Microbial ecology of sulfur biogeochemical cycling at a mesothermic hot spring atop Northern Himalayas, India

Shekhar Nagar¹, Chandni Talwar¹, Mikael Motelica-Heino², Hans-Hermann Richnow³, Mallikarjun Shakarad⁴, Rup Lal⁵ and Ram Krishan Negi^{1*}

¹Fish Molecular Biology Laboratory, Department of Zoology, University of Delhi, Delhi 110007, India

²Institut des Sciences de la Terre D'Orleans (ISTO), UMR 7327 CNRS-Université d'Orleans-Brgm, 1A rue de la Férollerie, CEDEX2 45071 Orleans, France

³Department of Isotope Biogeochemistry, Helmholtz Centre for Environmental Research -UFZ, D-04318 Leipzig, Germany

⁴Evolutionary Biology Laboratory, Department of Zoology, University of Delhi, Delhi 110007, India

⁵NASI Senior Scientist Platinum Jubilee Fellow, The Energy and Resources Institute, Darbari Seth Block, IHC Complex, Lodhi Road, New Delhi-110003, India

*Corresponding Author

Email: negigurukul@gmail.com

Department of Zoology, University of Delhi, Delhi 110007, India.

Keywords

Metagenomics; Hot spring; Sediment; Water; Microbial mat; Sulfur Related Bacteria.

Graphical Abstract

Abstract

Sulfur Related Prokaryotes (SRP) residing in hot spring present good opportunity for exploring the limitless possibilities of integral ecosystem processes. Metagenomic analysis further expands the phylogenetic breadth of these extraordinary sulfur metabolizing microorganisms, as well as their complex metabolic networks and syntrophic interactions in environmental biosystems. Through this study, we explored and expanded the microbial genetic repertoire with focus on sulfur cycling genes through metagenomic analysis of sulfur (S) contaminated hot spring, located at the Northern Himalayas. The analysis revealed rich diversity of microbial consortia with established roles in S cycling such as *Pseudomonas*, *Thioalkalivibrio*, *Desulfovibrio* and

Desulfobulbaceae (*Proteobacteria*). The major gene families inferred to be abundant across microbial mat, sediment and water were assigned to *Proteobacteria* as reflected from the RPKs (reads per kilobase) categorized into translation and ribosomal structure and biogenesis. Analysis of sequence similarity showed conserved pattern of both *dsrAB* genes (n=178) retrieved from all metagenomes while other sulfur disproportionation proteins were diverged due to different structural and chemical substrates. The diversity of sulfur oxidizing bacteria (SOB) and sulfate reducing bacteria (SRB) with conserved (r)*dsrAB* suggests for it to be an important adaptation for microbial fitness at this site. Here, we confirm that (i) SRBs belongs to δ -*Proteobacteria* occurring independent LGT of *dsr* genes to different and few novel lineages (ii) also, the oxidative and reductive *dsr* evolutionary time scale phylogeny, proved that the earliest (not first) *dsrAB* proteins belong to anaerobic *Thiobacillus* with other (r)*dsr* oxidizers. Further, the structural prediction of unassigned DsrAB proteins confirmed their relatedness with species of *Desulfovibrio* (TM score= 0.86; 0.98; 0.96) and *Archaeoglobus fulgidus* (TM score= 0.97; 0.98). We proposed that the genetic repertoire might provide the basis of studying time scale evolution and horizontal gene transfer of these genes in biogeochemical S cycling and the complementary genes could be implemented in biotechnology and bioremediation applications.

Introduction

The untapped sulfur compounds oxidizing microorganisms (SOM) and sulfur compounds reducing microorganisms (SRM) microbial communities residing in extreme and contaminated environmental conditions such as hot water, sulfide contaminated springs offer an intriguing opportunity to explore the unique microbial diversity with uncovered metabolic potential (Ghilamical et al., 2017). The investigations of such microbiota began with focus on identifying and culturing novel thermostable biocatalysts with huge biotechnological applications (Inskeep et al., 2010, Li et al., 2014, Ayangbenro and Babalola., 2017). Over the decades, metagenomics studies have revealed the crucial role of abiotic ecology that guides the microbiota at such sites (Lozupone et al., 2007, Inskeep et al., 2013). Evidences of such environmentally adapted microbiota are vast in terms of diversity and its correlation with temperature (Blank et al., 2002; Miller et al., 2009; Reigstad et al., 2010; Ghosh et al., 2012; Cole et al., 2013; Chan et al., 2015). However, little progress has been made in exploring the correlation between microbiome and geochemistry of hot spring systems particularly that possess mesothermic hot waters with neutral

pH and elemental sulfur and sulfate richness (Roy et al., 2020; Ghosh et al., 2012). Moreover, the survival of microbes in these niches is often supported by community dynamics and interactions. Studies of such ecosystems may provide insights into the microbial evolution of specific pathways for microbial biogeochemical cycling of minerals. However, with over 400 thermal hot water springs located in India, less than 15% have been explored for biogeochemical and taxonomical classification using genomics and metagenomics approaches (Cinti et al., 2009; Saxena et al., 2017). Sulfur springs provide harsh physiochemical conditions to sustain the growth of only meso and hyper-thermophilic microbes which includes sulfur oxidizers, sulfate reducers (Chan et al., 2015; Gonsior et al., 2018). The survival could also be achieved with “microorganism adaptation” by several resistance mechanism such as activity of bioprecipitation, biosorption, extracellular sequestration and/or chelation (Haferburg and Kothe 2007). During these changes, the exchange of genetic material by means of horizontal gene transfer (HGT) is prevalent and necessary for the adaptation of microbes through acquisition of novel genes.

Khirganga, the mesothermal sulfur spring in Northern Himalayas discharging waters rich in sulfate, chlorine, sodium and magnesium ions has remained uncharted so far (Shirkot and Verma 2015; Poddar and Das, 2018). High levels of sulfides in the environment accounts for the milky appearance of the hot spring water with white microbial mats predicted to be formed from sulfide reduction by the sulfur related prokaryotes (SRP) enriched at this site (Sharma et al., 2004; Dong et al., 2019). The microbial disproportionation one of the oldest (3.5 billion years ago; Finster 2011) biological processes on Earth producing sulfide, sulfite and sulfate compounds establishes a complex network of pathways in the biogeochemical sulfur cycle. Heretofore, it is the very foremost metagenomic investigation of microbial communities in Khirganga (avg. atmospheric temperature 6.9 ± 0.3) focused on exploring the microbial biogeochemical sulfur cycling with a complex of disproportionation of elemental sulfur conforming intermediary compounds. The current study was carried out via microbial mats, sediments and hot spring water samples in hot spring to decipher the stabilized and diversified genes involved in sulfur cycle intermediary process in sulfate-reducing bacteria and anoxygenic, photolithotrophic and chemolithotrophic sulfur-oxidizing bacteria (Dahl & Truper, 1994; Hipp et al., 1997). The work expands the genetic and evolutionary information for sulfur cycling genes and evaluates the biodiversity and applications for screening of the novel thermostable enzymes

from microorganisms. Further, understanding these adaptations vis-à-vis the physiological properties and metabolic processes in these springs could be monitored as the engineered SRP consortia could develop into an effective tool in optimizing degradation of sewage waste in industrial processes (Ayangbenro et al., 2018). Also, the sulfate-reducing bacteria (SRB) implied to treat various environment contaminants including metals (Zhang et al., 2016, Mothe et al. 2016), metalloids (Bataglia-brunet et al., 2012, Sahinkaya et al. 2015), various non-methane hydrocarbons (Callaghan et al., 2012), alicyclic hydrocarbons (Jaekel et al., 2015), nitroaromatic compounds (Boopathy 2014, Mulla et al., 2014) and aromatic hydrocarbons (Stasik et al., 2015, Meckenstock et al., 2016, Kamarisima et al., 2019).

Material and Methods

Sample collection, Physicochemical analysis and Helium Ion Microscopy

Samples of water (5L), sediment (250 g) and microbial mat deposits (250 g) were collected from Khirganga hot water spring (31°59'34" N, 77°30'35" E) in February, 2017. Sampling was performed in two replicates for each habitat from two closely located primary thermal outlets (31°99'18" N, 77°50'96" E) and secondary outlets (31°99'19" N, 77°50'96" E). The surface temperature and pH of each habitat were recorded on site.

Priorly, water samples were filtrated to 0.1 µm, sediment and microbial mats were digested in pure nitric acid prior to chemical analysis. All samples were subjected to physicochemical analysis for major elements. Concentrations of major cations (Na^+ , K^+ , Mg^{2+} and Ca^{2+}) and anions (SO_4^{2-} , and Cl^-) were analyzed by ionic chromatography (Dionex ICS-2000, Sunnyvale, CA) using the columns CS16A for measuring cations and AS17 for anions. Elemental analysis of minor and trace elements through inductively coupled plasma mass spectrometry (ICP-MS) Agilent ICP-MS 7900 with Ultra High Matrix Introduction (UHMI). Samples of sediments and microbial mats were desiccated overnight followed by ethanolic dehydration and microstructure was studied on scanning electron microscope at the Centre for Chemical Microscopy (ProVIS). Images were captured using high efficiency detector.

Metagenomic DNA Extraction, Sequencing and Assembly

Total community DNA from 5 liters of filtered water (0.45µm) and 0.25g sediment samples were extracted using PowerMax Soil DNA isolation kit (MoBio Laboratories Inc., Carlsbad, CA, USA) following the manufacturer's instructions. For extraction of total DNA from microbial mats, 0.25g samples were processed following a method described by Varin *et al.*, 2010. Sequencing was performed at Beijing Genome Institute (BGI), Hongkong, China using Illumina HiSeq 2500 platform. Paired end libraries of read length 100bp were generated with insert size of 350bp. The raw sequences were quality filtered using SolexaQA (Cox *et al.*, 2010), and the low-quality sequences <Q₂₀ quality cut-off and artificially duplicated reads (ARDs) were castoff using Illumina-Utils (Eren *et al.*, 2013) and duplicate read inferred sequencing error estimation (DRISSE) (Gomez-Alvarez *et al.*, 2009), respectively. Further, the assembly was integrated in IDBA-UD (Peng *et al.*, 2012) with insertion length 50 bp, min. *k-mer*: 31, max. *k-mer*: 93 (61 for water) using seed *k-mer* size for alignment 30 bp and min. size of contig as 200 bp while allowing min. multiplicity for filtering *k-mer* while building the graph.

Taxonomic and Functional Assignments

Alpha diversity within each sample was estimated as abundance-weighted average of annotated species from source databases built in MG-RAST v3.0 (Meyer *et al.*, 2008) and expressed as Shannon diversity index transformed based on rarefaction curve using the following formula:

$$\text{Log}_{10} (\square\text{-diversity}) / \ln (10)$$

Diversity at phylum level was inferred from MG-RAST (maximum e-value, 1×10^5 and minimum percentage identity cutoff, 60%). A paired-sample t-test was applied on the phylum determined in any habitat pair to estimate significant similarities based on taxonomic mean abundance using SPSS (SPSS Inc., version 20.0, IBM). Microbial genera were deciphered based on clade-specific markers to identify taxonomy up to species level using MetaPhlAn v2.0 (Truong *et al.*, 2015) and a heatmap was constructed using Bray-curtis dissimilarity with supporting dendrograms for both species and samples. We used HUMAnN2 (Franzosa *et al.*, 2018) to perform phylum-resolved functional profiling of the communities that maps contigs onto the pangenomes of the known species of the community and quantifies the pathways and uniref90 gene families database (Suzek *et al.*, 2015). Later, these UniRef90 families were regrouped as COG annotations based on eggNOG (Huerta-Cepas *et al.*, 2017). Open Reading

Frames (ORFs) of assembled metagenome were predicted using Prodigal v2.6.1 (Hyatt *et al.*, 2010) and annotated at hierarchy levels *viz.* subsystems, protein families and individual enzymes using Prokka v1.12 (Seemann, 2014). The amino acid sequences were mapped against KEGG database (Kanehisa *et al.*, 2004) and top 50 metabolic pathways in all of the six samples were compared through heatmap constructed using package ‘*pheatmap*’ (Kolde and Kolde, 2015) and ‘*ggplot2*’ (Wickham, 2009) in R (R Development Core team). Identification of S substrates disproportionation genes was performed by mapping all predicted ORFs on the HMM databases obtained from TIGRfam v10 (Haft *et al.*, 2012) and Pfam (Finn *et al.*, 2014) using hmmscan v3.1b2 (Eddy, 2011). Abundance of each enzyme was plotted as number of copies annotated within each sample. The sequences with >150 amino acids were queried against the NCBI Microbial proteins from RefSeq *nr* database (4th April, 2020) using Blastp (Altschul *et al.*, 1990) to identify the sequences producing significant alignments for taxonomic confirmation.

Analysis of diversity of sulfate reduction proteins

For sequence similarity networks (SSN), amino acid sequences of sulfide oxidation and sulfate reduction proteins annotated in all six samples annotated by KEGG Ids were implied over an empirical measurement of diversity. For this, an all-vs-all BLAST was performed to define the similarities/variations between sequence pairs of diversifying sulfate reduction proteins. A user defined threshold was optimized according to the alignment score and maximum length of BLAST results in diversifying and stabilized protein sequences. Clustering was performed using CD-HIT (Li & Godzik, 2006) on the scores of BLASTp pairwise alignments at a threshold value (e-value of 1e-30). The networks were visualized in Cytoscape v3.7.1 (Shannon *et al.*, 2003). The average number of degree and neighbours for a protein sequence or a node was calculated under:

$$k = \frac{2K}{N}$$

Where K is denoted with number of edges and N is denoted with total number of nodes. Also, to determine the divergence/similarity among nodes or protein sequences was calculated under:

$$D = \frac{2K}{N(N-1)}$$

Attributes of node degree distribution, average clustering coefficient, average neighborhood connectivity and closeness centrality were studied through power law fits to determine their correlation with number of neighbors. Sulfur oxidizing bacteria (SOB) and sulfate reducing bacteria (SRB) were identified for the sequences that could be classified up to genus level to study the distribution of S substrates oxidation and reduction genes in the different clusters.

Sequence alignment, phylogeny and structure prediction of putative unidentified Dsr and Asr enzymes

To elucidate the diversity of key sulphite reductases, the DsrA/B and AsrA/B protein sequences were individually aligned using MUSCLE v3.8.31 (Edgar 2004) and clustered using UPGMB. All the alignments were end trimmed manually and maximum likelihood (ML) phylogeny was inferred with 500 bootstrap resampling using RAXML v8.0.26 (Stamatakis 2014). For this, we used standalone version of RAXML which was called as follows:

```
raxmlHPC-PTHREADS -s input -N 500 -n result -f a -p 12345 -x 12345 -m
PROTGAMMAGTR.
```

The resulting phylogenies were also confirmed using most complex general time-reversible model (GTR; Tavare, 1986) with PhyloBayes v1.7b using CIPRES Science Gateway v 3.3 (Lartillot et al., 2009) that incorporates different rates for every change and different nucleotide frequencies.

For proteins showing similarity with those from uncultured bacteria, we determined the structures using I-TASSER suite (Yang and Zhang, 2015). These predicted structures were then aligned onto their top structural analogs and C-scores, TM-scores and RMSD were computed and ligand binding sites with conserved residues were identified. The TM-score is to compare two models based on their given residue equivalency (i.e., based on the residue index in the PDB file). It is usually NOT applied to compare two proteins of different sequences. The TM-score predicted from structural alignment of two proteins while comparing them based on residue equivalency such that a score of 0.6 and above denote the two proteins to be fairly aligned (Yang and Zhang, 2015). The TM-align will first find the best equivalent residues of two proteins based on the structure similarity and then output a TM-score.

Results and Discussion

Description of sampling site and microscopic analysis of samples

The Khirganga is a natural hot spring setting that lies in the Parvati Valley in the Northern hemisphere of the great Himalayas (31°59'34" N, 77°30'35" E, altitude 2978) at district Kullu, Himachal Pradesh, India (Figure 1a-1b). The Parvati valley emerges from the confluence of rivers Parvati and Beas adjoined the hot water springs in the valley, including the hottest at Manikaran (Sangwan et al., 2015) have long been associated with high sulfur contents characterized by the distinct smell from sulfur dioxide or hydrogen sulfide gases escaping into the air (Shirkot and Verma 2015). Hot water is continually discharged from a major outlet at Khirganga from where it flows in a stream and deposits sulfur upon white microbial mats over the sediments along its course (Sharma et al., 2004). For this study, samples from all three habitats *viz.* microbial mats, sediments and water were collected proximal to the major opening (KgM1, KgS1, KgW1) and from a distance of 10m (KgM2, KgS2, KgW2) shown in Figure 1c.

Using helium-ion microscopy, we dissected the microstructure of the niches and were able to visualize cellular structures on complex sample matrices. The microbial diversity was visualized as numerous filamentous structures in microbial mats and sediments that resembled Cyanobacteria. Cyanobacteria are widely distributed in mats and sedimentary deposits of thermal springs (Gemerden et al., 1993; Podar et al., 2020). In addition, rod and cocci shaped cells of varying sizes were also observed in the sample matrices that providing a visual insight into the microbial diversity at this mesothermic site (Figure 1d).

Physicochemical and Elemental Analysis

The *in-situ* measures of water temperature were from 59 °C at the outlet to 55 °C at 10m distance (Table 1). Sediments and microbial mat deposits had much lower temperature (42-45 °C) than water. The pH of the hot spring water was 6.7 while sediments and mats were slightly acidic with pH 6.1 and 6.3 respectively. Thus, all three habitats were recorded to be mesothermic. The physicochemical composition of the hot spring is dominated by anions of chloride (up to 11024 mgL⁻¹) and sulfate (up to 10079 mgL⁻¹) while ions of calcium and potassium were abundant (Table 1; Supplementary File: 1). Importantly, sulfates (SO₄²⁻) concentration in microbial mats and sediments were higher (9529 ± 313.29 mgL⁻¹; 10079 ± 863.29 mgL⁻¹, respectively) and

exceeded the limit of 8000 mgL⁻¹ standardized by Environment Protection Act (EPA, 2001) and also found to be exceeded the limit of 53 mgL⁻¹ in surface waters (79.82 ± 1.85 mgL⁻¹) (EPA, 2001; Table 1). The chlorides (1456.77 ± 367.27 mgL⁻¹), manganese, sodium and silicon constituents in the hot spring waters were surpassing the normal average concentrations of 250, 0.05, 200 and 4 mgL⁻¹ respectively (EPA, 2001) in surface water samples (Supplementary File: 1). Among others, the predominant elements and minerals in water samples were aluminium, magnesium, copper, zinc and arsenic. The data signified that the hot spring waters have high concentrations of sulfates and chlorides that are characteristic of majority of other hot springs in the Himalayan ranges (Cinti et al., 2009).

In general, hydrogen sulfide account for the sulfur present in the underground geothermal waters originating from pyrites or leaching of other sulfides by deep hypothermal waters (Picard *et al.*, 2016). Sulfide (S²⁻) is oxidized to sulfate (SO₄²⁻) as the water rises to the surface and under mild oxidizing conditions, sulfide is only oxidized to sulfur or sulfur dioxide (Picard *et al.*, 2016; Wu *et al.*, 2021). The results provided pertinent information on the geochemical composition of the three habitats to be correlated with the microbial diversity and community functions. More importantly, high concentrations of sulfate ions in microbial mats and sedimentary deposits supported the hypothesis of a key role of the bacterial sulfur cycle in sustaining the microbial community at the hot water spring.

Metagenomic DNA Sequencing and Assembly

A large metagenomic dataset was obtained from sequencing having sufficient no. of reads sized up to ~18Gb for each sample. We retrieved a total number of reads ranging between 110,861,650 to 152,895,302 in all samples which were assembled into 180,849 - 519,194 (>200 bp) contigs. After assembly, the metagenomes sizes varied between 329 and 600 Mbp. A summary of characteristics of the datasets and assembled metagenomes is provided in Table 2. The alpha diversity estimated as the Shannon diversity indices ranged between 2.5 and 3 (Table 2; Supplementary File: 2). The alpha diversity was higher in the water samples than microbial mat and sediments as has been reported previously (Ghilamical *et al.*, 2017). The species richness as rarefaction curves obtained for all samples attained a plateau indicating optimum sequencing and sampling of a reasonable number of species for all metagenomes. The Bray-Curtis index calculated and plotted using NMDS demonstrated a significant difference in the

beta-diversity of all three habitats at phyla level (PERMANOVA; $P < 0.01$). Abundance of class *γ-Proteobacteria* in microbial mats significantly distinguished latter from the other two habitats. Similar results have been reported in previous studies (Selvarajan et al., 2018, Pohlner et al., 2019). Communities in sediment and water samples varied from each other majorly due to differences in the abundance of *δ-Proteobacteria*.

Dominance of Sulfur oxidizing bacteria (SOB) and sulfate reducing bacteria (SRB) in microbial consortia

Bacteria belonging to 15 different phyla dominated the microbial communities. The average percentage relative abundances of major phyla in the three habitats shown in parentheses in the order microbial mat, sediment and water as follows: *Proteobacteria* (62.1%, 50.5%, 58.7%), *Bacteroidetes*, *Firmicutes*, *Cyanobacteria*, *Planctomycetes*, *Chloroflexi* (Figure 2A). Species belonging to phylum *Proteobacteria* are found in varied temperature ranges which results in their dominance in various hot springs (De León et al., 2013; Singh et al., 2016; Saxena et al., 2017) and disproportionation of sulfur compounds is mainly carried out by SRMs of *Proteobacteria* (Finster, 2011). Besides, *Actinobacteria*, *Spirochaetes*, *Verrucomicrobia*, *Acidobacteria*, *Deinococcus-Thermus*, *Deferribacteres*, *Chlorobi*, *Gemmatimonadetes* and *Nitrospirae* were also detected with relative abundances less than 3%. Archaea superkingdom was represented by methane-producing phyla *Euryarchaeota* (2.5% in mat, 0.6% in sediment, 0.4% in water) and *Thaumarchaeota* which accounted for 0.5% of all classified sequences. These archaeal members are well known for actively carrying out ammonia-oxidation, nitrification and carbon fixation in marine and terrestrial environments in hot spring environments (Eme et al., 2013; de la Torre et al., 2008; Hatzenpichler et al., 2008). Among the three habitats, microbial diversity profiles of water and sediments were more similar compared to those of microbial mats.

Highest genus level diversity was revealed in sediment (n=196) followed by water (n= 132) and mat (n= 63). Top 50 genera in all habitats were plotted (Figure 2B). Microbial mats were dominated by *Pseudomonas* (61.8%) followed by unclassified genera of family *Desulfobulbaceae* (SRB; 5.2%) and *Flavobacterium* (4.3%). Among the abundant genera in the microbial mats were *Aeromonas*, *Flavobacterium*, *Exiguobacterium*, *Enterobacter*, *Escherichia* and *Klebsiella* (Figure 2B). These taxa are often found associated with mat deposits and are active biofilm producers that are thought to use adherence to the surface as a strategy to survive,

evolve and to cope with various abiotic stresses at such extreme habitats (López et al., 2006; Wright et al., 2013). On the other hand, sediment habitats were found to be enriched in *Thioalkalivibrio* (SOB; 18.9%), *Desulfobulbaceae* (SRB; 14%), *Burkholderia* (6.3%) and unclassified genera of families *Acetobacteraceae* (8%). Hot spring waters with highest diversity of bacterial genera were dominated by *Thioalkalivibrio* (9%), *Burkholderia* (6.5%), *Desulfovibrio* (SRB; 5.9%). Other genera with less than 6% abundances in all three habitats were also detected as shown in Figure 2B.

One important observation was the dominance of sulfur-oxidizing bacteria (SOB) and sulfate-reducing bacteria (SRB) in all three habitats (Agostino and Rosenbaum, 2018; Nagar et al., 2021). SOB and SRB are usually categorized as lithoautotrophs that play key microbial role in biogeochemical cycling of sulfur in various habitats. SOB oxidize the reduced sulfur compounds such as hydrogen sulfide (H_2S), elemental sulfur (S^0), sulfite (SO_3^{2-}), thiosulfate ($\text{S}_2\text{O}_3^{2-}$), and various polythionates ($\text{S}_n\text{O}_6^{2-}$ or $-\text{S}_n\text{O}_6^{6-}$) into sulfate (SO_4^{2-}). On the contrary, SO_4^{2-} can serve as an electron acceptor of SRB under anaerobic conditions, and they reduce the SO_4^{2-} and other oxidized sulfur compounds ($\text{S}_2\text{O}_3^{2-}$, SO_3^{2-} , S^0) into H_2S (Agostino and Rosenbaum, 2018). The abundance of SOB such as *Thioalkalivibrio* and *Burkholderia* as well as SRB such as *Desulfobulbaceae unclassified* and *Desulfovibrio* in spring is not surprising as high levels of sulfate dominate the site and relative abundance of these bacteria provide evidence of an active S cycling mediated by microbial communities. We identified the sequences producing significant alignments from the nr database for taxonomic confirmation and assigned each sequence that could be classified up to genus level to either SOB or SRB (Figure 4A, Supplementary File: 5). The taxonomy and phylogenetic topologies are discussed in detail in the next section.

Metabolic functions of the community and sulfur metabolism among the three sampling sites

The major gene families inferred to be abundant across all three habitats were assigned to *Proteobacteria* followed by *Chloroflexi*, *Firmicutes*, *Bacteroidetes* and *Spirochaetes* as reflected from the RPKs (reads per kilobase) in the metagenomes. These gene families were then regrouped as clusters of orthologous groups (COGs) and the top functions were determined to be translation, ribosomal structure and biogenesis (COG: J), amino acid transport and metabolism (E), general function prediction (R), energy production and conversion (C), replication,

recombination and repair (L) and carbohydrate transport and metabolism (G). These functions in microbial mats were carried out by *Proteobacteria* (COGs: J, E, C, G) and unclassified bacteria (COGs: R, L); in sediments by unclassified bacteria (COG: J) and *Proteobacteria* (COGs: E, R, C, L, G) and in water by *Proteobacteria* (COGs: J, C), *Firmicutes* and *Chloroflexi* (COGs: E, R, G, L) (Figure 2C; Supplementary File: 3).

The ORFs that were categorized on the basis of KEGG categories were mapped onto the metabolic functions and the pathways that could be reconstructed with more than 60% completeness were used to define the metabolic potential of the habitats. Based on this criterion, we studied the top 50 functional pathways of each habitat and identified the core functions (n=37 pathways) of the communities that included the common pathways for metabolism of nucleotides, carbohydrates and amino acids (Figure 3A, Supplementary File: 4). In addition, we determined differentially abundant pathways in each habitat: microbial mats (n=7), sediments (n=7) and water (n=2). Sediments were enriched in functions such as glycosaminoglycan degradation, metabolism of cyanoamino acid and riboflavin as well as degradation of nitrotoluene, geraniol, lysine and fatty acids while communities in water showed abundance of fatty acid and novobiocin biosynthetic pathways (Figure 3A). The community functional profiles of sediment and water were more similar compared to those of microbial mats which may be due to the stratified layered organization of the mats which are different in sediment and water. The microbial communities in mats were optimized for metabolism of methane specifically, members of genus *Methanospirillum*, which were abundant in mats (Figure 2B). Methanogens are shown to compete more effectively with SRB in attached-growth systems such as those of mats in which kinetic and thermodynamic considerations are the dominant mechanisms regulating the competition (Raskin et al., 1996). Other metabolic pathways such as lysine biosynthesis, C5-branched dibasic acid metabolism, thiamine metabolism, pyrimidine metabolism, vitamin B6 metabolism and other glycan degradation were also found to be abundant in microbial mat.

The enriched diversity for sulfur oxidation and sulfate reduction as well as the geochemical analysis of sulfide rich habitats compelled us to mine the regulatory genes involved in the different pathways of S cycle. In natural system, the sulfur intermediates viz., sulfate (SO_4^{-2}), thiosulfate ($\text{S}_2\text{O}_3^{2-}$), sulfite (SO_3^{-2}) or elemental sulfur (S^0) are reduced by different bacteria through two different reduction processes, viz., dissimilatory and assimilatory reactions (Vermeij

and Kertesz 1999; Zavarzin 2008; Figure 3B). In dissimilatory reduction, SRB utilize three enzymes (ATP sulfurylase, APS reductase, and sulfite reductase) to reduce sulfate and produce toxic hydrogen sulfide (Agostino and Rosenbaum, 2018; Kushkevych et al., 2020). On the contrary, sulfate is assimilated into organic compounds under assimilatory process (Figure 3B; Kushkevych et al., 2020).

The S metabolism pathway could be reconstructed within a range of 76.31 - 84.21% which was maximum in sediment and minimum in microbial mat. Total 75 genes were responsible for sulfur metabolism present in all the samples with a mean copy number of 1580 ± 249.2 . In order to gain insights into the sulfur oxidation as well as assimilatory and dissimilatory sulfate reduction potential across all the habitats, we mapped the TIGRFAM and Pfam (Supplementary File: 4) ids of the 25 associated genes (mean copy number of 980 ± 222.1) on to the ORFs and copy numbers of these S cycling genes involved in sulfur oxidation, sulfide oxidation and sulfate reduction (as described in Figure 3B) were estimated. Thiosulfate ions are fused to the carrier complex of sulfur-oxidizing proteins *soxYZ* (*soxY*=145, *soxZ*=71), while L-cysteine S-thiosulfotransferase (*soxX*=76, *soxA*=86) and S-sulfosulfanyl-L-cysteine sulfohydrolase (*soxB*=145) mediate the hydrolytic release of reduced sulfur ions from sulfur-bound *soxYZ*. Besides, the dissimilatory sulfite reductase encoding genes: *dsrA*=95, *dsrB*=83 and anaerobic sulfite reductase subunits *asrA*=44, *asrB*=30, *asrC*=9 reduces sulfite to sulfide (Supplementary File: 4). The other 15 genes for reduction of sulfate ions included solute binding protein (*cysP*=37), ATP binding protein (*cysA*=168), transport system permease proteins (*cysU*=158, *cysW*=154), ATP sulfurylase (*sat*=254), transferases (*cysC*=348, bifunctional *cysNC*=155, *cysN*=75), phosphoadenosine phosphosulfate reductase (*cysH*=314), adenylylsulfate reductase subunit A and B (*aprA*=88, *aprB*=104), sulfite reductase flavoprotein alpha-component (*cysJ*=111), sulfite reductase hemoprotein beta-component (*cysI*=180), homocysteine desulfhydrase (*mccB*=6), and cysteine synthase (*ATCYSC1*=4). Figure 3C shows the comparative abundances of these proteins in the three habitats. Here, the results revealed that the microbial reduction of sulfate occurs largely through the dissimilatory pathway carried out by dissimilatory sulfite reductase (DsrAB) as compared to assimilatory sulfite reductase (AsrABC) mediated reduction. In nature, the dissimilatory pathway is shorter and thus, preferred route of microbial sulfate reduction (Kushkevych et al., 2020).

Diversification, interactions and evolution of key enzymes involved in sulfur oxidation and reduction

To study the diversity and evolution of the key enzymes of SOB and SRB communities in this environmental biosystem, we employed a two-step strategy of comparing similarities of all sequences in a pairwise fashion through sequence similarity networks (SSN) analysis and further estimated the measures of the rates of non-synonymous to synonymous substitutions in their orthologous proteins between each pair of habitats. SSN effectively resolves the pairwise similarities of each sequence (node) with every other sequence of an enzyme or a group of enzymes for a pathway such that any two nodes are connected by edges only if they share sequence homology above a certain cutoff (here e-value of $1e-30$). Thus, SSN provides for an accurate placement of a sequence among its putative homologs (Talwar et al., 2020).

Here, we examined the diversity among 19 key S substrates oxidizing and reducing proteins determined from the communities as discussed in Figure 3C, except membrane permeases (CysPUWA) and genes for cysteine synthesis (MccB, ATCYSC1). In total, we retrieved 2413 protein sequences (mean; $M=254$, $S=480$, $W=472$) denoted by nodes in SSN. The network was organized into 88 connected components including 46 isolated nodes with an average clustering coefficient of 0.84. The connected components were represented by the homologous and heterologous clusters depending on whether they were constituted by the same gene or a number of different genes involved in a pathway, respectively. The number of connected components formed through SSN analysis of S metabolic proteins distributed into gene clusters and the isolated nodes denoted the diverging and highly diverged sequences, respectively (Figure 4A). Hence, we looked into these components to study the diversity of each gene that were distributed as shown in Table 3. The proteins CysNC, CysH, CysI and CysJ catalyse important steps and act as co-factors for the assimilatory sulfite reductase (AsrABC) which were all found to be diverging with many isolated nodes and loosely connected components. The enzymes for S oxidation (Sox) were also found to be diverging as observed from loosely formed clusters. On the other hand, all sequences of the key enzyme of dissimilatory pathway, DsrAB, formed only one connected component which suggested that they might be under convergent evolution at this site (Figure 4A). Further, we compared the distribution of diverged sequences that could be

separated as isolated nodes and found that hot spring sediments harbored a high diversity of these enzymes (n=23) in comparison with microbial mats (n=12) and water (n=11).

Based on sequence similarity, majority of the AsrABC genes that were taxonomically related to SRB could be distinguished from the rest of the sequences that were not identified as either SOB or SRB. Thus, assimilatory reduction was diverged among the SRB communities while those for dissimilatory pathway were rather conserved at the site. Previously, sequence comparisons have confirmed that dissimilatory sulfite reductase, Dsr, to be highly conserved enzyme that could serve as marker gene for SRBs (Loy et al., 2009). The DsrAB and AprAB enzymes were contributed by both SOB and SRB with syntrophic interaction which suggests for the presence of reverse dsr (rdsr) mediated oxidation of S substrates in addition to dissimilatory reduction. Here, the inherent complexity of sulfur-based metabolic network revealed that there are controlled mutation rates in dsrAB genes in presence or increased selective pressure of contamination and extreme conditions. The syntrophy of SOB and SRB prevailing in anoxic and anaerobic conditions governs the dissimilatory sulfur metabolism (oxidation and reduction parallelly) and indirectly promoting the growth of diverse microbes in this natural ecosystem (Bhatnagar *et al.*, 2020).

Through network analysis, we observed that, the node degree distribution estimated to be decreasing with increasing protein quantity (correlation=0.52, $r^2=0.29$), average neighborhood connectivity within the networks interpreted as function in k was increasing and positively correlated (correlation=0.90, $r^2=0.74$) (Figure 4B). Furthermore, closeness centrality curve that measures closeness between nodes was unable to reach the bench top (correlation=0.03, $r^2=0.01$), might be due to maximum number of connected components and less sequence homology. So, we also analyzed each protein cluster individually by using network analysis, 178/2413 nodes of the network DsrAB protein cluster found to be conserved showed higher clustering coefficient values (0.94), followed by AprA (0.93), AprB, CysNC (0.88) and others (Table 3). A relatively high diversity of the sulfate reduction proteins in all three segments of this habitat unveiled the high nutritional demands and efficiency of the microbes towards uptake of a wide range of structurally and chemically diverse amino acid side chains from environment (Talwar et al., 2020).

The evolutionary selection pressures on these genes were studied through estimation of dN/dS values calculated for a subset of conserved gene sequences in all three habitats (Figure 4C). The number of core genes and the range of dN/dS values identified for each gene are shown in Table 3. The CysJ and CysI genes were found to be under moderate selection pressures with dN/dS values in the range 0.4-0.7 (Figure 4C, Supplementary material: 6). The results supported the observation as these enzymes code for important co-factors for the assimilatory sulfite reductase (AsrABC) that were found to be diverging in through SSN analysis. Therefore, the microbial genes for assimilatory reduction pathway are diversifying under moderate selection pressures.

Divergence, phylogeny and structural relationships of DsrAB and AsrAB

The enzymes catalyzing the reductive (dissimilatory) or oxidative (rDSR) transformation between sulfite and sulfide appear to be related with respect to their subunit composition and catalytic properties (Loy et al., 2009). The dsr genes have been characterized from bacterial as well as archaeal domains (Chang et al., 2001; Grim et al., 2011; Colman et al., 2020). However, their evolution in these domains has long remained a subject of discussion. Our preliminary results showed that both subunits of dsr genes (dsrA=79, dsrB=72) corresponds to about 70 newly identified organism for both oxidation and reduction processes (Supplementary File 8). Through RAxML phylogenetic analysis, it can be confirmed that the dsrAB genes have been introduced in most of the newly identified members by a multiple independent Lateral Gene Transfer (LGT) (Anantharaman *et al.*, 2018). Importantly, organisms from *Acidobacteria*, *Candidate division Zixibacteria*, *Chloroflexi* and β -*Proteobacteria* form completely novel lineages other than known DsrAB clusters identified through RefSeq *nr* protein database with accession numbers (National Centre for Biotechnology Information; Supplementary File: 5). The result of the phylogeny suggested that an increase in substitution rate in both subunits of DSR may have occurred on the branch connecting δ -*Proteobacteria* to all other taxa as observed from the branch lengths. The different rates of substitution of the two DSR subunits has hitherto only been reported in δ -*Proteobacteria* lineage (Wagner et al., 1998). Hence, the dsr from sulfate reducers formed a separate cluster, with sequences from *Desulfarculus*, *Desulfocarbo*, *Desulfarcinum*, *Thermodesulfobacteria*, *Syntrophobacter*, *Desulfomonile*, *Desulfovibrio*, *Desulfatirhabdium* and *Desulfobacteriaceae* in both DsrA and DsrB phylogenies and additionally, *Dissulfuribacter* in DsrB (Supplementary File 8). We proposed that these

organisms with newly identified lineages of dsr genes involved in sulfite/sulfate oxidation and reduction likely serves an important control on sulfur cycling on terrestrial subsurface. The divergence of dsrAB between unrelated taxa could be driven through combination of speciation, functional diversification and LGT. Also, there is equal possibility of non-functionality of the genes in these taxa (Loy *et al.*, 2008; Anantharaman *et al.*, 2014).

Through time-scale evolutionary phylogeny of the sequences of DsrAB and AsrAB identified from the metagenomes, we determine the most basal and earliest evolved lineages involved in dsr and rsdr pathways (Figure 5; Supplementary File: 7). Our results suggested that both the subunits of the oxidative type reverse-DSR evolved much earlier than the reductive type DSR subunits (Figure 5). The templates of these *dsrAB* genes have potential to study the genotypic and phenotypic traits in SRPs and the dissimilatory sulfur metabolism processes which will expand the gene-environment interaction mechanism. Also, prior analysis has proved that the evolution of *dsrAB* have been influenced by LGT only among major taxonomic lineages (Klein *et al.*, 2001; Muller *et al.*, 2015) but the findings here provide evidence of independent multiple LGT events distributed throughout the dissimilatory gene clusters. Currently, the time scale study of this site cannot produce evidence of the progenitor lineages, as the evolutionary history of dissimilatory reduction is complex and yet ascertain. Although it had provided information of the earliest lineage where sulfate/sulfite oxidation and reduction appeared.

We used the phylogenetic analysis to further assign taxa to the sequences that showed similarity with yet uncultured bacteria and predicted their structures to gain insights into the more common phylogenetic ancestor of the two DSR subunits. Interestingly, these sequences of the oxidative type Dsr (rDSR) formed monophyletic clade with a more recently identified genus, *Sulfuritortus* in both DsrA and DsrB phylogenies. The genus of sulfur oxidizing bacteria currently has only one recognized species and is closely related to the members of genus *Thiobacillus* (Kojima *et al.*, 2017). Our analysis revealed similar tree topologies with these unassigned sequences forming clade with *Sulfuritortus*, *Thiobacillus* and *Hydrogenophilales bacterium* in both DsrA (n=1) and DsrB (n=3). Although the sequences were similar to *Thiobacillus* phylogenetically, prediction of their structures revealed DsrA to be highly similar to that of *Desulfovibrio gigas* (TM score= 0.75; Analog TM= 0.86; C-score=0.25, 1.6, 1.58; Supplementary File 9) and structures of DsrB proteins aligned most closely with *Archaeoglobus fulgidus* (TM score= 0.94; Analog TM= 0.97; C-score=1.64; Supplementary File 9, 10). While the other DsrB subunit (n=2) formed clades

with *Thioalkalivibrio* (β -Proteobacteria) also showing similarity to *A. fulgidus* (TM score= 0.94; Analog TM= 0.98; C-score=1.64).

In the reductive type Dsr clades, two DsrA sequences from uncultured bacteria formed clades with *Nitrospiraceae* and Chloroflexi bacteria (Figure 5). However, a consistent observation was seen in the similarity of all these structures of DsrA proteins with *Desulfovibrio vulgaris* (TM score= 0.94; Analog TM= 0.96; C-score= 1.58). Thus, we predict that the genus derives its two subunits of Dsr from different ancestors. A plausible explanation for this is observed in the previous reports of high sequence homology between the Dsr of *A. fulgidus* and *Desulfovibrio vulgaris* that suggest a common origin of archaeal and bacterial DSRs or their horizontal gene transfer (Karkhoff-Schweizer et al., 1995). In addition to homology in their sequences, the evolutionary distance separating the enzymes from *A. fulgidus* and *Desulfovibrio vulgaris*. For DsrB subunits, the archaeal and bacterial sequences were not particularly distant; such that the branches with structural homology to *A. fulgidus* were approximately the same length as branches leading to bacteria such as *Thiobacillus*, *Thioalkalivibrio* and β -Proteobacteria *bacterium* (Larsen et al., 1999).

Conclusion

To date, the microbial community and functional dynamics including genes of sulfur hot spring in Himalayas are poorly understood. This study was designed to investigate the syntrophic interactions between SOB and SRB in sulfur contaminated spring. Through analysis of microbial mat and sediment deposits along with water samples, we traced the microbial diversity and evolutionary dynamics of genes involved in S cycling within the Khirganga ecosystem. The mesothermic hot spring have been composed of a diverse group of microbes (Bacteria and Archaea) and genotypes (*dsrAB*) that could be screened out as novel thermozymes that cannot be underestimated. From the results, it could be concluded that the microbial community functions were distinguished in microbial mat from water and sediment. These thermophilic microbial inhabitants played a very crucial role in understanding the metal toxification, ion exchange and biogeochemical cycling of the elements that later became acquainted with physicochemical and elemental analysis of the habitat. Here, the genomic repertoire suggested the ongoing specific adaptations to cope up with extreme values of chloride, sulfate ions and manganese, silica contamination in this ecological setting. The sulfur metabolic pathways are completed where

inorganic sulfur compounds being the main source for sulfate reducing bacteria releasing toxicity in the form of sulfides (S^2-). The sulfate reduction profiling in all three habitats reveals dissimilatory sulfate reduction process (*dsrAB*) is active than assimilatory sulfate reduction (*asrAB*). Later, the genes involved in sulfur reduction/oxidation were classified and belong mostly to *Proteobacteria* with maximum homologous proteins classified in anoxygenic SOB. In all sulfur disproportionation proteins, the sulfite reductase DsrAB proteins showed conserved behavior with 0/1 isolated nodes that have been signified as phylogenetic markers for SRBs. The evolutionary phylogenetic analysis showed that the oxidative rDsr were the earliest then the reducing Dsr which may predict that the condition with more sulfides oxidized in more sulfates, directing sulfate reducing bacteria (SRBs) to perform dissimilatory reduction later. Phylogenetic clades of DsrAB proteins showed unanimous distribution of taxa except the δ -*Proteobacteria* which could be the reason for occurring LGT to other phyla. On the basis of structural alignments, the lineages with unclassified clades have shown different analogy in both Dsr subunits where DsrB derived from Archaea and DsrA are δ -*Proteobacteria* in origin at this meso-thermic niche.

Availability of Data

All sequence data presented in this article is submitted and is available in the public repository of NCBI under Bioproject number PRJNA673998, experiment Biosample ID and SRA number are shown in Supplementary File 1.

Acknowledgements

The sequence data were produced by the Beijing Genome Institute, China (BGI Tech. Solutions Co. Ltd, Hongkong) in collaboration with the user community. The authors are thankful for the use of the scanning electron microscope at the Centre for Chemical Microscopy (ProVIS) scanning at UFZ Leipzig, which is supported by European Regional Development Funds (EFRE-Europe funds Saxony) and the Helmholtz Association and The Energy and Resources Institute (TERI).

Funding

This work was supported by funds from the Department of Biotechnology (DBT), National Bureau of Agriculturally Important Microorganisms (NBAIM), University Grants Commission-Career Advancement Scheme and Department of Science and Technology-Purse grant. S.N. and C.T. thank Council of Scientific and Industrial Research (CSIR) for providing doctoral fellowships.

References

1. Agostino V and Rosenbaum MA (2018) Sulfate-Reducing ElectroAutotrophs and Their Applications in Bioelectrochemical Systems. *Front. Energy Res.* 6:55. doi: 10.3389/fenrg.2018.00055
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
3. Anantharaman K, Duhaime MB, Breier JA, Wendt KA, Toner BM, Dick GJ (2014) Sulfur oxidation genes in diverse deep-sea viruses. *Science*. **344**:757–760
4. Anantharaman K, Hausmann B, Jungbluth SP. *et al.* (2018) Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. *ISME J.* 12, 1715–1728 <https://doi.org/10.1038/s41396-018-0078-0>.
5. Ayangbenro AS, Babalola OO (2017) A New Strategy for Heavy Metal Polluted Environments: A Review of Microbial Biosorbents. *Int J of environ res publ heal*, 14(1), 94. <https://doi.org/10.3390/ijerph14010094>
6. Ayangbenro AS, Olanrewaju OS, Babalola OO (2018) Sulfate-Reducing Bacteria as an Effective Tool for Sustainable Acid Mine Bioremediation. *Front. Microbiol.* 9:1986. doi: 10.3389/fmicb.2018.01986
7. Battaglia-Brunet F, Crouzet C, Burnol A, Coulon S, Morin D, Joulain C (2012) Precipitation of arsenic sulphide from acidic water in a fixed-film bioreactor. *Water Res* 46:3923–3933.
8. Bhatnagar, S., Cowley, E.S., Kopf, S.H. *et al.* (2020) Microbial community dynamics and coexistence in a sulfide-driven phototrophic bloom. *Environmental Microbiome* **15**, 3. <https://doi.org/10.1186/s40793-019-0348-0>
9. Blank CE, Cady SL, Pace NR (2002) Microbial composition of near-boiling silica-depositing thermal springs throughout Yellowstone National Park. *Appl. Environ. Microbiol.* 68, 5123–5135.

10. Boopathy R (2014) Biodegradation of 2, 4, 6-trinitrotoluene (TNT) under sulfate and nitrate reducing conditions. *Biologia* 69:1264–1270
11. Bowen de León K, Gerlach R, Peyton BM, Fields MW (2013) Archaeal and bacterial communities in three alkaline hot springs in Heart Lake Geyser Basin, Yellowstone National Park. *Front. Microbiol.* 4:330. doi: 10.3389/fmicb.2013.00330
12. Callaghan AV, Morris BE, Pereira IA, McInerney MJ, Austin RN, Groves JT, Kukor JJ, Suflita JM, Young LY, Zylstra GJ, Wawrik B (2012) The genome sequence of *Desulfatibacillum alkenivorans* AK-01: a blueprint for anaerobic alkane oxidation. *Environ Microbiol* 14:101–113
13. Chan CS, Chan KG, Tay YL, Chua YH, Goh, KM (2015) Diversity of thermophiles in a Malaysian hot spring determined using 16S rRNA and shotgun metagenome sequencing. *Front. Microbiol.* 6, 177, <https://doi.org/10.3389/fmicb.2015.0017>.
14. Chang YJ, Peacock AD, Long PE, Stephen JR, McKinley JP, Macnaughton SJ, Hussain AK, Saxton AM, White DC (2001) Diversity and characterization of sulfate-reducing bacteria in groundwater at a uranium mill tailings site. *Appl. environ. Microb.* 67(7), 3149–3160. <https://doi.org/10.1128/AEM.67.7.3149-3160.2001>
15. Cinti D, Pizzino L, Voltattorni N, Quattrocchi F, Walia V (2009) Geochemistry of thermal waters along fault segments in the Beas and Parvati valleys (north-west Himalaya, Himachal Pradesh) and in the Sohna town (Haryana), India *Geochemical jour.* 43(2),65-76
16. Cole JK, Peacock JP, Dodsworth JA, Williams AJ, Thompson DB, Dong H, Wu G, Hedlund BP (2013) Sediment microbial communities in Great Boiling Spring are controlled by temperature and distinct from water communities. *ISME J.* 7, 718–729.
17. Colman DR, Lindsay MR, Amenabar MJ, Fernandes-Martins MC, Roden ER, Boyd ES (2020). Phylogenomic analysis of novel *Diaforarchaea* is consistent with sulfite but not sulfate reduction in volcanic environments on early Earth. *ISME J.* 14(5), 1316-1331.
18. Cox MP, Peterson DA, Biggs PJ. (2010) SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* 11:485. <https://doi.org/10.1186/1471-2105-11-485>.
19. Dahl C, Truper HG (1994). Enzymes of dissimilatory sulfide oxidation in phototrophic sulfur bacteria. *Methods Enzymol* 243, 400–421.

20. de la Torre JR, Walker CB, Ingalls AE, Konneke M, Stahl DA (2008) Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ Microbiol.* 10:810–818. [PubMed]
21. De León KB, Gerlach R, Peyton BM, Fields MW (2013) Archaeal and bacterial communities in three alkaline hot springs in Heart Lake Geyser Basin, Yellowstone National Park. *Front Microbiol.* 4:330
22. Dong Y, Sanford RA, Inskeep WP, Srivastava V, Bulone V, Fields CJ, Yau PM, Sivaguru M, Ahrén D, Fouke KW, Weber J, Werth CR, Cann IK, Keating KM, Khetani RS, Hernandez AG, Wright C, Band M, Imai BS, Fried GA, Fouke BW (2019). Physiology, Metabolism, and Fossilization of Hot-Spring Filamentous Microbial Mats. *Astrobio.*, 19(12), 1442–1458. <https://doi.org/10.1089/ast.2018.1965>
23. Eddy SR (2011). Accelerated profile HMM searches. *PLoS comput. biology*, 7(10), p.e1002195.
24. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–7.
25. Eme L, Reigstad LJ, Spang A, Lanze'n A, Weinmaier T, Rattei T, Schleper C, Brochier-Armanet C (2013) Metagenomics of Kamchatkan hot spring filaments reveal two new major (hyper) thermophilic lineages related to Thaumarchaeota. *Res. in Microbiol* 164 425e438.
26. EPA (2001) Parameters of Water Quality: Interpretation and Standards. Environmental Protection Agency, Wexford. https://www.epa.ie/pubs/advice/water/quality/Water_Quality.pdf
27. Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, Eddy SR (2014) Pfam: the protein families database. *Nucleic Acids Res.* 42: D222–30.
28. Finster K. (2011) Microbiological disproportionation of inorganic sulfur compounds. *J. Sulf. Chem.* 29: 281-292. <https://doi.org/10.1080/17415990802105770>
29. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, Lipson KS, Knight R, Caporaso JG, Segata N, Huttenhower C (2018) Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* 15: 962-968.
30. Gemerden H (1993) Microbial mats: a joint venture. *Mar Geol.* 113:3–25. doi: 10.1016/0025-3227(93)90146-M

31. Ghilamical AM, Budambula NLM, Anami SE (2017) Evaluation of prokaryotic diversity of five hot springs in Eritrea. *BMC Microbiol* 17, 203. <https://doi.org/10.1186/s12866-017-1113-4>
32. Ghosh W. et al. Molecular and cellular fossils of a mat-like microbial community in geothermal boratic sinters. *Geomicrobiol. J.* 29, 879–885 (2012).
33. Gomez-Alvarez V, Teal TK, Schmidt TM. (2009) Systematic artifacts in metagenomes from complex microbial communities. *ISME J.* 3:1314–1317. <https://doi.org/10.1038/ismej.2009.72>.
34. Gonsior M, Hertkorn N, Hinman N. (2018) Yellowstone Hot Springs are Organic Chemodiversity Hot Spots. *Sci Rep* 8, 14155. <https://doi.org/10.1038/s41598-018-32593-x>
35. Grim F, Franz B, Dahl C (2011) Regulation of dissimilatory sulfur oxidation in the purple sulfur bacterium *Allochrochromatium vinosum*. *Front. Microbio.* 2:51. doi: 10.3389/fmicb.2011.00051
36. Haferburg G, Kothe E (2007) Microbes and metals: interactions in the environment. *J Basic Microbiol.* 47(6):453–67. doi: 10.1002/jobm.200700275. PMID: 18072246
37. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E (2012) TIGRFAMs and genome properties in 2013. *Nucleic acids research*, 41(D1), pp.D387–D395.
38. Hatzenpichler R, Lebedeva EV, Spieck E, Stoecker K, Richter A, Daims H, Wagner M (2008) A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc Natl Acad Sci U S A.* 2008:2134–2139.
39. Hipp WM, Pott AS, Thum-Schmitz N, Faath I, Dahl C, Truper HG (1997). Towards the phylogeny of APS reductases and sirohaem sulfite reductases in sulfate-reducing and sulfur-oxidizing prokaryotes. *Microbiology* 143, 2891–2902
40. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Von Mering C (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol Biol Evol.* 34:2115–22.
41. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11: 119.
42. Inkscape Project 2020, <https://inkscape.org>

43. Inskeep WP, Jay ZJ, Tringe SG, Herrgard MJ, Rusch DB (2013) The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem *Frontiers in Microbiology* 4 :10.3389
44. Inskeep WP, Rusch DB, Jay ZJ, Herrgard MJ, Kozubal MA, Richardson TH, Macur RE, Hamamura N, Jennings R, Fouke BW, Reysenbach AL, Roberto F, Young M, Schwartz A, Boyd ES, Badger JH, Mathur EJ, Ortmann AC, Bateson M, Geesey G, Frazier M (2010) Metagenomes from High-Temperature Chemotrophic Systems Reveal Geochemical Controls on Microbial Community Structure and Function. *PLoS ONE*. 5:e9773.
45. J. Yang, Y. Zhang (2015) Protein structure and function prediction using I-TASSER *Curr. Protoc. Bioinforma* 10.1002/0471250953.bi0508s52
46. Jaekel U, Zedelius J, Wilkes H, Musat F (2015) Anaerobic degradation of cyclohexane by sulfate-reducing bacteria from hydrocarbon-contaminated marine sediments. *Front Microbiol* 6:116. doi:10.3389/fmicb.2015.00116
47. Kamarisima, K. Miyanaga, Y. Tanji (2019) The utilization of aromatic hydrocarbon by nitrate- and sulfate-reducing bacteria in single and multiple nitrate injection for souring control *Biochem. Eng. J.*, 143 (2019), pp. 75-80
48. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32: 277–280.
49. Karkhoff-Schweizer RR, Huber DP, Voordouw G (1995) Conservation of the genes for dissimilatory sulfite reductase from *Desulfovibrio vulgaris* and *Archaeoglobus fulgidus* allows their detection by PCR. *Applied and Environmental Microbiology*, 61(1), pp.290-296.
50. Klein M, Friedrich M, Roger AJ, Hugenholtz P, Fishbain S, Abicht H (2001). Multiple lateral transfers of dissimilatory sulfite reductase genes between major lineages of sulfate-reducing prokaryotes. *J Bacteriol.* 183:6028–35
51. Kojima H, Watanabe M, Fukui M (2017) *Sulfuritortus calidifontis* gen. nov., sp. nov., a sulfur oxidizer isolated from a hot spring microbial mat. *Int. j. syst. evol. microbiol.*, 67(5), pp.1355-1358.
52. Kolde R, Kolde, MR (2015) Package ‘pheatmap’. *R package*, 1(7), p.790.

53. Kushkevych I, Abdulina D, Kováč J, Dordević D, Vítězová M, Iutynska G, Rittmann SKM, (2020) Adenosine-5'-Phosphosulfate-and Sulfite Reductases Activities of Sulfate-Reducing Bacteria from Various Environments. *Biomolecules*, 10(6), p.921.
54. Larsen O, Lien T, Birkeland NK (1999) Dissimilatory sulfite reductase from *Archaeoglobus profundus* and *Desulfotomaculum thermocisternum*: phylogenetic and structural implications from gene sequences. *Extremophiles*, 3(1), pp.63-70.
55. Lartillot N, Rodrigue N, Stubbs D, Richer J (2015) Phylobayes-MPI 1.7: a Bayesian software for phylogenetic reconstruction using mixture models.
56. Li H, Yang Q, Li J, Gao H, Li P, Zhou H (2015) The impact of temperature on microbial diversity and AOA activity in the Tengchong Geothermal Field, China. *Sci Rep* 5, 17056 <https://doi.org/10.1038/srep17056>
57. Li SJ, Hua ZS, Huang LN (2014) Microbial communities evolve faster in extreme environments. *Sci Rep* 4, 6205 <https://doi.org/10.1038/srep06205>
58. Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, *Bioinformatics* 22: 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>
59. López MA, Zavala-Díaz de la Serna FJ, Jan-Roblero J, Romero JM, Hernández-Rodríguez C (2006) Phylogenetic analysis of a biofilm bacterial population in a water pipeline in the Gulf of Mexico. *FEMS Microbiol. Ecol.* 58: 145–154.
60. Loy A, Duller S, Baranyi C, Musmann M, Ott J, Sharon I, Béjà O, Le Paslier D, Dahl C, Wagner M (2009) Reverse dissimilatory sulfite reductase as phylogenetic marker for a subgroup of sulfur-oxidizing prokaryotes. *Environ Microbiol.* 11(2):289-99. doi: 10.1111/j.1462-2920.
61. Loy A, Duller S, Wagner M (2008) Evolution and ecology of microbes dissimilating sulfur compounds: insights from siroheme sulfite reductases In: Dahl C (ed) *Microbial Sulfur Metabolism* Springer: Berlin, Germany; 46–59
62. Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. *Proc. Natl Acad. Sci. USA* 104, 11436–11440 (2007).
63. Meckenstock RU, Boll M, Mouttaki H, Koelschbach JS, Cunha Tarouco P, Weyrauch P, Dong X, Himmelberg AM (2016) Anaerobic degradation of benzene and polycyclic aromatic hydrocarbons. *J Mol Microbiol Biotechnol* 26:92–118

64. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9: 386. doi: 10.1186/1471-2105-9-386.
65. Miller SR, Strong AL, Jones KL, Ungerer, MC (2009) Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Appl. Environ. Microbiol.* 75, 4565–4572.
66. Mothe GK, Pakshirajan K, Das G (2016) Heavy metal removal from multicomponent system by sulfate reducing bacteria: mechanism and cell surface characterization. *J Hazard Mater.* doi:10.1016/j.jhazmat.2015.12.042
67. Mulla SI, Talwar MP, Ninnekar HZ (2014) Bioremediation of 2, 4, 6-trinitrotoluene explosive residues. In: Singh SN (ed) *Biological remediation of explosive residues*. Springer, Switzerland, pp 201–233. doi:10.1007/978-3-319-01083-0
68. Müller AL, Kjeldsen KU, Rattei T, Pester M, Loy A (2015) Phylogenetic and environmental diversity of DsrAB-type dissimilatory (bi)sulfite reductases. *ISME J.* 9:1152–65. doi.org/10.1038/ismej.2014.208
69. Nagar S, Talwar C, Bharti M, Yadav S, Siwach S, Negi RK (2021) Metagenome-assembled genomes recovered from the datasets of a high-altitude Himalayan hot spring Khirganga, Himachal Pradesh, India, Data in Brief, 39:107551. doi.org/10.1016/j.dib.2021.107551
70. Peng Y, Leung HC, Yiu SM, Chin FY. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics.* 28: 1420-8. <https://doi.org/10.1093/bioinformatics/bts174>.
71. Picard A, Gartman A and Girguis PR (2016) What Do We Really Know about the Role of Microorganisms in Iron Sulfide Mineral Formation? *Front. Earth Sci.* 4:68. doi: 10.3389/feart.2016.00068.
72. Podar PT, Yang Z, Björnsdóttir SH, Podar M (2020) Comparative analysis of microbial diversity across temperature gradients in hot springs from yellowstone and Iceland. *Front. in microbiol*, 11, p.1625.
73. Poddar A, Das SK (2018) Microbiological studies of hot springs in India: a review. *Arch Microbiol* 200, 1–18 <https://doi.org/10.1007/s00203-017-1429-3>.

74. Pohlner M, Dlugosch L, Wemheuer B, Mills H, Engelen B and Reese BK (2019) The Majority of Active *Rhodobacteraceae* in Marine Sediments Belong to Uncultured Genera: A Molecular Approach to Link Their Distribution to Environmental Conditions. *Front. Microbiol.* 10:659. doi: 10.3389/fmicb.2019.00659
75. Raskin L, Rittmann, BE, Stahl DA (1996) Competition and coexistence of sulfate-reducing and methanogenic populations in anaerobic biofilms. *Applied and Environmental Microbiology*, 62(10), pp.3847-3857.
76. Reigstad LJ, Jorgensen SL, Schleper C (2010) Diversity and abundance of Korarchaeota in terrestrial hot springs of Iceland and Kamchatka. *ISME J.* 4, 346–356.
77. Roy C, Rameez MJ, Haldar PK (2020) Microbiome and ecology of a hot spring-microbialite system on the Trans-Himalayan Plateau. *Sci Rep* **10**, 5917 <https://doi.org/10.1038/s41598-020-62797-z>
78. Sahinkaya E, Yurtsever A, Toker Y, Elcik H, Cakmaci M, Kaksonen AH (2015) Biotreatment of As-containing simulated acid mine drainage using laboratory scale sulfate reducing upflow anaerobic sludge blanket reactor. *Miner Eng* 75:133–139
79. Sangwan N, Lambert C, Sharma A, Gupta V, Khurana P, Khurana JP, Sockett RE, Gilbert JA, Lal R (2015) Arsenic rich Himalayan hot spring metagenomics reveal genetically novel predator-prey genotypes. *Environ Microbiol Rep.* 7(6):812-23. doi: 10.1111/1758-2229.12297
80. Saxena R, Dhakan DB, Mittal P, Waiker P, Chowdhury A, Ghatak A and Sharma VK (2017) Metagenomic Analysis of hot Springs in Central India Reveals Hydrocarbon Degrading Thermophiles and Pathways Essential for Survival in Extreme Environments. *Front. Microbiol.* 7:2123. doi: 10.3389/fmicb.2016.02123.
81. Saxena R, Dhakan DB, Mittal P, Waiker P, Chowdhury A, Ghatak A, et al. Metagenomic analysis of hot springs in Central India reveals hydrocarbon degrading thermophiles and pathways essential for survival in extreme environments. *Front Microbiol.* 2017;7
82. Seemann T (2014) Prokka: rapid prokaryotic genome annotation *Bioinformatics* 30(14):2068-9.
83. Selvarajan R, Sibanda T, Tekere M (2018) Thermophilic bacterial communities inhabiting the microbial mats of "indifferent" and chalybeate (iron-rich) thermal springs:

- Diversity and biotechnological analysis. *Microbiologyopen*. 7(2): e00560.
doi:10.1002/mbo3.560
84. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13(11), 2498–2504.
<https://doi.org/10.1101/gr.1239303>
85. Sharma AK, Sharma R, Dandi HR (2004) Mineral Resources of Himachal Pradesh http://emerginghimachal.hp.gov.in/miningstone/survey_docs/Mineral_Resources_of_Himachal_Pradesh.pdf
86. Shirkot P, Verma A (2015) Assessment of thermophilic bacterial diversity of thermal springs of Himachal Pradesh. *ENVIS Bulletin Himalayan Ecology*. 23. 27-34.
87. Singh A, Subudhi E (2016) Structural insights of microbial community of Deulajhari (India) hot spring using 16s-rRNA based metagenomic sequencing. *Genom. Data* 7:101–2.
88. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), pp.1312-1313.
89. Stasik S, Wick LY, Wendt-Potthoff K (2015) Anaerobic BTEX degradation in oil sands tailings ponds: impact of labile organic carbon and sulfate-reducing bacteria. *Chemosphere* 138:133–139
90. Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH (2015) UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformat*. 31:926–32.
91. Talwar C, Nagar S, Kumar R, Scaria J, Lal R, Negi RK (2020) Defining the environmental adaptations of genus *Devosia*: Insights into its expansive short peptide transport system and positively selected genes. *Scientific reports*, 10(1), pp.1-18.
92. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, Tett A, Huttenhower C, Segata N (2015) MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. met*. 12(10), pp.902-903.
93. Varin T, Lovejoy C, Jungblut AD, Vincent WF, Corbeil J (2010) Metagenomic profiling of Arctic microbial mat communities as nutrient scavenging and recycling systems. *Limnol. Oceanogr*. 55:1901–1911

94. Vermeij P, Kertesz M (1999). Pathways of Assimilative Sulfur Metabolism in *Pseudomonas putida*. *J. of bacteriol.* 181. 5833-7. 10.1128/JB.181.18.5833-5837.1999.
95. Wagner M, Roger AJ, Flax JL, Brusseau GA, Stahl DA (1998) Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration. *Journal of bacteriology*, 180(11), pp.2975-2982.
96. Wickham H (2009) Elegant graphics for data analysis. *Media*, 35(211), pp.10-1007.
97. Wright KE, Williamson C, Grasby SE, Spear JR and Templeton AS (2013) Metagenomic evidence for sulfur lithotrophy by Epsilonproteobacteria as the major energy source for primary productivity in a sub-aerial arctic glacial deposit, Borup Fiord Pass. *Front. Microbiol.* 4:63. doi: 10.3389/fmicb.2013.00063.
98. Wu B, Liu F, Fang W, Yang T, Chen GH, He Z, Wang S (2021) Microbial sulfur metabolism and environmental implications. *Sci Total Environ.* 15; 778:146085. doi: 10.1016/j.scitotenv.2021.146085
99. Zavarzin GA (2008) Microbial Cycles, Encyclopedia of Ecology, Academic Press, 2335-2341, <https://doi.org/10.1016/B978-008045405-4.00745-X>.
100. Zhang J, Tian Y, Zhang J, Li N, Kong L, Yu M, Zuo W (2016) Distribution and risk assessment of heavy metals in sewage sludge after ozonation. *Environ Sci Pollut Res.* doi:10.1007/s11356-016-6313-1.

Table 1: Physicochemical and elemental analysis of the niche samples. Values for microbial mats and sediments samples are given in (ug/g) and those of water samples are shown in ppm.

Value for indicated sampling sites			
Environmental data	Mat	Sediment	Water
Temperature (°C)	42.4 ± 0.56	44.55 ± 1.06	57 ± 2.2
pH	6.2 ± 0.14	6.5 ± 0.14	6.7 ± 0
Major ions/metals	(ug.g ⁻¹)	(ug.g ⁻¹)	(mg.L ⁻¹)
Cl ⁻	11024.88 ± 983.17	15547.08 ± 204.96	1456.77 ± 367.27
SO ₄ ²⁻	9529.06 ± 313.29	10079 ± 863.29	79.82 ± 1.85
Na ⁺	976.88 ± 64.83	991.47 ± 7.88	1338.52 ± 167.97
Ca ²⁺	7529.71 ± 1034.68	5111.4 ± 588.21	30.27 ± 7.32
K ⁺	3681.66 ± 886.65	7743.91 ± 412.12	41.56 ± 9.03
Al	7205.41 ± 1685.37	22011.1 ± 2312.9	0.05 ± 0
Mg	3528.62 ± 735.04	8204.28 ± 964.02	8.83 ± 0.02
Si	3.22 ± 0.66	5.28 ± 0.42	31.5 ± 1.36
Mn	1149.7 ± 178.85	1219.29 ± 13.93	0.34 ± 0.03
Cu	14.4 ± 2.62	27.23 ± 1.71	<0.01
Zn	27.44 ± 6.36	67.35 ± 0.47	<0.01
As	3.9 ± 0.64	6.87 ± 0.22	<0.01
Ag	5.51 ± 3.7	5.1 ± 0.03	<0.01
Pb	17.74 ± 2.54	11.4 ± 1.27	<0.00

Table 2: Characteristics of sequenced datasets generated and assembled metagenomes obtained for each sample from different habitats.

Attributes	Microbial mat		Sediment		Water	
Sample name	KgM1	KgM2	KgS1	KgS2	KgW1	KgW2
(Clean reads x 100) bases	110,861,650 x100	152,895,302 x100	148,273,248 x100	145,473,498 x100	144,915,928 x100	141,139,772 x100
No. of contigs	180,849	216,350	404,684	398,499	519,194	506,942
Size of Assembly (Mbp)	329	397	592	583	600	591
N_{50}	7668	8123	2672	2664	3945	3985
GC- content (%)	55.55 ± 12.97	54.0 ± 12	59.70 ± 10.74	59.61 ± 11	56.43 ± 12.97	56.34 ± 13
Shannon Diversity Index	2.6	2.6	2.7	2.7	3.0	3.03
Paired-sample t-test (p-value < 0.05)	0.194		0.334		0.416	
Predicted proteins	305,932	369,390	570,660	562,428	545,606	537,498
CRISPRs	452	595	560	570	1105	1081

Table 3: Attributes of the sequence similarity networks and dN/dS analysis of S metabolism genes.

Network Parameters	Network Analysis of Homologous and Heterologous clusters															
	Sulfate reduction genes	CysNC	Sat	CysH	DsrAB	CysI	CysJ	AprA	AprB	AsrAB	AsrC	SoxA	SoxB	SoxX	SoxY	SoxZ
No. of Nodes	2413	578	254	346	178	170	102	87	98	68	9	86	145	76	145	71
No. of Edges	1,62,389	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Average degree	134.59	264.2	205.1	121.2	77.8	65.2	57.1	72.4	61.0	17.8	3.4	35.3	119.6	52.1	43.3	21.7
Connected components	88	17		14	1	11	8	2	3	6	4	4	2	2	7	7
Isolated nodes	46	12		8	0	10	5	1	0	4	2	0	0	0	2	2
Network Density	0.05	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Characteristic path length	1.89	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Shortest path	14%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Network centralization	0.16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Clustering coefficient	0.84	0.88	0.86	0.80	0.94	0.85	0.86	0.93	0.88	0.79	0.76	0.97	0.91	0.89	0.83	0.83
Core genes clusters	-	14	6	9	5	3	2	1	2	1	0	2	4	3	6	2
Range dN/dS (≤ 0)	-	≥ 0.4	≥ 0.5	≥ 0.5	≥ 0.6	≥ 0.5	≥ 0.6	≥ 0.4	≥ 0.1	0	-	0	≥ 0.4	≥ 0.1	≥ 0.3	≥ 0.3

Figures

Figure 1: a) Geographical location of Himachal Pradesh, Northern Himalayas, b) Khirganga hot spring as shown in the maps are located in Parvati valley in Kullu district of Himachal Pradesh, India, c) the different habitats from where the samples were collected are shown: microbial mat (green), sediment (yellow) and water (blue). Samples were collected in replicates from two outlets located 10m distance apart d) the scanning electron micrographs of different habitat samples shown with arrow demarcating the filamentous (cyanobacteria), cocci-shaped and rod-shaped bacteria (SRB) in pink, yellow and white respectively.

Figure 2: Relative abundance of phylum and genera. A) The stacked bar representation shows the dominating phylum in all three habitats. The dendrograms show hierarchical clustering between species and samples. B) The abundant common genera *Pseudomonas* (20–60%), *Desulfobulbaceae_unclassified* (15–20%), *Burkholderia* (10–20%), *Desulfovibrio* (5–10%) and *Thioalkalivibrio* (5–7%) are shown here relative abundance $\times \log_x$ scale. C) Taxa based functional profiles demonstrating the major phylum contributing towards COG subsystems and percentage of proteins annotated within each COG category for the habitat sites. The distribution of the RPKs were mapped in accordance to the member abundance in the habitats. Dots were colored according to the phylum.

Figure 3: A) Reconstruction of top 50 pathways annotated using KEGG Automatic Annotation Server. Heatmap matrix representation and clustering was performed by using 'pheatmap' package (Kolde and Kolde, 2015) in R (R Development Core Team, 2011 <http://www.R-project.org/>). B) The sulfate reduction pathway involved a group of reductases, kinases and transferases with the product chemical structures generated through chemDraw7 and Inkscape v0.9 (Inkscape project, 2020). C) The gene copy number of both sulfate reduction and sulfide oxidation pathway that were partitioned in different habitats showed here using ggplot2 in R (R Development Core Team, 2011 <http://www.R-project.org/>).

Figure 4: Sequence similarity network analyses. A) Diversity of sulfate reduction genes of both assimilatory and dissimilatory pathways in microbial mat (diamond-shaped), sediment (square) and water (spherical) habitats visualized in cytoscape v3.7.1. Highlighted only the classified taxa, where color cyan belong to SRBs and yellow to SOBs. The network was set at threshold e-value cutoff of $1e-30$ and nodes represent sequences connected through edges if the similarity exceeds the cutoff score. Here, the clusters and isolated nodes were showing the conserved pattern and diversified pattern of the proteins significantly playing an important role in sulfate reduction, B) Topological properties of the similarity networks: degree distribution, average clustering coefficient, average neighborhood connectivity and closeness centrality are plotted against the number of neighbors. The power law fit curves are shown within each graph. C) Habitat vs. habitat dN/dS values of all S cycle genes were estimated and plotted using xyplot in R (R Development Core Team, 2011).

Figure 5: Divergence estimation over time: Reconstruction of the phylogenic tree of optimized full length dissimilatory sulfite reductase and assimilatory sulfite reductase subunits in three habitats using PhyloBayes with the CAT-GTR model; The highlighted squares consist of clades with proteins that were remained unclassified through *nr* database. A. (i) Among, 78 DsrA nodes that showed here the earliest evolution of the rDsr oxidative proteins occurred in *Thiobacillus* sp. (ii) 72 nodes of DsrB proteins with similar results; B. (i) 38 nodes of AsrA (ii) 21 nodes of AsrB were also compared as a control for branch length shown here.

Supplementary material

Supplementary File 1: NCBI accession and in-situ measures of temperatures, pH and elemental analysis of each sample from three studied niches.

Sample ID	KgM1	KgM2	KgS1	KgS2	KgW1	KgW2
Accession number	SAMN16657637, SRX9422217	SAMN16657637, SRX9428981	SAMN16657991, SRX9429254	SAMN16673719, SRX9441237	SAMN16683881, SRX9449098	SAMN16683882, SRX9449099
Temperature (°C)	42.8	42	45.3	43.8	58.6	55.4
pH	6.1	6.3	6.4	6.6	6.7	6.7
ions/metals	(ug/g)	(ug/g)	(ug/g)	(ug/g)	(mg/L)	(mg/L)
Cl⁻	10041.71	12008.05	15342.12	15752.05	1089.50	1824.05
SO₄²⁻	9842.36	9215.77	10942.36	9215.77	81.68	77.97
Na⁺	1041.71	912.05	983.59	999.36	1170.54	1506.49
Ca²⁺	8564.4	6495.03	5699.61	4523.19	37.59	22.95
K⁺	4548.32	2815	7331.79	8156.03	32.53	50.59
Al	8890.79	5520.04	19698.23	24324.04	0.05	0.05
Mg	4266.67	2790.58	7240.26	9168.3	2.71	2.76
Si	2.56	3.88	4.86	5.71	32.86	30.14
Mn	1328.55	970.85	1233.23	1205.36	0.27	0.34
Cu	17.03	11.78	25.52	28.94	<0.01	<0.01
Zn	33.80	21.07	66.88	67.83	<0.01	<0.01
As	4.54	3.26	7.09	6.65	<0.01	<0.01
Ag	9.22	1.81	5.13	5.07	<0.01	<0.01
Pb	15.19	20.28	12.67	10.13	<0.00	<0.00

Supplementary File 2: Rarefaction curves based on measures of alpha diversity species richness calculated for each of the six samples (B) The interquartile range of alpha diversity measures based on Shannon indices. (C) Non-metric multidimensional scaling (NMDS) plots based on Bray-Curtis indices of community dissimilarities in the six samples types ($p < 0.05$, PERMANOVA).

Supplementary File 3: The list of all metabolic gene families in each sample (KgM1, KgM2, KgS1, KgS2, KgW1, KgW2) and the RPKs abundances of taxa (*Proteobacteria*, *Bacteroidetes* etc.) summarized in details.

Supplementary File 4: Detailed information of percentage reconstruction of metabolic pathways; Common and unique pathways in all six samples; Identification sources for 19 important proteins involved in sulfur disproportionation using TIGRfam and Pfam conserved domain database, ‘x’ indicates no hits found; Also, the copy number of 25 sulfur disproportionation proteins recorded.

Supplementary File 5: Details of taxonomic information of 2413 nodes of 19 sulfur cycle genes identified from nr NCBI microbial proteins database.

Supplementary file 6: Habitat vs. habitat dN/dS values of the sulfur cycling genes.

Supplementary File 7: Sequence alignment of DsrA, DsrB, AsrA, AsrB proteins identified in this study.

DsrA

DsrB

AsrA

AsrB

Tree scale: 0.1

DsrA

Species names (from top to bottom):

- Candidatus Magnetococcus marshallii
- Solibacteraceae uncultured
- Bromobacteriales bacterium
- Acidobacteria bacterium 21
- Acidobacteria bacterium
- Bacteroidetes bacterium
- Candidatus Division Zetabacteria
- Nitrospiraceae bacterium
- np506880
- np481424
- Chloroflexi bacterium
- Syntrophobacter sp.
- Desulfosarcina infernus
- Desulfocarbado indanensis
- Desulfuratus sp.
- Desulfotomaculum sp.
- Desulfotomaculum sp.
- Desulfotomaculum sp.
- Desulfotomaculum sp.
- Desulfotomaculum sp.
- Thermodesulfobacteria bacterium
- Hydrogenispirillum bacterium
- Gallinella bacterium RB
- Betaisproteobacteria bacterium
- Cronamonomadaceae SYSL
- Magnetospirillum sp.
- Aeropyrum haloproferens
- Gammaglutinisphaera bacterium
- Sulfuricoccus limicola
- Acidithiobacillus bacterium SM23
- Rhodovibrio unclassified
- Heteroproteobacteria bacterium R11
- Sulfurivibrio caldionitis
- np252935
- Sulfurivibrio caldionitis
- Thiohalophilus sp.
- Heteroproteobacteria bacterium SCN1
- Hydrogenispirillum bacterium 12-64-6
- Hydrogenispirillum bacterium 12-61-10
- Thiohalophilus sp. SCN1
- Heteroproteobacteria bacterium SCN1
- Thiohalophilus sp. SCN1

Tree scale: 0.1

DvR

The following table lists the species names and their corresponding DvR values as indicated by the color bar:

Species Name	DvR Value (approximate)
np199275	Low (Blue)
<i>Thioalkalibacter denitrificans</i>	Low (Blue)
<i>Gammaglutobacter bacterium</i>	Low (Blue)
np59533	Low (Blue)
<i>Acidobacterium bacterium</i>	Low (Blue)
<i>Chlorobis MS-B</i>	Low (Blue)
Candidate division LCP-14	Low (Blue)
<i>Igarivibrio bacterium</i>	Low (Blue)
<i>Nitrospirillum bacterium</i>	Low (Blue)
<i>Chloroflexus bacterium</i>	Low (Blue)
<i>Syntrophobacter sp.</i>	Low (Blue)
<i>Desulfatium infernum</i>	Low (Blue)
<i>Desulfarculus sp.</i>	Low (Blue)
<i>Disulfibacter thermophilus</i>	Low (Blue)
<i>Thermodesulfobacter bacterium</i>	Low (Blue)
<i>Devosimonia teijdei</i>	Low (Blue)
<i>Deltaproteobacterium CG2</i>	Low (Blue)
<i>Devolovibrio gigas</i>	Low (Blue)
<i>Deltaproteobacteria haiterum</i>	Low (Blue)
<i>Desulfobacteriaceae bacterium</i>	Low (Blue)
<i>Azospirillum haloproferens</i>	Low (Blue)
<i>Betaproteobacteria bacterium</i>	Low (Blue)
<i>Commamonadaceae bacterium</i>	Low (Blue)
<i>Rhodocyclaceae bacterium</i>	Low (Blue)
<i>Steroidobacteriaceae unclassified</i>	Low (Blue)
<i>Thiosphaellaceae bacterium</i>	Low (Blue)
<i>Betaproteobacteria bacterium</i>	Low (Blue)
<i>Sulfuriturus caldifontis</i>	Low (Blue)
np24977	Low (Blue)
np24617	Low (Blue)
<i>Sulfuriturus caldifontis</i>	Low (Blue)
np116684	Low (Blue)
<i>Sulfuriturus caldifontis</i>	Low (Blue)
<i>Hydrogenophiles bacterium</i>	Low (Blue)
RIFOXYD1	Low (Blue)
<i>Thiosphaella sp.</i>	Low (Blue)
<i>Thiosphaella denitrificans</i>	Low (Blue)

Tree scale: 0.1

AsrB

Phylogenetic tree showing the relationships between various bacterial taxa, with bootstrap values indicated at the nodes. The taxa are listed on the right, color-coded to match the branches.

- Candidatus Amniomonas bacterium*
- Volutellaceae bacterium*
- Clostridium bogliense*
- Clostridium sp.FP2*
- Bryobacteriales bacterium*
- Magnesiostirillum sp. 64-120*
- Deltaproteobacteria bacterium*
- Utsusumud acetothermia bacterium*
- Nitrospina gracilis*
- Elusimicrobium bacterium*
- Physciophaeus bacterium*
- Syntrophobacter sp. sbD1*
- Deltaproteobacteria bacterium*
- Desulfosporobium gigas*
- Deltaproteobacteria bacterium*
- Bacteroides bacterium*
- Desulfurculus sp.*
- Thermodesulfator sp.*
- Syntrophobacter-oxa bacterium*
- Desulfosoma caldarium*

Phyla

- Euryarchacota
- Proteobacteria
- Chloroflexi
- Bacteroidetes
- Firmicutes
- Spirochaetes
- Others
- Unclassified

— Newly identified lineages
— Previously known lineages

Supplementary File 8: Phylogenetic analysis of DsrAB proteins: showing percentage similarity with taxa on ncbi *nr* database; Few newly identified lineages were marked in red. In DsrAB maximum no. of LGT events were occurred in Proteobacteria class itself *clade A2; B2* while the independent events of LGT towards different phyla was occurred through δ -Proteobacteria depicted in *clade A1, B1*.

Supplementary File 9: Structural similarity and analogy of unclassified proteins from PDB (Protein data bank) using i-TASSER. A) DsrA protein subunits i) np252939 and ii); iii) np500880; np481424 showing Tm-align similarity with PDB ids 3or1 and 2v4j (*Desulfovibrio gigas* and *Desulfovibrio vulgaris*) respectively; with SF4 (iron sulfur cluster) ligand binding sites B) DsrB protein subunits i) np24977, np24617, np116484 and ii) np199275, np59533 showing Tm-align similarity with PDB ids 3c7b; with SRM (siroheme) ligand binding sites.

Supplementary File 10: Structural models for other ligand binding sites of unclassified proteins of DsrA and DsrB subunits.