that corresponding genes were not detected. In the deep layer, where oxygen concentrations are low (Fig. 1), competition for other electron acceptors is removed and sulfide oxidation coupled to ferric reduction becomes active.

Nitrate and dissolved iron(III) are present at low concentrations in the deep layer (0.4 μM and <0.01 μM , respectively) [20] and could also serve as electron acceptors for S(0) or sulfide oxidation. *Thermosulfurimonas dismutans* and *Dissulfuribacter thermophilus*, for example, can grow autotrophically with nitrate as the electron acceptor and S(0) as the electron donor [64]. Archaea of the genus *Sulfolobus* can reduce iron(III) coupled to S(0) oxidation [65]. Potential sources of oxidized nitrogen in the deep layer include explosives used in mine construction [10]. Although no transcripts involved in N or Fe cycling were related to Thermoplasmatales, they were present in other less abundant taxa (see below).

Iron cycling

The microbial community in the upper layer primarily expressed genes for iron storage and transport (Fig. 5). Transcripts involved in iron cycling in the upper layer were not as abundant (TPM < 100) as those involved in carbon cycling (TPM > 5000, Fig. 3), sulfur cycling (TPM > 1000, Fig. 5), or nitrogen cycling (TPM > 5000, Fig. 5). The most abundant transcripts (TPM > 50) involved in iron storage and transport were affiliated with Chlorophyta, Acidobacteria, and Armatimonadetes. Such transcripts may be required to protect cells from dissolved Fe(II) or from being covered with Fe(III) precipitates. Other less abundant transcripts (TPM < 10) were affiliated with Acidobacteria and Proteobacteria and annotated as *cyc2* (Table S4), a central gene involved in Fe(II) oxidation [66]. Dissolved iron in the upper layer is

almost 100% Fe(III). Photoreduction of Fe(III) [38], however, likely contributes to the temporal availability of Fe(II) (midday when sunlight is intense enough) to iron oxidizers such as *Ferrihithrix* spp. in the upper layer.

The most intense iron cycling activity via gene expression in CM was detected in the chemocline and included Fe(II) oxidation likely carried out by Proteobacteria (Ca. *Acidulodesulfobacterales* and *Ferroplasma*, TPM > 1000), Nitrospirae (*Leptospirillum*, TPM > 10), Acidobacteria (TPM > 50), and *Armatimonadetes* (TPM > 10) (Fig. 5, Table S5). Consistent with previous work by Tan et al. [22], Ca. *Acidulodesulfobacterales* expressed genes involved in Fe(II) oxidation and siderophore transport. *Ferroplasma* and *Leptospirillum* species are well known acidophilic iron oxidizers [41, 67, 68]. Some of the Acidobacteria transcripts were affiliated with the genus *Acidipila* but were not retrieved in our 16S rRNA gene data from the chemocline (Table S1). *Armatimonadetes* might be involved in Fe(II) oxidation activity, but they are usually found in low abundance in other acidic systems [25, 26] with no evidence on their ability for Fe(II) oxidation activity. Finally, relatives of *Acidicapsa* spp. (Acidobacteria, Table S1) in the chemocline have the potential to carry out Fe(III) reduction but no transcripts related to this metabolism were detected. Since we assume that Fe(III) reduction does occur in the chemocline, this finding was unexpected and suggests that there are alternative genes, pathways, or organisms carrying out Fe(III) reduction in this environment.

In the deep layer, the most abundant transcripts involved in iron cycling were mainly related to Fe(III) reduction (*OmcS* and *OmcB* polyheme membrane-associated cytochrome c) (TPM > 100), storage (TPM > 50), and transport (TPM ~100) carried out by Proteobacteria (Fig. 5). Most Proteobacteria transcripts were affiliated with the genus *Geobacter* (Table S6). Fe(III)-reducers like *Geobacter* could potentially use the iron in the Fe(III) minerals (e.g., schwertmannite) settling through the deep water column [15]. Transcripts related to iron storage or transport were affiliated with Chloroflexi, Firmicutes, Nitrospirae, and Euryarchaeota (Fig. 5). We found no evidence for microbial Fe(II)-oxidizing activity in the deep layer despite the likely presence of low micromolar concentrations of dissolved oxygen.

Theoretical consistency between thermodynamic favorability and gene expression differed for several iron cycling reactions (Table 1). Ferrous oxidation coupled to oxygen reduction (R20) was thermodynamically favorable in the upper layer but became increasingly unfavorable with depth because of the decreasing oxygen concentration. Corresponding genes for this process were detected in the upper layer (green color code) and the chemocline (orange color code because of thermodynamic unfavourability) but were not detected in the deep layer (red color code). Ferric reduction coupled to sulfide oxidation (R11) or sulfur(0) oxidation (R14) were thermodynamically favorable in all layers yet corresponding genes were only detected in the deep layer (green color code in the deep layer compared to yellow color code in the upper layer and chemocline). As noted above, sulfide or S(0) oxidation coupled to oxygen reduction may limit S oxidation coupled to ferric reduction such that these processes are only active when competition for other electron acceptors is removed. Ferric reduction coupled to acetate oxidation (R19) was thermodynamically favorable in all layers yet corresponding genes were only detected in the deep layer. Once again, ferric reduction becomes active only when competition for other electron acceptors is removed. Based on speciation analyses of filtered water from the deep layer, dissolved ferric iron is essentially absent while a limited flux of Fe(III) solids settle through the water column. Thus, 'cryptic' iron cycling is occurring in the deep layer based on gene expression (Fig. 5).

Nitrogen cycling

Nitrogen cycling has been reported in acidic systems [69] but has been poorly studied in acidic pit lakes. In CM, microbial

contributions to nitrogen cycling were predicted in all three layers. *Coccomyxa* are likely assimilating nitrogen in the upper layer and chemocline as evidenced by the high abundance of transcripts for transport and assimilation of nitrate/nitrite (TPM > 5000, Fig. 5, Table S4). *Coccomyxa onubensis*, the nearest cultivated relative of the most abundant eukaryotic ASVs in the upper layer and chemocline of CM, is known to grow photoautotrophically with either nitrate or ammonium or mixotrophically with urea as a nitrogen source [70].

The activity of nitrogen-fixers and nitrate-reducers in the chemocline and deep layer may explain increasing concentrations of ammonium with depth in CM. Previous geochemical studies in CM showed that nitrate concentrations decreased with depth (from ~7 μ M in the upper layer to ~0.4 μ M in the deep layer) while the ammonium concentrations increased with depth (from ~1 μ M to 32 μ M) [20]. In the chemocline, Proteobacteria contributed may contribute to nitrogen cycling through dissimilatory nitrate reduction, denitrification, and nitrogen fixation (TPM > 10, Fig. 5). Specifically, transcripts for denitrification (*nosZ* and *norB*), dissimilatory nitrate reduction (*narIG*), and nitrogen fixation (*nifKD*) were affiliated with the order Ca. *Acidulodesulfobacterales* (Table S5). Other transcripts for nitrogen fixation were affiliated with the genus *Ferroplasma* (Table S5). Both *Ferroplasma* and Ca. *Acidulodesulfobacterales* were previously reported to contribute to nitrogen cycling under acidic conditions [22, 41]. In the deep layer, genes involved in dissimilatory nitrate reduction, denitrification, and/or nitrogen fixation were present and affiliated with Actinobacteria, Chloroflexi, Firmicutes, Nitrospirae, and Proteobacteria (Fig. 5). However, only Actinobacteria and Proteobacteria had detectable transcripts for nitrate reduction. The only transcript within the phylum Actinobacteria that had a high abundance (TPM > 100) was annotated as a *narI* affiliated with the family Gaiellales (Table S6) consistent with nitrate reduction as a phenotypic characteristic of this order [71]. Within the Proteobacteria, some transcripts were annotated as *narIG* affiliated with the genera *Geobacter* and *Nitrobacter* (Table S6).

Biogeochemical model of acidic pit lake

In this study, we present taxonomically resolved contributions to microbial carbon, sulfur, iron, and nitrogen cycling activity via gene expression including the previously unstudied deep layer of CM. We used the most relevant findings to construct a conceptual biogeochemical model of elemental cycles occurring in the different layers of CM (Fig. 6). We revealed the dominance of the green algae *Coccomyxa onubensis* in the upper photic zone actively contributing to carbon fixation. Organic carbon produced by these phytoplanktonic cells supported the less abundant organoheterotrophic bacteria in the upper layer, chemocline, and deep layer. Active green algae were also detected in the low-light chemocline. Such activity could be related to access to nutrients such as phosphorus that becomes more available as Fe(III) minerals (formed in the upper layer) are bioreduced at depth. The chemocline was expected to have reductive and oxidative microbial active metabolisms that contributed to the sharp change in geochemical conditions at 11 m. One of the most abundant taxa in the chemocline were *Acidicapsa* populations associated with Fe(III) reduction. However, no genes related to this metabolism were annotated in either metagenomes or metatranscriptomes in the chemocline. Chemocline activity for Fe(II) oxidation was associated with abundant Ca. *Acidulodesulfobacterium* and less abundant *Leptospirillum*, *Ferroplasma*, and *Armatimonadetes* populations. Oxidation and reduction of sulfur compounds were likely carried out by *Desulfomonile* and Ca. *Acidulodesulfobacterium*, the most abundant taxa in the chemocline. The activity of nitrogen-fixers and nitrate-reducers in the chemocline associated with Ca. *Acidulodesulfobacterium* and *Ferroplasma* may explain increasing concentrations of ammonium with depth. It is hypothesized that organic particles from the

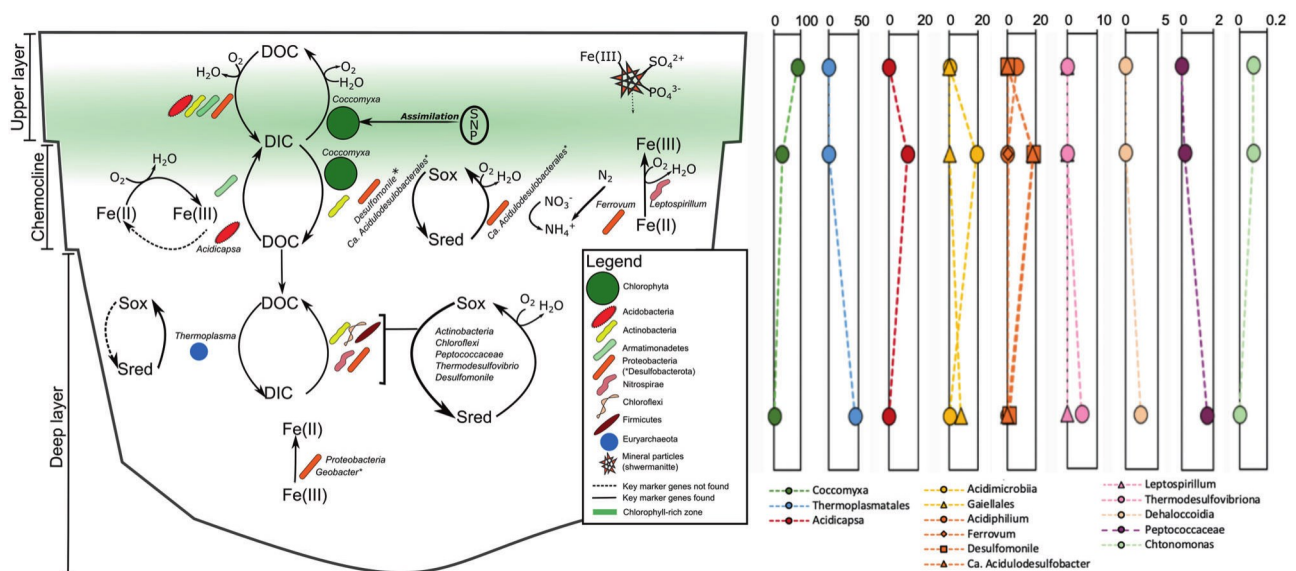


Fig. 6 Conceptual network of the major microbial taxa predicted to be actively mediating carbon, sulfur, iron, and nitrogen cycling in the upper layer, chemocline, and deep layer of Cueva de la Mora (center). Physicochemical profiles displayed on left-side, and relative abundance (% based on 18 S rRNA gene for *Coccomyxa* and 16S rRNA gene for archaea and bacteria) profiles of important taxa displayed on right-side. Figure not drawn to scale.

upper layers were transported with sinking minerals supporting the growth of heterotrophic populations such as *Thermoplasmatales*, the most abundant taxa in the deep layer.

Active Fe(III) reduction was associated with low abundant *Geobacter* spp., and no Fe(II) oxidation was detected in the deep layer. Activity for sulfur compounds oxidation was not expected to happen in the anoxic deep layer. However, our thermodynamic calculations showed that sulfide oxidation becomes thermodynamically feasible ($\sim 20 \text{ kJ mol}^{-1}$) even at low dissolved oxygen concentrations ($< 1 \times 10^{-5} \text{ M}$). Accordingly, sulfur compounds oxidation was predicted for unknown active and abundant populations from Euryarchaeota, Actinobacteria, Chloroflexi, Firmicutes, Nitrospirae, and Proteobacteria. Among these phyla, Actinobacteria, Chloroflexi, and Nitrospirae were the most abundant and had high gene expression for reactions in reductive sulfur metabolisms.

Potential for bioremediation

These integrated findings shed new light on remediation possibilities. Our results reveal how microbial activities in the deep layer of CM could be exploited for the remediation of acidic pit lakes. Our data show that sulfide production in the deep layer, where toxic metal(oid)s could best be sequestered and/or recovered by precipitation in sulfide minerals, is not limited by genetic potential. Instead, sulfide production is potentially limited by episodic oxygen addition from groundwater that allows for sulfide/sulfur oxidation and by the quantity and quality of organic carbon reaching the deep layer from primary producers in the overlying water column. Increasing photosynthetic primary production by adding limiting nutrients such as P and N to the upper layer (i.e., artificial eutrophication as an indirect strategy to deliver C and nutrients to the deep layer) or direct additions of electron donors to the deep layer should enhance biosulfidogenesis and metal(oid) sequestration in the most highly polluted layer of the lake.

A simple framework for the in situ remediation of the deep layer of acidic pit lakes would be to promote all processes that produce sulfide and inhibit all processes that oxidize reduced S species. Reoxidation of metal sulfides (targeted formation for metal sequestration) could occur by any oxygen influx (abiotic process) or by biological S oxidation. Both processes could be inhibited by

delivering sufficient organic C into the deep layer. Because oxygen is an energetically favorable electron acceptor, preferential consumption of oxygen can be promoted by organic C addition. Because organic C is a more energetically favorable electron donor compared to Fe(II) or reduced S species, reoxidation of metal sulfides can be inhibited by organic C addition. More organic C could be fixed into the water column by nutrient addition to stimulate phototrophic microbes or by addition of a C substrate such as glycerol.

There are some uncertainties with fertilizing the upper layer to promote primary production of organic carbon by phototrophic microbes in the upper layer. Based on concentrations of N and P in the upper layer and chemocline (Fig. 1), the system is likely limited by P (because N/P molar ratios are 32 and 44, respectively). However, sorption of P to abundant Fe(III) minerals (schwertmanite and jarosite) in the upper layer may control the concentration of dissolved and bioavailable P. If the concentration of bioavailable P cannot be substantially increased by fertilizer addition, organic C fixed by phototrophs may not substantially increase. Assuming that fertilization can increase bioavailable P leading to a corresponding increase in organic C in the upper layer, it is unknown how much of this organic C will be delivered into the deep layer. As shown in this study, various metabolisms are particularly active in the chemocline (Figs. 3–5). Increased phototrophic growth in the upper layer may further limit light penetration and disrupt low-light phototrophic activity in the chemocline. Less primary production of organic C in the chemocline could affect other metabolisms in this layer. Yet more transport of organic C from the upper layer could increase heterotrophic activity in the chemocline such that the quantity and quality of organic C reaching the deep layer is not substantially changed by fertilization. We speculate that fertilization of the upper layer should increase organic C in the deep layer but are uncertain of the magnitude of the increase.

Another option would be to add a C substrate such as glycerol to the upper layer or via pressurized subaqueous injection directly into the deep layer. Surface addition of glycerol would be simpler and less costly as compared to subaqueous injection. However, it is uncertain how much of the glycerol added to the upper layer (or biomass produced via glycerol utilization) would be transported to the deep layer (similar to issues with upper-layer fertilization).

Because N and P concentrations are lower in the upper layer (Fig. 1), the extent of glycerol utilization in the upper layer is uncertain. Less glycerol utilization in the upper layer should allow more diffusive transport of glycerol into the chemocline and deep layer. Because N and P concentrations are higher in the deep layer, pressurized subaqueous injection of only glycerol may be enough to stimulate biosulfidogenesis and limit sulfide reoxidation.

These different approaches to remediate the lake are based on the understanding of the microbial ecology and geochemistry of the lake. Future work will focus on using a combination of culture-dependent and culture-independent studies boosting biosulfidogenesis in the deep layer of Cueva de la Mora.

MATERIALS AND METHODS

Sample collection

Three biological replicates within the 2- μ m to 0.22- μ m size fraction were sampled in the water column of Cueva de la Mora at depths representing the upper layer (3-m, CM03), chemocline (11-m, CM11) and deep layer (35-m, CM35). In May 2018, water samples were collected in CM for meta-omics analysis. Physicochemical profiling and sample collection were previously described in Ayala-Muñoz et al. [18]. Briefly, physicochemical profiles were acquired with a Hach MS5 datasonde (Loveland, CO) and confirmed the physical and chemical stability of CM with respect to many previous sampling campaigns since 2007. Water samples were collected at 3-m (upper layer), 11-m (chemocline), and 35-m (deep layer) with a 5-L Van Dorn limnological horizontal sampling bottle (KC Denmark A/S, Silkeborg, Denmark). Three replicate sample bottles were collected per depth. On shore, water was immediately prefiltered through a 2- μ m pore size glass fiber filter and subsequently filtered through 0.22- μ m polyethersulfone (PES) Sterivex filters. At least two Sterivex filters were collected per replicate, one for DNA extraction (up to 3-L of water filtered), and one for RNA extraction (up to 1-L of water filtered). The Sterivex filters were immediately frozen on dry ice and cryo-shipped to the US where they were stored at -80°C until analysis. All material used for sampling was RNase free or cleaned with RNase AWAY prior to each water collection and filtering. Water for RNA extraction was filtered in less than 30 min after collection to minimize the risk of degradation.

DNA and RNA extraction

Sterivex filters were processed using the DNAeasy PowerWater Kit and the RNeasy PowerMicrobiome Kit (Qiagen, Venlo, Netherlands) for DNA and RNA extraction, respectively. DNA extraction was conducted following the manufacturer's instructions with a modification to include extractions from difficult-to-lyse cells. One Sterivex filter per replicate per sample was selected and thawed in ice. The plastic capsule of each Sterivex filter was cracked using pliers (without touching the filter). The tip of the capsule was taken out of the base so that the filter came out with the tip. Once the filter was out and while holding the tip with the plier in one side, the filter was cut off in two pieces using a sterile exacto knife. Using sterile straight fine point tweezers, each half of the filter was taken off the base and cut in two pieces that were placed in a bead beating tube to proceed with the DNA extraction. A separate lysis step which involved heating at 65°C for 10 min followed by 5-minute vortexing was carried out in a half of each filter. The remaining half was subjected to only 5-minute vortexing. Two extracts per filter were obtained and pooled together to provide one DNA extract per replicate per sample. Extraction blanks were included following the described steps but using an unused Sterivex filter. RNA extractions were conducted according to the manufacturer's instruction. All extractions were performed in a laminar-flow hood to minimize risk of contamination. DNA extracts were quantified using the Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). RNA extracts were quantified using both the Qubit RNA high sensitivity Fluorometer (Invitrogen, Carlsbad, CA, USA) as well as a Bioanalyzer 2100 RNA 6000 Pico Assay (Agilent, Santa Clara, CA, USA).

Library preparation and sequencing

Two biological replicates per depth with the highest DNA/RNA concentration and quality were selected for library preparation and sequencing. With the DNA extracts from samples from the three depths, we carried out 16S rRNA gene and metagenomic sequencing, and with the RNA extracts, we carried out metatranscriptomic sequencing. DNA extracts were also used for 18S rRNA gene sequencing for the upper layer and chemocline.

The Earth Microbiome 16S and 18S rRNAs protocols were applied for DNA-PCR amplification [72]. For 16S rRNA genes, PCR amplifications were set up with 11.375 μL sterile water, 10.625 μL of Ex Taq master mix with 806 R reverse primer (0.2 μM final concentration), 2 μL of previously extracted DNA (serial dilutions also used for troublesome samples), and 1 μL of 515 F forward barcoded primer (0.2 μM final concentration). For 18S rRNA genes, PCR amplifications were set up with 11.375 μL sterile water, 10.625 μL of Ex Taq master mix with Eukbr reverse primer (0.2 μM final concentration), 2 μL of previously extracted DNA (serial dilutions also used for troublesome samples), and 1 μL of 1391F forward primer (0.2 μM final concentration). The V4 region of the 16S rRNA gene and V9 region of the 18S rRNA gene were amplified using the following thermal cycler conditions: 3-minute denaturation at 94°C ; then 35 cycles at 94°C for 45 s, annealing at 50°C (16S rRNA gene) and 57°C (18S rRNA gene) for 60 s, and extension at 72°C for 90 s; followed by a 10-minute hold at 72°C for final extension of amplicons. Reactions were then checked on a 2% agarose gel (ThermoFisher, Massachusetts, USA) for successful amplicons at ~ 390 bp (16S rRNA gene) and ~ 250 (± 50) bp (18S rRNA gene). Following completion of PCR, each PCR product were pooled together in approximately equimolar amounts. The pooled sample was then loaded on a 2% agarose gel for purification to remove primer dimers and other non-specific amplified DNA. Bands at 390 bp (16S rRNA gene) and between 200–300 bp (18S rRNA gene) were then isolated from the gel and purified with the QIAquick Gel extraction kit (Qiagen, Venlo, Netherlands) to recover pure DNA. Finally, the purified library was quantified using a Qubit 2.0 fluorometer with the High Sensitivity double-stranded DNA assay (Invitrogen, Carlsbad, CA, USA). Libraries were multiplexed. This final multiplexed library was sent for MiSeq (Illumina, San Diego, California, USA) 2 $\times$ 250 sequencing using v2 chemistry.

The metagenomics and metatranscriptomics sequencing were previously described in Ayala-Muñoz et al. [18]. Briefly, preparation of the metagenome libraries was performed using the Illumina's NexteraXT library preparation kit (Illumina, San Diego, California, USA). The metatranscriptome library preparation was performed using the Tecan RNA Trio library preparation kit (Tecan, Männedorf, Switzerland) without a ribosomal RNA depletion step. Sequencing was conducted on an HiSeq 4000 platform (Illumina, San Diego, California, USA) using 150 bp paired end chemistry. Two metagenomic and two metatranscriptomic datasets were obtained per sampling depth. General statistics of these datasets are provided in Table S3.

Bioinformatic analyses

The paired-ended V4 (16S rRNA gene) and V9 (18S rRNA gene) amplicon sequences were bioinformatically processed using the Quantitative Insights Into Microbial Ecology 2 program (QIIME2 v2018.8, <https://qiime2.org/>). Quality filtering, trimming, de-noising, and merging of sequences were carried out with DADA2 [73]. The consensus method in DADA2 was used to identify and remove chimeric sequences. A feature-classifier (sklearn) in QIIME2 was utilized for taxonomic classification of all ribosomal sequence variants. The reference databases used to train the classifier were from SILVA release 132 for both 16S and 18S rRNA genes [74]. Mitochondrial sequences and chloroplasts were removed from the analysis. The core-diversity plugin within QIIME2 calculated diversity metrics with a resulting Amplicon Sequence Variant or ASV table (Table S1). A beta diversity analysis including only prokaryotic (16S rRNA gene) ASVs was conducted by Bray Curtis metrics based on pairwise distance values using R and the rarefied ASV table.

The bioinformatic analyses with respect to DNA and RNA reads preprocessing, metagenome assembly, gene calling, and reads mapping were described previously in Ayala-Muñoz et al. [18]. Briefly, trimmomatic v0.36 [75] was used for quality-filtering of raw reads, and Nonpareil 3 [76] was used to estimate the metagenomic coverage. Quality-filtered metagenomic reads were co-assembled with Megahit v1.1.2 with default parameters [77]. Eukaryotic contigs from co-assembled metagenomes were selected with EukRep [78] and genes were predicted with MetaEuk [79] using the NCBI-nr database (downloaded May 2020). Prodigal v2.6.3 [80] was used to predict genes from non-eukaryotic contigs in all co-assembled metagenomes. In-silico removal of ribosomal RNAs was conducted with sortmeRNA v2.1 [81] in the quality-filtered metatranscriptomic reads. The mRNAs were mapped to the co-assembled metagenomic contigs with BBMap (min_id = 0.95 and slow mode). Around 50% or more of mRNA reads were mapped to the respective metagenome contigs, except for one metatranscriptome from CM35_2 (<1%) that was not used for downstream analysis. Mapping of DNA and mRNA reads to genes was

conducted with BBmap (min_id = 0.95, slow mode), and TPM (transcripts per million) values per predicted gene were calculated as a proxy for gene and transcript abundance (TPM = number of reads from metagenome or metatranscriptome mapped to one gene ($\times 10^6$) divided by its length/sum of number of reads from metagenome or metatranscriptome mapped to all genes divided by respective lengths).

Functional annotation of genes was conducted with GhostKOALA [82], KOFAMScan v1.3.0 [83] and FeGenie [36]. KOFAM functional annotations with an E-value $< 10^{-4}$ and a matching KO with those from the GhostKOALA annotations were chosen for downstream analysis. For nutrient cycling analysis, we selected the annotated genes involved in carbon, sulfur, oxygen, and nitrogen cycling using a list of marker genes with assigned KEGG Orthology identifier -KOs (Table S2). Finally, for annotation of Fe cycling genes, we used FeGenie with default options.

Taxonomic annotation of the predicted genes was conducted with three different bioinformatic tools. First, genes were annotated with DIAMOND v0.9.32.133 [84] against the NCBI-nr database (downloaded in May 2020) using 50% identity over at least 80% length as the annotation threshold. The top hits per gene were visualized with Megan v6.18.10 [85]. Second, genes were taxonomically annotated with the Contig Annotation Tool-CAT [86], using the CAT_prepare_20200618 database files and flag -top 11. Third, genes were annotated with GhostKOALA [82] in the web server option. The first approach (DIAMOND + Megan) assigned between 55 to 70% at the phylum level, 51 to 67% at the class level, 43 to 62% at the order level, 36 to 57% at the family level, and 22 to 48% at the genus level of the genes with a functional annotation. The second approach (CAT) assigned between 14 to 55% at the phylum level of the functionally annotated genes. Between 6 to 40% of the genes were assigned to other lower taxonomic levels by CAT. GhostKOALA assigned ~98% of the genes at the phylum level. The final taxonomic assignment for each gene was reconciled based on agreement between at least two of the three approaches at the phylum level. We also kept the taxonomic annotation at the phylum level of genes when assigned only by one of the three methods. In total, around 98% of the genes with a functional annotation were assigned at the phylum level (Table S3).

Although the metagenomic and metatranscriptomic results represent the specific conditions at the time of sampling, they can be interpreted as the overall contributions of microbial populations to elemental cycling in diurnal conditions given the stability and maturity of the lake.

Thermodynamic calculations

The bioenergetic favorability of various metabolic processes occurring under the measured conditions in each layer of the lake were evaluated by thermodynamic calculations. A script in R (https://github.com/DianaKarina/DeltaG_calculations/blob/main/energetics-calculations.R) using the package CHNOSZ [87] was utilized to calculate the non-standard state Gibbs free energies (ΔG in kJ mol^{-1}) of the metabolic processes potentially occurring in each layer. The script calculates the Gibbs of formation based on standard conditions and the Gibbs of reaction considering the actual conditions of the lake to get ΔG . The script requires as input the species and their respective activities, followed by the balanced reactions of interest. Activities of species per layer were calculated in PHREEQC based on molar concentrations (Table S7).

DATA AVAILABILITY

Raw data for amplicons, metagenomes, and metatranscriptomes are available in the SRA database as bioproject PRJNA646106.

REFERENCES

- Schultze M, Boehrer B, Wendt-Potthoff K, Sánchez-España J, Castendyk D, Meromictic Pit Lakes: Case Studies from Spain, Germany and Canada and General Aspects of Management and Modelling. In: Gulati RD, Zadereev ES, Degermendzhi AG, editors. *Ecology of Meromictic Lakes*. Cham: Springer International Publishing; 2017. p. 235–75.
- Blanchette ML, Lund MA. Pit lakes are a global legacy of mining: an integrated approach to achieving sustainable ecosystems and value for communities. *Curr Opin Environ Sustain*. 2016;23:28–34.
- Whitton BA. Acidic pit lakes—The legacy of coal and metal surface mines. *J Limnol*. 2013;72:1.
- Schultze M, Limnology of Pit Lakes. In: Geller W, Schultze M, Kleinmann R, Wolkersdorfer C, editors. *Acidic Pit Lakes: The Legacy of Coal and Metal Surface Mines*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 23–224.
- Grettenberger CL, McCauley Rench RL, Gruen DS, Mills DB, Carney C, Brainard J, et al. Microbial population structure in a stratified, acidic pit lake in the Iberian Pyrite Belt. *Geomicrobiol J*. 2020;37:623–34.
- Santofimia E, González-Toril E, López-Pamo E, Gomariz M, Amils R, Aguilera Á. Microbial diversity and its relationship to physicochemical characteristics of the water in two extreme acidic pit lakes from the Iberian Pyrite Belt (SW Spain). *Plos One*. 2013;8:e66746.
- Falagán C, Sánchez-España J, Johnson DB. New insights into the biogeochemistry of extremely acidic environments revealed by a combined cultivation-based and culture-independent study of two stratified pit lakes. *FEMS Microbiol Ecol*. 2014;87:231–43.
- Van der Graaf CM, Sánchez-España J, Yusta I, Ilin A, Shetty SA, Bale NJ, et al. Biosulfidogenesis mediates natural attenuation in acidic mine pit lakes. *Microorganisms* 2020;8:1275.
- Sánchez-España J, Pamo EL, Díez M, Santofimia E. Physico-chemical gradients and meromictic stratification in Cueva de la Mora and other acidic pit lakes of the Iberian Pyrite Belt. *Mine Water Environ*. 2009;28:15–29.
- Wendt-Potthoff K, Koschorreck M, Díez Ercilla M, Sánchez, España J. Microbial activity and biogeochemical cycling in a nutrient-rich meromictic acid pit lake. *Limnologia - Ecol Manag Inland Waters*. 2012;42:175–88.
- Sánchez-España J, Díez Ercilla M, Pérez Cerdán F, Yusta I, Boyce AJ. Hydrological investigation of a multi-stratified pit lake using radioactive and stable isotopes combined with hydrometric monitoring. *J Hydrol*. 2014;511:494–508.
- Sánchez-España J, Pamo EL, Pastor ES, Ercilla MD. The acidic mine pit lakes of the Iberian Pyrite Belt: An approach to their physical limnology and hydrogeochemistry. *Appl Geochem*. 2008;23:1260–87.
- Bigham JM, Nordstrom DK. Iron and aluminum hydroxysulfates from acid sulfate waters. *Rev Mineral Geochem*. 2000;40:351–403.
- Sánchez-España J, Yusta I, López GA. Schwertmannite to jarosite conversion in the water column of an acidic mine pit lake. *Mineral Mag*. 2012;76:2659–82.
- Sánchez-España J, Yusta I, Díez-Ercilla M. Schwertmannite and hydrobasaluminite: A re-evaluation of their solubility and control on the iron and aluminium concentration in acidic pit lakes. *Appl Geochem*. 2011;26:1752–74.
- Cameron D, Willett M, Hammer L. Distribution of organic carbon in the Berkeley Pit Lake, Butte, Montana. *Mine Water Environ*. 2006;25:93–9.
- Spiteri C, Cappellen PV, Regnier P. Surface complexation effects on phosphate adsorption to ferric iron oxyhydroxides along pH and salinity gradients in estuaries and coastal aquifers. *Geochimica et Cosmochimica Acta*. 2008;72:3431–45.
- Ayala-Muñoz D, Burgos WD, Sánchez-España J, Couradeau E, Falagán C, Macalady JL. Metagenomic and metatranscriptomic study of microbial metal resistance in an acidic pit lake. *Microorganisms* 2020;8:1350.
- Falagán C, Geomicrobiology of Meromictic Metal-mine pit lakes in the Iberian Pyrite belt and biotechnological applications [Doctoral dissertation]: Universidad del País Vasco; 2015.
- Sánchez-España J, Falagán C, Ayala D, Wendt-Potthoff K. Adaptation of *Coccomyxa* sp. to extremely low light conditions causes deep chlorophyll and oxygen maxima in acidic pit lakes. *Microorganisms* 2020;8:1218.
- Miller CS, Baker BJ, Thomas BC, Singer SW, Banfield JF. EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol*. 2011;12:R44.
- Tan S, Liu J, Fang Y, Hedlund BP, Lian Z-H, Huang L-Y, et al. Insights into ecological role of a new deltaproteobacterial order Candidatus Acidulodesulfobacterales by metagenomics and metatranscriptomics. *ISME J*. 2019;13:2044–57.
- DeWeerd KA, Mandelco L, Tanner RS, Woese CR, Suflita JM. Desulfomonile tiedjei gen. nov. and sp. nov., a novel anaerobic, dehalogenating, sulfate-reducing bacterium. *Arch Microbiol*. 1990;154:23–30.
- Lu S, Chourey K, Reiche M, Nietzsche S, Shah MB, Neu TR, et al. Insights into the structure and metabolic function of microbes that shape pelagic iron-rich aggregates (“iron snow”). *Appl Environ Microbiol*. 2013;79:4272–81.
- Volant A, Bruneel O, Desoeuvre A, Héry M, Casiot C, Bru N, et al. Diversity and spatiotemporal dynamics of bacterial communities: physicochemical and other drivers along an acid mine drainage. *FEMS Microbiol Ecol*. 2014;90:247–63.
- Mesa V, Gallego JLR, González-Gil R, Lauga B, Sánchez J, Méndez-García C, et al. Bacterial, archaeal, and eukaryotic diversity across distinct microhabitats in an acid mine drainage. *Front Microbiol*. 2017;8:1756.
- Lee KCY, Morgan XC, Dunfield PF, Tamas I, McDonald IR, Stott MB. Genomic analysis of *Chthonomonas calidirosea*, the first sequenced isolate of the phylum Armatimonadetes. *The ISME J*. 2014;8:1522–33.
- Holanda R, Hedrich S, Nancucio I, Oliveira G, Grail BM, Johnson DB. Isolation and characterisation of mineral-oxidising “*Acidibacillus*” spp. from mine sites and geothermal environments in different global locations. *Res Microbiol*. 2016;167:613–23.
- Sánchez-Andrea I, Stams AJM, Amils R, Sanz JL. Enrichment and isolation of acidophilic sulfate-reducing bacteria from Tinto River sediments. *Environ Microbiol Rep*. 2013;5:672–8.

30. Senko JM, Zhang G, McDonough JT, Bruns MA, Burgos WD. Metal reduction at low pH by a desulfosporosinus species: implications for the biological treatment of acidic mine drainage. *Geomicrobiol J*. 2009;26:71–82.
31. Méndez-García C, Peláez AI, Mesa V, Sánchez J, Golyshina OV, Ferrer M. Microbial diversity and metabolic networks in acid mine drainage habitats. *Front Microbiol*. 2015;6:475.
32. García-Moyano A, Austnes AE, Lanzén A, González-Toril E, Aguilera Á, Øvreås L. Novel and unexpected microbial diversity in acid mine drainage in Svalbard (78° N), revealed by culture-independent approaches. *Microorganisms*. 2015;3:667–94.
33. Gavrilov SN, Korzhnikov AA, Kublanov IV, Bargiela R, Zamana LV, Popova AA, et al. Microbial communities of polymetallic deposits' acidic ecosystems of continental climatic zone with high temperature contrasts. *Front Microbiol*. 2019;10:1573.
34. Pereira IAC, Ramos A, Grein F, Marques M, Da Silva S, Venceslau S, A comparative genomic analysis of energy metabolism in sulfate reducing bacteria and archaea. *Front Microbiol*. 2011;2:69.
35. Fuentes JL, Huss VAR, Montero Z, Torronteras R, Cuaresma M, Garbayo I, et al. Phylogenetic characterization and morphological and physiological aspects of a novel acidotolerant and halotolerant microalga *Coccomyxa onubensis* sp. nov. (Chlorophyta, Trebouxiophyceae). *J Appl Phycol*. 2016;28:3269–79.
36. Garber AI, Nealson KH, Okamoto A, McAllister SM, Chan CS, Barco RA, et al. FeGenie: a comprehensive tool for the identification of iron genes and iron gene neighborhoods in genome and metagenome assemblies. *Front Microbiol*. 2020;11:37.
37. Peng H, Wei D, Chen G, Chen F. Transcriptome analysis reveals global regulation in response to CO₂ supplementation in oleaginous microalga *Coccomyxa subellipsoidea* C-169. *Biotechnol Biofuels*. 2016;9:151.
38. Díez Ercilla M, López Pamo E, Sánchez España J. Photoreduction of Fe(III) in the acidic mine pit lake of San Telmo (Iberian Pyrite Belt): Field and experimental work. *Aquat Geochem*. 2009;15:391–419.
39. Barcytė D, Nedbalová L. *Coccomyxa*: a dominant planktic alga in two acid lakes of different origin. *Extremophiles*. 2017;21:245–57.
40. Norris PR, Davis-Belmar CS, Brown CF, Calvo-Bado LA. Autotrophic, sulfur-oxidizing actinobacteria in acidic environments. *Extremophiles*. 2011;15:155–63.
41. Johnson DB, Hallberg KB, Hedrich S. Uncovering a microbial enigma: isolation and characterization of the streamer-generating, iron-oxidizing, acidophilic bacterium “*Ferrovum myxofaciens*”. *Applied. Environ Microbiol*. 2014;80:672–80.
42. Lee KC, Stott MB, Dunfield PF, Huttenhower C, McDonald IR, Morgan XC. The *Chthonomonas calidirosea* genome is highly conserved across geographic locations and distinct chemical and microbial environments in New Zealand's Taupō Volcanic Zone. *Appl Environ Microbiol*. 2016;82:3572–81.
43. Johnson DB, Hallberg KB. Carbon, iron and sulfur metabolism in acidophilic micro-organisms. *Adv Microb Physiol*. 2008;54:201–55.
44. Golyshina OV, Kublanov IV, Tran H, Korzhnikov AA, Lünsdorf H, Nechitaylo TY, et al. Biology of archaea from a novel family *Cuniculiplasmataceae* (Thermoplasmata) ubiquitous in hyper acidic environments. *Sci Rep*. 2016;6:39034.
45. Yelton AP, Comolli LR, Justice NB, Castelle C, Denef VJ, Thomas BC, et al. Comparative genomics in acid mine drainage biofilm communities reveals metabolic and structural differentiation of co-occurring archaea. *BMC Genom*. 2013;14:485.
46. Korzhnikov AA, Toshchakov SV, Bargiela R, Gibbard H, Ferrer M, Teplyuk AV, et al. Archaea dominate the microbial community in an ecosystem with low-to-moderate temperature and extreme acidity. *Microbiome*. 2019;7:11.
47. Hug LA, Castelle CJ, Wrighton KC, Thomas BC, Sharon I, Frischkorn KR, et al. Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. *Microbiome*. 2013;1:22.
48. Watanabe M, Kojima H, Fukui M. Review of *Desulfotomaculum* species and proposal of the genera *Desulfallus* gen. nov., *Desulfofundulus* gen. nov., *Desulfofarcimen* gen. nov. and *Desulfohalotomaculum* gen. nov. *Int J Syst Evolut Microbiol*. 2018;68:2891–9.
49. Lehours A-C, Borrel G, Morel-Desrosiers N, Bardot C, Grossi V, Keraval B, et al. Anaerobic microbial communities and processes involved in the methane cycle in freshwater lakes—a focus on Lake Pavin. In: Sime-Ngando T, Boivin P, Chapron E, Jezequel D, Meybeck M, editors. *Lake Pavin: History, geology, biogeochemistry, and sedimentology of a deep meromictic maar lake*. Cham: Springer International Publishing; 2016. p. 255–84.
50. Meyer O, Fiebig K, editors. *Enzymes Oxidizing Carbon Monoxide* 1985; Dordrecht: Springer Netherlands.
51. Tejada-Jimenez M, Chamizo-Ampudia A, Llamas A, Galvan A, Fernandez E, Chapter 8—Roles of Molybdenum in plants and improvement of its acquisition and use efficiency. In: Hossain MA, Kamiya T, Burritt DJ, Phan Tran L-S, Fujiwara T, editors. *Plant Micronutrient Use Efficiency*: Academic Press; 2018. p. 137–59.
52. Zhang L, Liu X, Liu J, Zhang Z. Characteristics and function of sulfur dioxygenase in *Echiurum worm Urechis unicinctus*. *PLoS One*. 2013;8:e81885-e.
53. Díez-Ercilla M, Sánchez-España J, Yusta I, Wendt-Potthoff K, Koschorreck M. Formation of biogenic sulphides in the water column of an acidic pit lake: biogeochemical controls and effects on trace metal dynamics. *Biogeochemistry*. 2014;121:519–36.
54. Anantharaman K, Hausmann B, Jungbluth SP, Kantor RS, Lavy A, Warren LA, et al. Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. *ISME J*. 2018;12:1715–28.
55. van Vliet DM, von Meijenfildt FAB, Dutilh BE, Villanueva L, Sinninghe Damsté JS, Stams AJM, et al. The bacterial sulfur cycle in expanding dysoxic and euxinic marine waters. *Environ Microbiol*. 2021;23:2834–57.
56. Sekiguchi Y, Muramatsu M, Imachi H, Narihiro T, Ohashi A, Harada H, et al. *Thermodesulfovibrio aggregans* sp. nov. and *Thermodesulfovibrio thiophilus* sp. nov., anaerobic, thermophilic, sulfate-reducing bacteria isolated from thermophilic methanogenic sludge, and emended description of the genus *Thermodesulfovibrio*. *Int J Syst Evolut Microbiol*. 2008;58:2541–8.
57. Oude Elferink SJWH, Akkermans-van Vliet WM, Bogte JJ, Stams AJM. *Desulfobacca acetoxidans* gen. nov., sp. nov., a novel acetate-degrading sulfate reducer isolated from sulfidogenic granular sludge. *Int J Syst Evolut Microbiol*. 1999;49:345–50.
58. Davidova IA, Duncan KE, Choi OK, Suflija JM. *Desulfoglaeba alkanexedens* gen. nov., sp. nov., an n-alkane-degrading, sulfate-reducing bacterium. *Int J Syst Evolut Microbiol*. 2006;56(Pt 12):2737–42.
59. Umezawa K, Watanabe T, Miura A, Kojima H, Fukui M. The complete genome sequences of sulfur-oxidizing Gammaproteobacteria *Sulfurifustis variabilis* skN76(T) and *Sulfuricaulis limicola* HAS(T). *Standard Genomic. Science* 2016;11:71.
60. Inceoglu Ö, Lirós M, Crowe SA, García-Armisen T, Morana C, Darchambeau F, et al. Vertical distribution of functional potential and active microbial communities in meromictic lake Kivu. *Microb Ecol*. 2015;70:596–611.
61. Segerer A, Langworthy TA, Stetter KO. *Thermoplasma acidophilum* and *Thermoplasma volcanium* sp. nov. from Solfataria Fields. *Syst Appl Microbiol*. 1988;10:161–71.
62. Ruepp A, Graml W, Santos-Martinez M-L, Koretke KK, Volker C, Mewes HW, et al. The genome sequence of the thermophilic scavenger *Thermoplasma acidophilum*. *Nature*. 2000;407:508–13.
63. Stolper DA, Revsbech NP, Canfield DE. Aerobic growth at nanomolar oxygen concentrations. *Proc Natl Acad Sci* 2010;107:18755–60.
64. Slobodkina GB, Mardanov AV, Ravin NV, Frolova AA, Chernyh NA, Bonch-Osmolovskaya EA, et al. Respiratory ammonification of nitrate coupled to anaerobic oxidation of elemental sulfur in deep-sea autotrophic thermophilic bacteria. *Front Microbiol*. 2017;8:87-.
65. Brock TD, Gustafson J. Ferric iron reduction by sulfur- and iron-oxidizing bacteria. *Appl Environ Microbiol*. 1976;32:567–71.
66. Ilbert M, Bonnefoy V. Insight into the evolution of the iron oxidation pathways. *Biochimica et Biophysica Acta (BBA) - Bioenerg*. 2013;1827:161–75.
67. Grettenberger CL, Havig JR, Hamilton TL. Metabolic diversity and co-occurrence of multiple *Ferrovum* species at an acid mine drainage site. *BMC Microbiol*. 2020;20:119.
68. Ojumu TV, Petersen J, Hansford GS. The effect of dissolved cations on microbial ferrous-iron oxidation by *Leptospirillum ferriphilum* in continuous culture. *Hydrometallurgy*. 2008;94:69–76.
69. Hua Z-S, Han Y-J, Chen L-X, Liu J, Hu M, Li S-J, et al. Ecological roles of dominant and rare prokaryotes in acid mine drainage revealed by metagenomics and metatranscriptomics. *ISME J*. 2014;9:1280.
70. Casal C, Cuaresma M, Vega JM, Vilchez C. Enhanced productivity of a lutein-enriched novel acidophile microalga grown on urea. *Mar Drugs*. 2011;9:29–42.
71. Albuquerque L, Costa MSd. The Family Gaiellaceae. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. *The Prokaryotes*. Berlin, Heidelberg: Springer; 2014.
72. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locay KJ, et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*. 2017;551:457–63.
73. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:581.
74. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41(D1):D590–D6.
75. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
76. Rodríguez-R LM, Konstantinidis KT. Nonpareil: a redundancy-based approach to assess the level of coverage in metagenomic datasets. *Bioinformatics*. 2013;30:629–35.
77. Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, Sadakane K, et al. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*. 2016;102:3–11.
78. West PT, Probst AJ, Grigoriev IV, Thomas BC, Banfield JF. Genome-reconstruction for eukaryotes from complex natural microbial communities. *Genome Res*. 2018;28:569–80.

79. Levy Karin E, Mirdita M, Söding J. MetaEuk—sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics. *Microbiome*. 2020;8:48.
80. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinforma*. 2010;11:119.
81. Kopylova E, Noe L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 2012;28:3211–7.
82. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol*. 2016;428:726–31.
83. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 2019;36:2251–2.
84. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;12:59–60.
85. Huson DH, Beier S, Flade I, Górski A, El-Hadidi M, Mitra S, et al. MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. *PLOS Comput Biol*. 2016;12:e1004957.
86. von Meijenfeldt FAB, Arkhipova K, Cambuy DD, Coutinho FH, Dutilh BE. Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biol*. 2019;20:217.
87. Dick JM, CHNOSZ: Thermodynamic Calculations and Diagrams for Geochemistry. *Front Earth Sci*. 2019;7:180.

ACKNOWLEDGEMENTS

This research was partially supported by the US National Science Foundation project EAR-2016826 to WB, JM, and JSE, by the Spanish Ministry of Economy, Industry, and Competitiveness (MINECO) grant number CGL2016-74984-R to JSE, and by a scholarship from Fulbright-Ecuador to D.A.M.

AUTHOR CONTRIBUTIONS

DA-M, WB, and JM contributed to the conceptualization of this work, data interpretation, and continuous review, discussion, and feedback. JS-E and CF contributed to sampling and geochemical and microbiological characterization. EC contributed to data analysis, interpretation, and discussion of the results. All contributed to reviewing and editing the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41396-022-01320-w>.

Correspondence and requests for materials should be addressed to Diana Ayala-Muñoz or William D. Burgos.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.