Metagenomic insights into sulfate-reducing bacteria in a revegetated acidic mine wasteland

Jin-tian Li

South China Normal University College of Life Sciences

Pu Jia

South China Normal University College of Life Sciences

Xiao-juan Wang

South China Normal University College of Life Sciences

Shi-wei Feng

South China Normal University College of Life Sciences

Tao-tao Yang

South China Normal University College of Life Sciences

Zhou Fang

South China Normal University College of Life Sciences

Jun Liu

South China Normal University College of Life Sciences

Bin Liao

South China Normal University College of Life Sciences

Wen-sheng Shu

South China Normal University College of Life Sciences

Jie-Liang Liang (✉ liangjl@m.scnu.edu.cn)

South China Normal University College of Life Sciences <https://orcid.org/0000-0002-3949-7486>

Research

Keywords: Degraded lands, Facultative anaerobes, Metagenome-assembled genomes, Soil phages, Sulfate-reducing bacteria, Terrestrial ecosystems

Posted Date: October 21st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-92205/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

The widespread occurrence of sulfate-reducing microorganisms (SRMs, which are typically considered anaerobic organisms) in temporarily oxic/hypoxic aquatic environments indicates an intriguing possibility that SRMs can prevail in continuously oxic/hypoxic terrestrial environments rich in sulfate. However, little attention has been paid to such a possibility, leading to an incomplete understanding of microorganisms driving terrestrial part of the global sulphur cycle.

Results

In this study, genome-centric metagenomics was employed to explore SRMs in a revegetated acidic mine wasteland under continuously oxic/hypoxic conditions. We reconstructed 12 *Acidobacteria* and four *Deltaproteobacteria* genomes encoding reductive DsrAB, of which five represented three new SRM genera. Our results showed that *Acidobacteria*-related SRMs differed considerably from *Deltaproteobacteria*-related SRMs in metabolic potentials. Genomes of *Acidobacteria*-related SRMs harbored more glycoside hydrolase (GH) genes than those of previously known SRMs. They also tended to encode more oxygen-tolerant hydrogenases and cytochrome c oxidases, but less methyl-accepting chemotaxis proteins (MCPs) than genomes of *Deltaproteobacteria*-related SRMs. More importantly, we discovered that SRM-infecting viruses can contribute to glycoside hydrolysis, chemotaxis and antioxidation of their hosts. Remarkably, one GH encoded by a SRM-infecting virus is responsible for the liberation of rhamnose (a monosaccharide that is accessible directly to SRMs for dissimilatory sulfate reduction) from plant cell-wall-derived oligosaccharides.

Conclusions

Taken together, our results do not only improve our understanding of microorganisms driving dissimilatory sulfate reduction in terrestrial environments under continuously oxic/hypoxic conditions but also provides the first evidence for putative roles of viruses in S biogeochemical cycle in terrestrial ecosystems.

Background

Sulfate-reducing microorganisms (SRMs) are characterized by their ability to grow with energy derived from the reduction of sulfate to sulfide, which is coupled to the oxidation of organic compounds or molecular hydrogen (H₂) [1, 2]. The canonical dissimilatory sulfate reduction pathway in SRMs is driven by a set of enzymes [3], including sulfate adenylyltransferase (Sat), adenylyl-sulfate reductase (AprBA) and dissimilatory sulfite reductase (DsrAB). More specially, Sat and AprBA, encoded by *sat* and *aprBA*, cooperate to complete the reduction of sulfate to sulfite [2]. DsrAB, encoded by *dsrAB*, interacts with DsrC (a small protein, also known as DsvC) to reduce sulfite to sulfide [4, 5], a rate-limiting step in biogeochemical cycle of sulphur (S) on Earth [6, 7].

Due to the difficulty in isolating pure cultures of SRMs from the environment (especially acidic habitats) [8], as evidenced by the limited number of cultivated SRMs (approximately 240 species) [9], the functional genes *aprBA* and reductive *dsrAB* have been widely employed to explore SRMs diversity in the environment [9–13]. A striking example of such research was the work reported by Vigneron *et al.* [9]. The authors found 167,397 different species-level *dsrB* OTUs affiliated with 47 different families through high-throughput sequencing of *dsrB* genes in 200 environmental samples from 14 different ecosystems. Among these OTUs, approximately 99% were previously not detected, greatly improving our knowledge of the global species-level biodiversity of SRMs.

Recently, a new trend in SRMs research is the application of genome-centric metagenomics [14–18]. Compared to gene-centric approaches, the most important advantage of genome-centric metagenomics lies in the recovery of near-complete genomes representing species-level microorganisms of interest from complex environments [19]. Such an advantage renders not only discovery of previously uncultured (possibly uncultivable) microorganisms of interest but also insights into principal metabolic potentials of individual microorganisms of interest and putative interspecific interactions of microorganisms involved in a particular biogeochemical process [20–22]. Indeed, by retrieving and analyzing near-complete genomes containing reductive *dsrAB* from various environments, recent studies have revealed that: (1) eight prokaryotic phyla that were not previously reported to have the ability to reduce sulfate harbor the canonical functional genes required for dissimilatory sulfate reduction [14]; (2) two *Acidobacteria* genomes encoding reductive DsrAB but not Sat and AprBA contain additional genes for sulfite-producing enzymes that render the relevant SRMs to use organosulfonates as a growth substrate [15]; and (3) viruses can infect SRMs in wetland sediments and thus likely affect functions of their hosts in presently unknown ways [16].

Despite the aforementioned research advances, little is known about SRMs in terrestrial environments, which can be attributed largely to the fact that currently available studies on SRMs have focused almost exclusively on aquatic environments [2]. This situation represents an incomplete understanding of microorganisms responsible for terrestrial part of the global S cycle, as the geographic distribution of soils rich in sulfate (the most important growth substrate for SRMs) is by no means restricted to aquatic environments [23]. In fact, many natural processes (e.g. prolonged droughts and sea level declines) and anthropogenic interventions (including mining operations and draining wetlands for agricultural purposes) can lead to the distribution of sulfate-rich soils in terrestrial environments [23]. A major distinction between terrestrial and aquatic sulfate-rich soils lies in better aeration and continuously oxic/hypoxic conditions associated with the former, although oxic/hypoxic conditions can temporarily exist in the latter [24]. Such discrepancy, however, should not preclude the occurrence and activity of SRMs in terrestrial sulfate-rich soils.

The functioning of SRMs in temporarily oxic/hypoxic aquatic environments had long been recognized [25, 26], although SRMs are anaerobic organisms in nature. Moreover, there is considerable evidence that several *Deltaproteobacteria*-related model cultivated SRMs possess a variety of genes encoding enzymes conferring them a high tolerance to oxidative stress (see [24] for a review). A recent study even reported

the first SRM pure culture (i.e. *Desulfovibrio vulgaris* strains) that can grow using energy derived from oxygen reduction [27]. In contrast, existing literature provided only a few lines of evidence demonstrating the occurrence and activity of SRMs in terrestrial sulfate-rich soils under continuously oxic/hypoxic conditions (as indicated by the associated Eh values ranging from approximately 100 to 700 mV), which are related to sulfate-generating mine tailings deposited onto the surface soils of several Canadian mining areas (see [28] for a review). An attempt was made to characterize the community composition of SRMs in such mine tailings using *Bacteria* domain-, *Deltaproteobacteria* class- and *Desulfotomaculum*-specific primers, showing that the majority of potential SRMs were affiliated to *Firmicutes* and *Deltaproteobacteria* [29]. Additionally, two *Acidobacteria* genomes retrieved from an acidic sulfide mine waste rock site in Canada were found to encode canonical enzymes required for sulfate reduction [14]. However, no information regarding the geochemical properties of the site was available in the literature [14], which greatly limits our ability to exploit them to get a better understanding of the principal metabolic potentials (e.g. oxygen defense mechanisms) of SRBs in continuously oxic/hypoxic terrestrial environments. Nonetheless, more work is urgently needed, considering the importance of the topic.

In this study, we employed genome-centric metagenomics to characterize SRMs in a revegetated acidic mine wasteland rich in sulfate. Before revegetation, the wasteland (i.e. a mine tailings pond) was abandoned and drained for approximately eight years. This mine site was selected, as it consisted of different habitats with soil Eh values varying from approximately 180 to 680 mV [30], a representative Eh range of various terrestrial environments under continuously oxic/hypoxic conditions [31]. We recovered 12 *Acidobacteria* and four *Deltaproteobacteria* genomes encoding reductive DsrAB, of which five are members of three genera without previously known SRMs. Comparative genomic analysis revealed that the major metabolic potentials of *Acidobacteria*-related SRMs were considerably different from those of *Deltaproteobacteria*-related SRMs. Most importantly, we found for the first time that SRM-infecting viruses harbor auxiliary metabolic genes (AMGs) encoding proteins dedicated to glycoside hydrolysis, chemotaxis and antioxidation of their hosts.

Methods

Study site and soil sampling

We selected a revegetated acidic mine wasteland located in southern China (29°40'52"N, 115°49'21"E) as our study site. Briefly, this site was revegetated in spring of 2013 and thereby consisted of three different habitats: amended layer of the revegetated tailings (0–10 cm, ALRT), unamended layer of the revegetated tailings (11–20 cm, ULRT) and unrevegetated tailings (UT). These habitats were rich in sulfate (1.80–25.9 g SO₄²⁻ kg⁻¹ dry soil) and were under continuously oxic/hypoxic conditions (as indicated by a soil Eh range of approximately 180–680 mV). Three independent soil samples were collected from each of these habitats in July 2016 and 2017, respectively. More details on the study site, soil sampling and soil physic-chemical properties were presented elsewhere [30].

DNA extraction and sequence processing

DNA extraction, metagenomic sequencing and data processing (including assembly, binning, refinement, genome completion estimates, gene prediction, etc.) were described in detail elsewhere [30].

Metagenomic data analyzed in this study were deposited at EMBL under accession number PRJEB31441 [30] and the genome bins reported in this study have been deposited in GenBank under accession numbers SAMN15699825 and SAMN15808056–70.

Retrieval of key genes involved in dissimilatory sulfur metabolism

Genome-specific metabolic potential for sulfate/sulfite reduction was determined as follows. All predicted ORFs in a genome bin were searched against eggNOG [32] and KEGG [33] databases using Diamond [34] and against HMM profiles using InterProScan [35]. Then, the key sulfate reduction/sulfur oxidation genes (*dsrAB*, *dsrD*, *dsrT*, *dsrMKJOP*, *aprAB*, *sat* and *dsrEFH*) in genome bins were identified based on conserved domain hits elaborated in Anantharaman *et al.* [14].

Phylogenetic analysis of DsrAB proteins

A total of 214 DsrAB sequences, including those from previous studies [14, 36, 37], were used for phylogenetic analysis. The DsrAB sequences were aligned using MUSCLE [38] with default parameters. The alignments were then filtered by TrimAL [39] with the parameters -gt = 0.95 and -cons = 50. The concatenated DsrAB tree was constructed using RAxML [40] with the parameters set as -f a -m PROTGAMMAIJTT -p 12345 -x 12345 -# 100. The Newick files with the best tree topology were uploaded to the Interactive Tree of Life (iTOL) online interface [41] for visualization and formatting.

Taxonomic classification of reductive dsrAB-containing genome bins

16 genome bins retrieved in our study harbored reductive *dsrAB* sequences (Table S1). The direction of dissimilatory sulfur metabolism for each genomic bin was determined, according to the rules elaborated in Anantharaman *et al.* [14]. Taxonomic classification of the 16 SRM genome bins was inferred from two phylogenetic trees constructed with the reference genomes using GTDB-Tk [42] and PhyloPhlAn [43]. For *Acidobacteria* subdivision-level classification, 12 genome bins from *Acidobacteria* recovered in our study were used for phylogenetic analysis with published reference genomes spanned subdivisions 1, 3, 4, 6, 8 and 23 [15, 44]. One genome bin from *Syntrophobacteraceae* family without genus-level classification was used for phylogenetic tree construction with public reference genomes from *Syntrophobacteraceae* downloaded from NCBI. The maximum-likelihood phylogenetic trees were constructed based on a concatenated dataset of 400 universally conserved marker proteins using PhyloPhlAn and visualized using iTOL.

Calculation of relative abundances of genomic bins

Relative abundances of the 16 genome bins were calculated as previously described [30]. Briefly, the high-quality reads from each genomic dataset were mapped to all of the dereplicated genome bins using

BMap with the parameters $k = 14$, $\text{minid} = 0.97$, and $\text{build} = 1$. The coverage of the genome bin was calculated as the average scaffold coverage, and each scaffold was weighed by its length in base pairs. Then, the coverage of each genome bin divided by the total coverage of all genome bins in each sample was considered as its relative abundance.

Selection of genome bins for metabolic potential analysis

Genome bins with a completeness $> 90\%$ and contamination $< 10\%$ were chosen for further metabolic potential annotation, including six *Acidobacteria* genomes and one *Deltaproteobacteria* genome. The complete genomes of two model cultured SRMs (i.e. *D. vulgaris* Hildenborough (oxygen-tolerant) and *Desulfococcus multivorans* DSM 2059 (oxygen-sensitive) [45]) were also selected for comparative genomic analysis [46, 47].

Identification of carbohydrate-active enzymes

To identify carbohydrate-active enzyme genes, all predicted ORFs in the nine selected SRM genomes were searched against dbCAN2 meta server [48] with default parameters: HMMER (E-value $< 1e-15$, coverage > 0.35), Diamond (evalue $< 1e-102$) and Hotpep (Frequency > 2.6 , Hits > 6). Those identified by at least 2 tools were kept for further classification of glycoside hydrolase (GH) families using in-home Perl script.

Identification of hydrogenases

For identification of hydrogenases, hmm searches were performed by searching all predicted ORFs in the nine SRM genomes. Briefly, the individual HMM profiles for [NiFe] hydrogenases from Groups 1a–1 h, 2a–2d, 3a–3d and 4a–4e, [FeFe] hydrogenases from Groups A1–A4, B and C, and Fe hydrogenase were generated using the reference sequences retrieved from a previous study [49]. The reference sequences were aligned using MUSCLE with default parameters, then the alignment was converted into Stockholm format and databases were built using hmmscan [50]. The noise cutoffs for individual HMM profile were determined by manual inspection. Protein sequences that showed best hit with the HMM profiles with 1) bit-score greater than the calibrated threshold, and 2) over 90% sequence coverage, were retained.

Identification of proteins involved in respiration

All predicted ORFs in the selected genomes were searched for proteins involved in four respiratory complexes based on eggNOG annotation results, including NADH dehydrogenase, succinate dehydrogenase, quinol-cytochrome-c reductase, terminal oxidase and ATP synthase [15].

Identification of proteins involved in chemotaxis and oxidative stress

Besides methyl-accepting chemotaxis proteins (MCPs), the central components of bacterial chemotaxis system include CheA, CheB, CheR CheW and CheY [51]. MCPs were identified by Pfam annotation hits to PF00015, while the other central protein sequences were identified by KEGG annotation hits and were further confirmed based on eggNOG annotations. Classification of MCPs was according to Ud-Din & Roujeinikova [52]. Proteins involved in the core flagellum [53] and type IV pilus [54] systems were

identified by KEGG annotation hits and were further confirmed according to eggNOG annotation results. The antioxidative enzymes analyzed in this study were selected based on two previous reviews [24, 55] and were identified by KEGG, eggNOG and InterPro annotations.

Recovering and annotating viral scaffolds

VirSorter [56] was used to recover viral scaffolds from the 16 SRM genome bins as well as the two SRM reference genomes. The scaffolds from VirSorter categories 1, 2, 4 and 5 were retained. For scaffolds with predicted proviruses, only predicted proviral regions were retained. To taxonomically classify the viral scaffolds, a gene content-based network analysis was performed to cluster viral scaffolds into viral clusters at approximately the genus level, using vConTACT2 with ProkaryoticViralRefSeq94 database [57]. The ORFs in viral scaffolds were predicted with MetaProdigal. Virus signature proteins like terminase, integrase, capsid and tail were identified by Pfam hits. Viral sequences that encoded tail genes could be tentatively assigned to the order *Caudovirales*.

Estimation of viral–host abundances

Host (*Acidobacteria*-specific SRM) abundance and the abundance of virus for that host were both calculated as the normalized mean coverage depth as described elsewhere [58]. Briefly, the high-quality reads from each metagenomics dataset were mapped to all of the dereplicated viral scaffolds or dereplicated genome bins using BMap with the parameters $k = 14$, $\text{minid} = 0.97$, and $\text{build} = 1$. The viral or host abundances were pulled from the BMap mapping coverage output, normalized by the number of metagenomic reads in each sample, respectively. Pearson correlation analysis was used to correlate the viral and host abundances via the vegan package within the R statistical computing environment.

Identification of viral AMGs

To examine the potential roles of SRM-infecting viruses in S biogeochemistry, we assessed whether they contained AMGs. The predicted viral proteins were searched against dbCAN2 meta server and with HMM profiles using InterProScan as described above. For GHs, MCPs and nickel-containing superoxide dismutase (Ni-SOD) encoded by AMGs identified in this study, their protein sequences were structurally modeled using PHYRE2 in normal modeling mode (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) to confirm and further resolve functional predictions.

Results

Microorganisms harboring reductive dsrAB

We identified 50 DsrAB protein sequences in 982 genome bins recovered from 18 metagenomic datasets in the revegetated acidic mine wasteland [30]. 16 of these sequences belonged to the reductive bacterial-type DsrAB family (Figure S1). Accordingly, 16 reductive *dsrAB*-containing genome bins were retrieved, with 12 from *Acidobacteria* and four from *Deltaproteobacteria* (Fig. 1 & Table S1).

The 12 *Acidobacteria* genome bins were all affiliated to subdivision 1 of *Acidobacteria* (Figure S2). Among them, three (UT3_2.bins.71, UT3_3.bins.87 and UT4_3.bins.137) formed a monophyletic clade and had average 63% AAI to their closet relative *Granulicella tundricola* MP5ACTX9. Similarly, ULRT4_2.bins.48 formed a monophyletic clade and had 56%–62% AAI to its closest relatives. These four genomes represented two genera not previously reported to contain SRMs, given that no SRMs from *Acidobacteria* have been successfully cultivated and that currently known *Acidobacteria*-related genome bins containing *dsrAB* [15] were affiliated to other genera (Figure S2).

Three out of the four genome bins from *Deltaproteobacteria* were affiliated to the well-known SRM genus *Desulfovibrio* (Table S1). In contrast, another one (i.e. ULRT4_2.bins.61) can be only assigned to *Syntrophobacteraceae* and formed a monophyletic clade (Figure S3). It had 54–60% AAI to its closest relatives. Therefore, we inferred that it belonged to a new genus.

The Dsr operon structures of *Acidobacteria* were different from those of *Deltaproteobacteria* (Fig. 1). Multiple alignments of DsrD and DsrT sequences with published references confirmed highly conserved residues (Figures S4 & S5), indicating that these proteins are likely active. According to the rules for determination of direction of dissimilatory sulfur metabolism for uncultivated microorganisms [14], 11 genome bins (eight from *Acidobacteria* and three from *Deltaproteobacteria*) recovered in our study encoded the complete pathway for reduction of sulfate to sulfide (Fig. 1 & Table S2). Notably, seven genome bins (six *Acidobacteria* and one *Deltaproteobacteria*, Table S1) had a completeness > 90% and contamination < 10%.

Glycoside hydrolysis of SRMs

71 GH families were encoded by the nine SRM genomes considered in this study (Table S3). Among them, only one (i.e. GH 50) was not found in the six near-complete *Acidobacteria*-related genomes. This result is reasonable, as GH50 currently consists of β -agarase (EC 3.2.1.81), which is responsible for hydrolysis of (1→4)- β -D-galactosidic linkages in agarose (a polysaccharide produced by some aquatic red algae) [59]. In contrast, the three *Deltaproteobacteria*-related genomes encoded only 10 GH families. The numerical predominance of GH families observed for *Acidobacteria*-related genomes is remarkable even as compared to the results of Eichorst *et al.* [44] who identified 131 GH families in 24 non-SRB *Acidobacteria*-related genomes. According to the average number of GH genes (per genome), GH3, GH13, GH23, GH2, GH31, GH29, GH28, GH27, GH92 and GH35 were the top 10 most abundant ones across all investigated genomes (Fig. 2). Except for GH23, these abundant GH families were largely represented by *Acidobacteria*-related genomes. A striking example was GH3, which was encoded by 9–13 genes in each *Acidobacteria*-related genome but by only one gene in per *Deltaproteobacteria*-related genome (Fig. 2).

Hydrogen metabolism of SRMs

Genes encoding eight groups of hydrogenases (including Group A1, A2 of [FeFe]-hydrogenase and Groups 1a, 1b, 1d, 3d, 4c, 4e of [NiFe]-hydrogenase) were identified in this study (Table S4). Among them, [FeFe]-hydrogenase was encoded only by *D. vulgaris*. These results seem to agree with those of Hausmann *et al.* [15] who showed that genome bins of *Acidobacteria*-related SRMs harbored genes

encoding Groups 1 (excluding 1 h), 3, and 4 of [NiFe] hydrogenase. When individual genomes were taken into account, they differed considerably in the total number of hydrogenase genes (Table S4). The genome of *D. vulgaris* contained up to seven hydrogenase genes, whilst two *Acidobacteria* genomes (i.e. UT4_3.bins.137 and UT3_2.bins.71) lacked such genes. Despite this, seven out of the eight genes encoding oxygen-tolerant hydrogenases (i.e. Groups 1d, 3d of [NiFe]-hydrogenase) [49] were identified in the remaining four *Acidobacteria* genomes (Fig. 2).

Respiratory chain of SRMs

All investigated genomes encoded the major components of respiratory chain (Table S5). Specifically, the (near) complete operons for NADH dehydrogenases 1 (lacking in *D. vulgaris* and *D. multivorans*), NADH dehydrogenases 2 (lacking in UT4_3.bins.137 and UT3_2.bins.71), succinate dehydrogenase, quinol–cytochrome-c reductase, high-affinity terminal oxidase, low-affinity terminal oxidase (lacking in *D. multivorans*) and F-type ATP synthase were detected. Remarkably, high-affinity *bd*-type terminal oxidase was prevalent in all genomes, while high-affinity *cbb3*-type terminal oxidase was only detected in ULRT3_2.bins.110 (Table S5).

MCP system of SRMs

Two *Desulfovibrio*-related genomes (i.e. ULRT4_3.bins.101 and *D. vulgaris*) encoded much more MCPs than the other seven genomes (Fig. 2 & Table S6). Surprisingly, no MCP genes were detected in ULRT3_2.bins.110 and *D. multivorans*. While the majority of MCPs encoded by the two *Desulfovibrio* genomes belonged to class Ia (clusters I and II; Table S7 & Figure S6) that contained two experimentally validated redox and oxygen sensors (i.e. DcrA and DcrH) [52], acidobacterial MCPs were mainly from classes IVa and IVb. More specifically, three out of the five *Acidobacteria* genomes lacked genes encoding class Ia MCPs (Table S7).

The entire set of genes for core chemotaxis signaling complexes (CheB, CheR, CheW, CheA and CheY) [51] were detected in three acidobacterial genomes (i.e. ULRT4_1.bins.77, ULRT4_3.bins.75 and ULRT4_2.bins.48) and two *Desulfovibrio* genomes (i.e. ULRT4_3.bins.101 and *D. vulgaris*; Fig. 2 & Table S6). Additionally, the *che* operon structure in ULRT4_3.bins.101 was the same as that of *cheA3* in *D. vulgaris* (Figure S7), suggesting it had the ability to sense electron acceptors like sulfate and lactate [60].

Flagellum system of SRMs

The entire set of 24 core flagellar genes [53] was identified in four acidobacterial (i.e. ULRT4_1.bins.77, ULRT4_3.bins.75, ULRT3_2.bins.110 and ULRT4_2.bins.48) and two *Desulfovibrio*-related genomes (i.e. ULRT4_3.bins.101 and *D. vulgaris*; Fig. 2 & Table S6). Genes coding for highly conserved components of type IV pilus (e.g. *pilA*) were identified in UT4_3.bins.137 and UT3_2.bins.71 (Table S6), indicating that these two strains may move towards chemoattractants using pili-based “twitching” motility [54]. These results, together with those on MCP system, suggested that five investigated genomes (i.e. ULRT4_1.bins.77, ULRT4_3.bins.75, ULRT4_2.bins.48, ULRT4_3.bins.101 and *D. vulgaris*) have the

potential to utilize flagella-driven chemotaxis to sense surrounding chemoattractants and relocate themselves towards favorable microenvironments.

Antioxidative enzymes of SRMs

Among the four known enzymes involved in oxygen reduction by SRMs [24], only cytochrome bd oxygen reductase (Cbo, EC 7.1.1.7) was encoded by all the nine genomes (Fig. 2). The other three enzymes showed two contrastive patterns: (1) [Fe] hydrogenase (EC 1.12.7.2) and rubredoxin-oxygen oxireductase (ROO) were encoded largely by *Deltaproteobacteria*-related genomes; and (2) cytochrome c oxidase (Cco, EC 7.1.1.9) occurred mainly in *Acidobacteria*-related genomes (Fig. 2). Similarly, two opposite trends were observed for the two major enzymes responsible for eliminating superoxide anion radicals [24]: (1) all the investigated genomes encoded at least one SOD (EC 1.15.1.1), although the type of SOD differed between genomes; and (2) superoxide reductase (SOR, EC 1.15.1.2) genes were present only in *Deltaproteobacteria*-related genomes (Fig. 2). Note that the majority of *Acidobacteria*-related genomes lacked genes encoding catalase (EC 1.11.1.6) and thioredoxin peroxidase (Tpx, EC 1.11.1.15), whilst they harbored more genes encoding thioredoxin-dependent peroxiredoxin (BCP, EC 1.11.1.24), cysteine synthase (CysK, EC 2.5.1.47) and glutathione peroxidase (GPX, EC 1.11.1.9) than those of *Deltaproteobacteria*-related SRMs.

Viruses of SRMs

Six viruses (prophages) were identified across the 12 acidobacterial genome bins recovered in this study (Fig. 3 & Table S9), while no virus sequences were detected in the four deltaproteobacterial genome bins. Seven and three viruses were found in *D. vulgaris* and *D. multivorans* respectively, which seemed to be not in complete agreement with previous findings that eight and no viruses were identified in the two model SRMs correspondingly [47, 61]. This discrepancy likely resulted from the utilization of different viral prediction methods. Notably, PFAM annotations revealed that 11 out of the 16 viruses identified in this study harbored at least one virion-associated gene (Table S9), suggesting that these viruses still have the genetic potential to complete a lytic cycle [44].

Most of the identified 16 viruses could not be clustered with isolated viruses or those identified in publicly available microbial genomes or metagenomes using a gene-content based classification (genus-level grouping) [57], although half of them could be tentatively assigned to the order *Caudovirales* (Table S9). Specifically, four acidobacterial viruses formed an exclusive cluster, while the remaining two acidobacterial viruses were not closely related at the nucleotide level to any previously sequenced bacteriophages (i.e. singletons [62]; Table S9 & Fig. 3). Similarly, among the viruses of *D. vulgaris* and *D. multivorans*, one was affiliated to *Myoviridae* family, three were clustered exclusively, and the remaining six were singletons (Table S9 & Fig. 3). Additionally, the abundances of the viruses targeting acidobacterial SRMs were positively correlated with their host abundances (Figure S8).

Roles of viruses in glycoside hydrolysis of SRMs

Three genes encoding GHs were recovered from viral scaffolds, which were further predicted via three-dimensional protein structural modeling (Fig. 4 & Table S10). Among them, one was from the virus infecting ULRT3_2.bins.110 and encoded D-4,5-unsaturated β -glucuronidase (EC 3.2.1.172), which is able to release rhamnose from rhamnogalacturonan I oligomers (a major component of plant cell wall [63]; Fig. 4). The remaining two genes were identified on viral scaffolds D. vulgaris.2 and D. vulgaris.5, both of which encoded endochitinase (EC 3.2.1.14). This enzyme can cleave chitin randomly at internal sites, generating soluble low molecular mass multimers of N-acetyl-D-glucosamine, such as chitotetraose and chitotriose (Fig. 4) [64].

Roles of viruses in chemotaxis and antioxidation of SRMs

Three MCPs were encoded by viral scaffolds D. vulgaris.1 and D. vulgaris.2 (Fig. 4 & Table S11). Among them, two belonged to class Ia (cluster II) with double cache-like sensor domains, while the other one belonged to class Ia (cluster I) with single cache2 domain. According to previous findings [65] and the ligands confirmed in model protein structures, lactate and C2/C3 carboxylates (e.g. sodium acetate) could be the ligands for these MCPs (Fig. 4). On the other hand, one gene encoding Ni-containing SOD was identified on viral scaffold D. multivorans.2 (Fig. 4 & Table S11).

Discussion

To our knowledge, this work represents the first genome-centric metagenomic study that was specifically aimed at characterizing SRMs in a continuously oxic/hypoxic terrestrial environment. It seems somewhat surprising to find that SRMs affiliated to *Acidobacteria* prevailed in the mine wasteland (Fig. 1), since the limited evidence currently available indicates the dominance of *Firmicutes*- and *Deltaproteobacteria*-related SRMs in two other similar mine wastelands [29]. Nonetheless, the recovery of seven near-complete SRM genomes (six *Acidobacteria* and one *Deltaproteobacteria*) in this study provided us a particular opportunity to explore the principal metabolic potentials of these SRMs and their viruses, with emphasis on a comparison between *Acidobacteria* and *Deltaproteobacteria*. To this end, genomes of two *Deltaproteobacteria*-related model cultured SRMs with contrasting tolerance to oxygen stress (i.e. oxygen-tolerant *D. vulgaris* and oxygen-sensitive *D. multivorans*) were included [45].

The revegetated mine wasteland fostered novel SRM genera

The occurrence of SRMs in *Acidobacteria* was unveiled very recently [14]. So far, only 12 genomes of SRMs in *Acidobacteria* (with a completeness varying from 29.7–98.0%) have been reported [14–16]. In this context, this study not only doubled the number of such genomes but also detected the genus-level taxonomic novelty of *Acidobacteria*-related SRMs (Fig. 1). In addition, an unexpected finding of this study is that one *Deltaproteobacteria*-related SRM genome (i.e. ULRT4_2.bins.61) represents a genus without previously known SRMs, given that SRMs in *Deltaproteobacteria* were investigated extensively [2]. It thus becomes obvious that more work is needed to achieve a 'complete understanding' of the taxonomic

diversity of SRMs, although recent studies have greatly expanded the diversity of SRMs at the OTU and phylum levels [9, 14].

Acidobacteria-related SRMs encoded more GHs and oxygen-tolerant hydrogenases than Deltaproteobacteria-related SRMs

Many members of *Acidobacteria* are thought to have the ability to use a wide range of carbohydrates, as they dedicate a large portion of their genomes to carbohydrate metabolism [44]. Furthermore, there is evidence that three DsrAB-encoding *Acidobacteria* genomes recovered from peatland sediments harbor more GH genes (on average approximately 105 genes per genome) than not only the other known SRMs but also the majority of non-SRM *Acidobacteria* (approximately 60 genes per genome) [15]. Coincidentally, the average number of GH genes in the six near-complete DsrAB-encoding *Acidobacteria* genomes of this study was up to 120 per genome (Table S3), being much greater than that of the three *Deltaproteobacteria*-related SRMs. Our results indicate apparent adaptation of these *Acidobacteria*-related SRMs to the oligotrophic conditions of this mine wasteland [30]. Note that the most prevalent GH family across these *Acidobacteria* genomes is GH3 (Fig. 2), which consists of β -glucosidases (EC 3.2.1.21), β -xylosidases and N-acetylglucosaminidases (EC 3.2.1.52) [59]. These enzymes can liberate glucose and xylose (i.e. monosaccharides that can be used as a growth substrate by SRMs) [66] from plant cell-wall-derived oligosaccharides (e.g. cellobiose and xylan). In contrast, those enzymes belonging to the most abundant GH family across the *Deltaproteobacteria*-related SRMs (i.e. GH23; Fig. 2) are active on peptidoglycan and cannot release monosaccharides as a product [59].

Groups 1, 2a, 2b, 3a, 3b, 3c and 3d of [NiFe]-hydrogenase as well as groups A and B1/B2 [FeFe] hydrogenase are known to be involved in the oxidation of H_2 [67]. Two near-complete *Acidobacteria* genomes seemed to lack genes encoding such hydrogenases (Fig. 2), suggesting their inability to couple sulfate reduction to oxidation of H_2 . This characteristic likely puts the SRMs at a disadvantage in competition for growth substrates with co-existing microorganisms [49] and thus provides a possible explanation for the fact that they occurred in the wasteland at a lower relative abundance as compared to most of the other five SRMs possessing such genes (Table S1). Indeed, the hydrogenases encoded by the other four near-complete *Acidobacteria* genomes are oxygen-tolerant (i.e. 1d and/or 3d of [NiFe]-hydrogenase, Fig. 2) [49]. This seems to be in consistence with a previous finding that seven *Acidobacteria* genome bins, which can be putatively assigned as facultative anaerobes or aerobes (as they have low-affinity terminal oxidase genes) [68], all encoded oxygen-tolerant hydrogenases (i.e. 3b and/or 3d of [NiFe]-hydrogenase) [15]. Note, however, that *D. multivorans* rather than the other two *Deltaproteobacteria*-related SRMs also encoded an oxygen-tolerant hydrogenase (Fig. 2), which is different from our expectation. It is currently unclear what causes this discrepancy. Nonetheless, the specific functions of oxygen-tolerant hydrogenases in individual SRMs belonging to *Acidobacteria* and *Deltaproteobacteria* deserve further research.

Acidobacteria-related SRMs differed from Deltaproteobacteria-related SRMs in oxygen defense

The presence of genes encoding high-affinity terminal oxidases in all seven near-complete SRM genomes of this study (Table S5) suggests that these SRMs have the potential for microaerobic metabolism [68]. However, such genes are also present in the genome of the oxygen-sensitive model cultured SRM *D. multivorans* (Table S5), highlighting that elaborate antioxidant strategies are additionally required for the persistence of these seven SRMs in the mine wasteland.

According to the current knowledge of antioxidant mechanisms in SRMs, which are based largely on data from several *Deltaproteobacteria*-related cultured species, peculiar behaviors (including aggregation and aerotaxis) and antioxidant enzymes (e.g. SOD and SOR) are the two major antioxidant strategies [24]. MCPs and flagellum are indispensable components of the behavioral strategy [24]. The incompleteness of MCP and flagellum systems in *D. multivorans* (Fig. 2) is thus consistent with its sensitivity to oxygen, while the opposite is true for *D. vulgaris*. As expected, the near-complete *Deltaproteobacteria*-related SRM genome of this study encodes complete MCP and flagellum systems (Fig. 2). Moreover, the number (up to 34) and diversity of MCP genes on it seem to be higher than those of *D. vulgaris* (Table S7). On the other hand, it is actually somewhat surprising to find that the six near-complete *Acidobacteria*-related SRM genomes of this study harbor very few or even no MCP genes (Fig. 2). This surprising finding raises two possibilities: (1) the behavioral antioxidant strategy is insignificant for these *Acidobacteria*-related SRMs; and (2) there exist currently unknown proteins that have similar functions to MCPs. Note that the former possibility received little support from our results, as the *Acidobacteria*-related genomes encoding an incomplete MCP or flagellum system (i.e. ULRT3_2.bins.110, UT4_3.bins.137 and UT3_2.bins.71; Fig. 2) occurred in the wasteland at a lower relative abundance than their relatives with the ability to encode complete MCP and flagellum systems (Table S1).

SOD and SOR are thought to play an important role in scavenging superoxide ions in periplasm and cytoplasm of SRMs, respectively [24]. In agreement with this traditional wisdom, we found that genes encoding SOD are widespread across all SRM genomes considered in this study (Fig. 2). However, our results showed that genes encoding SOR are absent in the six near-complete *Acidobacteria*-related SRM genomes, despite their presence in those of *Deltaproteobacteria*-related SRMs. Similarly, most of the *Acidobacteria*-related SRM genomes lack genes encoding catalase (one major enzyme responsible for hydrogen peroxide elimination), although they all encode NPX (the other major enzyme responsible for hydrogen peroxide elimination; Fig. 2) [24]. These findings indicate a remarkable difference between the *Acidobacteria*- and *Deltaproteobacteria*-related SRMs in enzyme-based antioxidant strategy. It is likely that the *Acidobacteria*-related SRMs prefer to reduce oxygen to water before the powerful oxidant generates various types of damages in cytoplasm, as they tend to have more genes encoding CCO (i.e. a type of high-affinity terminal oxidases that are also involved in antioxidative defense) [24] than the *Deltaproteobacteria*-related SRMs (Fig. 2).

Virus infection in SRMs was widespread

Virus infection in SRMs was firstly recognized by Heidelberg *et al.* [46] who performed a whole-genome sequence analysis of *D. vulgaris*. Martins *et al.* [16] extended this work by documenting that *Acidobacteria*-, *Candidatus Aminicenantes*-, *Chloroflexi*-, *Deltaproteobacteria*-, *Nitrospirae*- and *Planctomycetes*-related SRM genomes retrieved from wetland sediments are hosts of viruses. In consistence with these previous studies, we found that five *Acidobacteria*-related SRM genomes recovered from the wasteland and the genome of *D. multivorans* were infected at least by one virus (Fig. 3). Taken together, the abovementioned findings indicate that virus infection in SRMs is more widespread than previously thought. Additionally, the *Acidobacteria*-specific virus/host abundance ratio recorded in this study (Figure S8) appeared to be greater than that for the phylum *Acidobacteria* in soils collected worldwide [58], which may be attributed to a scenario that the incidence of lysogeny in oligotrophic environments (e.g. mine wastelands) is high [69]. On the other hand, most of the viruses infecting SRMs in this study cannot be taxonomically assigned or even are not closely related at the nucleotide level to any known sequenced viruses (Fig. 3), supporting a notion that the diversity of environmental viruses is largely unexplored [16].

SRM-infecting viruses contributed to glycoside hydrolysis of their hosts

Viruses are widely thought to modulate S biogeochemical cycle in aquatic environments, as aquatic viruses can harbor several AMGs (i.e. *rdsrA*, *rdsrC* or *dsrC*) encoding enzymes directly involved in dissimilatory S oxidation or reduction [70, 71]. In contrast, little is known about the potential roles of viruses in S biogeochemical cycle in terrestrial environments. In this study, we failed to find viral AMGs encoding enzymes directly responsible for dissimilatory S oxidation or reduction. However, we identified two viral AMGs encoding enzymes dedicated to the oxidation of organic compounds (i.e. glycoside hydrolysis; Fig. 4 & Table S10). These results are remarkable, given that no previous studies have documented such metabolic potentials of viruses hosted by SRMs and that the oxidation of organic compounds is coupled to the reduction of sulfate in SRMs [1, 2]. In a wider context, a recent analysis of viral community composition and metabolic potential in mangrove sediments has speculated that viral carbohydrate AMGs may facilitate hosts to obtain energy for growth by decomposing complex carbohydrates in soil ecosystems [72]. Our finding on the viral AMG encoding D-4,5-unsaturated β -glucuronidase allows us to provide a novel mechanistic explanation for such an interesting notion, as this enzyme can degrade plant cell-wall-derived oligosaccharides into rhamnose (Fig. 4) [63], which is accessible directly to SRMs for dissimilatory sulfate reduction [66]. As to the endochitinase encoded by viruses in *D. vulgaris*, its products seem uncoupled with sulfate reduction in the host (Fig. 4) [66]. However, some of its products, such as chitobiose, can be metabolized further by the host for use in cell wall biogenesis [73]. Such a kind of synergism should be advantageous for viral hosts in oligotrophic environments, shedding some light on the 'black box' of soil virus–host interactions [69].

SRM-infecting viruses participated in chemotaxis and antioxidation of their hosts

Going beyond reporting viral GH genes, we found that three viruses infecting the two model cultured SRMs affiliated to *Deltaproteobacteria* are able to encode enzymes involved in chemotaxis or antioxidation (Fig. 4). There is only one previous study that documented viral AMG encoding MCP [74]. We extended this work by identifying a specific host of such AMGs for the first time (Fig. 4). Intriguingly, our in-depth analysis of the viral MCPs recorded in this study revealed that they possessed an extracellular LBD for C2/C3 carboxylates or alanine/lactate (Table S11). These characteristics of the viral MCPs may endow the host with a survival advantage in oligotrophic environments, considering that most of their targeted ligands are growth substrates for SRMs (Fig. 4) [66]. Similar to the situation for viral MCPs, only two lines of evidence showing the involvements of viral AMGs in hosts' resistance to oxidative stress are available in the literature. The first is that three viral AMGs (i.e. *yfdK*, *yfdO* and *yfdS*) encoding proteins with unknown functions in *Escherichia coli* are able to enhance the host's ability to resist oxidative stress [75]. The second lies in that viruses in oceanic cyanobacteria harbor genes encoding peroxiredoxin (a kind of peroxidases) and thus likely improve their hosts' antioxidative ability during oxygenic photosynthesis [76]. Therefore, our finding on the viral AMG encoding Ni-SOD deserves further research, although the generality of such a finding remains an open question.

Conclusions

Our results showed that members of *Acidobacteria* and *Deltaproteobacteria* were major microorganisms responsible for dissimilatory sulfate reduction in the revegetated acidic mine wasteland under continuously oxic/hypoxic conditions. Among them, five are novel SRMs at the genus level. *Acidobacteria*-related SRMs tend to encode more GHs, oxygen-tolerant hydrogenases and cytochrome c oxidases than *Deltaproteobacteria*-related SRMs, while an opposite trend was found for MCPs. More interestingly, a viral AMG associated with an *Acidobacteria*-related SRM is dedicated to liberation of rhamnose from plant cell-wall-derived oligosaccharides. Additionally, those viruses infecting *Deltaproteobacteria*-related SRMs can contribute to chemotaxis and antioxidation of their hosts by encoding MCPs and an antioxidative enzyme. Therefore, this study does not only improve our understanding of microorganisms driving dissimilatory sulfate reduction in terrestrial environments under continuously oxic/hypoxic conditions but also provides the first evidence for roles of viruses in S cycle in terrestrial ecosystems. Looking more widely, this study sheds some light on the 'black box' of soil virus–host interactions as well.

Abbreviations

ALRT: amended layer of the revegetated tailings; AMGs:auxiliary metabolic genes; BCP:thioredoxin-dependent peroxiredoxin; Cbo:cytochrome bd oxygen reductase; Cco:cytochrome c oxidase; CysK:cysteine synthase; *D. multivorans*:*Desulfococcus multivorans*; *D. vulgaris*:*Desulfovibrio vulgaris*; GH:Glycoside hydrolase; GPX:glutathione peroxidase; iTOL:Interactive Tree of Life; MCPs:Methyl-accepting chemotaxis proteins; Ni-SOD:Nickel-containing superoxide dismutase; ROO:rubredoxin-oxygen oxireductase;

S:sulphur; SOR:superoxide reductase; SRMs:Sulfate-reducing microorganisms; Tpx:thioredoxin peroxidase; ULRT:unamended layer of the revegetated tailings; UT:unrevegetated tailings.

Declarations

Acknowledgements

Not applicable

Authors' contributions

JTL and JLL designed all the experiments and wrote the manuscript. PJ, XJW, SWF, TTY, ZF and JL were responsible for performing the field and lab experiments. JTL, PJ, BL, WSS and JLL analyzed all data, discussed the results and critically reviewed the manuscript. All authors read the final manuscript and approved its publication.

Funding

This work was supported financially by the National Natural Science Foundation of China (Nos. 41622106, 31600082, 41561076, and 41603074), the Key-Area Research and Development Program of Guangdong Province (No. 2019B110207001) and the China Postdoctoral Science Foundation (No. 2018M640798).

Availability of data and materials

Metagenomic data analyzed in this study were deposited at EMBL under accession number PRJEB31441 and the genome bins reported in this study have been deposited in GenBank under accession numbers SAMN15699825 and SAMN15808056–70.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing Interests

The authors declare that they have no conflict of interest.

References

1. Barton LL, Hamilton WA. Sulphate-reducing bacteria: environmental and engineered systems. Cambridge: Cambridge University Press; 2007.

2. Rabus R, Venceslau SS, Wohlbrand L, Voordouw G, Wall JD, Pereira IAC. A post-genomic view of the ecophysiology, catabolism and biotechnological relevance of sulphate-reducing prokaryotes. *Adv Microb Physiol.* 2015;66:55–321.
3. Rückert C. Sulfate reduction in microorganisms – recent advances and biotechnological applications. *Curr Opin Microbiol.* 2016;33:140–6.
4. Venceslau SS, Stockdreher Y, Dahl C, Pereira IAC. The "bacterial heterodisulfide" DsrC is a key protein in dissimilatory sulfur metabolism. *BBA-Bioenergetics.* 2014;1837(7):1148–64.
5. Santos AA, Venceslau SS, Grein F, Leavitt WD, Dahl C, Johnston DT, et al. A protein trisulfide couples dissimilatory sulfate reduction to energy conservation. *Science.* 2015;350(6267):1541–5.
6. Bradley AS, Leavitt WD, Johnston DT. Revisiting the dissimilatory sulfate reduction pathway. *Geobiology.* 2011;9(5):446–57.
7. Leavitt WD, Bradley AS, Santos AA, Pereira IAC, Johnston DT. Sulfur isotope effects of dissimilatory sulfite reductase. *Front Microbiol.* 2015;6:1392.
8. Koschorreck M. Microbial sulphate reduction at a low pH. *FEMS Microbiol Ecol.* 2008;64(3):329–42.
9. Vigneron A, Cruaud P, Alsop E, de Rezende JR, Head IM, Tsessmetzis N. Beyond the tip of the iceberg; a new view of the diversity of sulfite- and sulfate-reducing microorganisms. *ISME J.* 2018;12(8):2096–9.
10. Chang YJ, Peacock AD, Long PE, Stephen JR, McKinley JP, Macnaughton SJ, et al. Diversity and characterization of sulfate-reducing bacteria in groundwater at a uranium mill tailings site. *Appl Environ Microbiol.* 2001;67(7):3149–60.
11. Friedrich MW. Phylogenetic analysis reveals multiple lateral transfers of adenosine-5'-phosphosulfate reductase genes among sulfate-reducing microorganisms. *J Bacteriol.* 2002;184(1):278–89.
12. Meyer B, Kuever J. Molecular analysis of the diversity of sulfate-reducing and sulfur-oxidizing prokaryotes in the environment, using *aprA* as functional marker gene. *Appl Environ Microbiol.* 2007;73(23):7664–79.
13. Watanabe T, Kojima H, Takano Y, Fukui M. Diversity of sulfur-cycle prokaryotes in freshwater lake sediments investigated using *aprA* as the functional marker gene. *Syst Appl Microbiol.* 2013;36(6):436–43.
14. 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(7):1715–28.
15. Hausmann B, Pelikan C, Herbold CW, Kostlbacher S, Albertsen M, Eichorst SA, et al. Peatland *Acidobacteria* with a dissimilatory sulfur metabolism. *ISME J.* 2018;12(7):1729–42.
16. Martins PD, Danczak RE, Roux S, Frank J, Borton MA, Wolfe RA, et al. Viral and metabolic controls on high rates of microbial sulfur and carbon cycling in wetland ecosystems. *Microbiome.* 2018;6(1):138.
17. Bell E, Lamminmaki T, Alneberg J, Andersson AF, Qian C, Xiong WL, et al. Active sulfur cycling in the terrestrial deep subsurface. *ISME J.* 2020;14(5):1260–72.

18. Colman DR, Lindsay MR, Amenabar MJ, Fernandes-Martins MC, Roden ER, Boyd ES. Phylogenomic analysis of novel Diaforarchaea is consistent with sulfite but not sulfate reduction in volcanic environments on early Earth. *ISME J.* 2020;14(5):1316–31.
19. Turaev D, Rattei T. High definition for systems biology of microbial communities: metagenomics gets genome-centric and strain-resolved. *Curr Opin Biotechnol.* 2016;39:174–81.
20. Raghoebarsing AA, Pol A, van de Pas-Schoonen KT, Smolders AJP, Ettwig KF, Rijpstra WIC, et al. A microbial consortium couples anaerobic methane oxidation to denitrification. *Nature.* 2006;440(7086):918–21.
21. Wrighton KC, Castelle CJ, Wilkins MJ, Hug LA, Sharon I, Thomas BC, et al. Metabolic interdependencies between phylogenetically novel fermenters and respiratory organisms in an unconfined aquifer. *ISME J.* 2014;8(7):1452–63.
22. Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, et al. Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature.* 2015;523(7559):208–11.
23. Ljung K, Maley F, Cook A, Weinstein P. Acid sulfate soils and human health – a millennium ecosystem assessment. *Environ Int.* 2009;35(8):1234–42.
24. Dolla A, Fournier M, Dermoun Z. Oxygen defense in sulfate-reducing bacteria. *J Biotechnol.* 2006;126(1):87–100.
25. Canfield DE, Des Marais DJ. Aerobic sulfate reduction in microbial mats. *Science.* 1991;251(5000):1471–3.
26. Minz D, Fishbain S, Green SJ, Muyzer G, Cohen Y, Rittmann BE, et al. Unexpected population distribution in a microbial mat community: sulfate-reducing bacteria localized to the highly oxic chemocline in contrast to a eukaryotic preference for anoxia. *Appl Environ Microbiol.* 1999;65(10):4659–65.
27. Schoeffler M, Gaudin AL, Ramel F, Valette O, Denis Y, Ben Hania W, et al. Growth of an anaerobic sulfate-reducing bacterium sustained by oxygen respiratory energy conservation after O₂-driven experimental evolution. *Environ Microbiol.* 2019;21(1):360–73.
28. Fortin D, Praharaj T. Role of microbial activity in Fe and S cycling in sub-oxic to anoxic sulfide-rich mine tailings. *J Nucl Radiochem Sci.* 2005;6(1):39–42.
29. Winch S, Mills HJ, Kostka JE, Fortin D, Lean DRS. Identification of sulfate-reducing bacteria in methylmercury-contaminated mine tailings by analysis of SSU rRNA genes. *FEMS Microbiol Ecol.* 2009;68(1):94–107.
30. Liang JL, Liu J, Jia P, Yang TT, Zeng QW, Zhang SC, et al. Novel phosphate-solubilizing bacteria enhance soil phosphorus cycling following ecological restoration of land degraded by mining. *ISME J.* 2020;14(6):1600–13.
31. Kludze HK, DeLaune RD, Patrick WH Jr. Aerenchyma formation and methane and oxygen exchange in rice. *Soil Sci Soc Am J.* 1993;57(2):386–91.

32. Huerta-Cepas J, Szklarczyk D, Heller D, Hernandez-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 2019;47(D1):D309-14.
33. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30.
34. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods.* 2015;12(1):59–60.
35. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, et al. InterPro: the integrative protein signature database. *Nucleic Acids Res.* 2009;37(suppl_1):D211-5.
36. Müller AL, Kjeldsen KU, Rattei T, Pester M, Loy A. Phylogenetic and environmental diversity of DsrAB-type dissimilatory (bi) sulfite reductases. *ISME J.* 2015;9(5):1152–65.
37. Tan S, Liu J, Fang Y, Hedlund BP, Lian ZH, Huang LY, et al. Insights into ecological role of a new deltaproteobacterial order *Candidatus Acidulodesulfobacterales* by metagenomics and metatranscriptomics. *ISME J.* 2019;13(8):2044–57.
38. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792–7.
39. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25(15):1972–3.
40. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30(9):1312–3.
41. Letunic I, Bork P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* 2011;39(suppl_2):W475-8.
42. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol.* 2018;36(10):996–1004.
43. Segata N, Börnigen D, Morgan XC, Huttenhower C. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Commun.* 2013;4:2304.
44. Eichorst SA, Trojan D, Roux S, Herbold C, Rattei T, Woebken D. Genomic insights into the *Acidobacteria* reveal strategies for their success in terrestrial environments. *Environ Microbiol.* 2018;20(3):1041–63.
45. Cypionka H, Widdel F, Pfennig N. Survival of sulfate-reducing bacteria after oxygen stress, and growth in sulfate-free oxygen-sulfide gradients. *FEMS Microbiol Ecol.* 1985;1(1):39–45.
46. Heidelberg JF, Seshadri R, Haveman SA, Hemme CL, Paulsen IT, Kolonay JF, et al. The genome sequence of the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *Nat Biotechnol.* 2004;22(5):554–9.
47. Dörries M, Wohlbrand L, Kube M, Reinhardt R, Rabus R. Genome and catabolic subproteomes of the marine, nutritionally versatile, sulfate-reducing bacterium *Desulfococcus multivorans* DSM 2059.

- BMC Genom. 2016;17:918.
48. Yin YB, Mao XZ, Yang JC, Chen X, Mao FL, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 2012;40(W1):W445-51.
 49. Greening C, Biswas A, Carere CR, Jackson CJ, Taylor MC, Stott MB, et al. Genomic and metagenomic surveys of hydrogenase distribution indicate H₂ is a widely utilised energy source for microbial growth and survival. *ISME J.* 2016;10(3):761–77.
 50. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol.* 2011;7(10):e1002195.
 51. Wuichet K, Zhulin IB. Origins and diversification of a complex signal transduction system in prokaryotes. *Sci Signal.* 2010;3(128):ra50.
 52. Ud-Din AIMS, Roujeinikova A. Methyl-accepting chemotaxis proteins: a core sensing element in prokaryotes and archaea. *Cell Mol Life Sci.* 2017;74(18):3293–303.
 53. Liu RY, Ochman H. Stepwise formation of the bacterial flagellar system. *Proc Natl Acad Sci USA.* 2007;104(17):7116–21.
 54. Wall D, Kaiser D. Type IV pili and cell motility. *Mol Microbiol.* 1999;32(1):1–10.
 55. Brioukhanov AL, Netrusov AI. Aerotolerance of strictly anaerobic microorganisms and factors of defense against oxidative stress: a review. *Appl Biochem Microbiol.* 2007;43(6):567–82.
 56. Roux S, Enault F, Hurwitz BL, Sullivan MB. VirSorter: mining viral signal from microbial genomic data. *PeerJ.* 2015;3:e985.
 57. Jang HB, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, et al. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat Biotechnol.* 2019;37(6):632–9.
 58. Emerson JB, Roux S, Brum JR, Bolduc B, Woodcroft BJ, Jang HB, et al. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat Microbiol.* 2018;3(8):870–80.
 59. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2014;42(D1):D490-5.
 60. Ray J, Keller KL, Catena M, Juba TR, Zemla M, Rajeev L, et al. Exploring the role of CheA3 in *Desulfovibrio vulgaris* Hildenborough motility. *Front Microbiol.* 2014;5:77.
 61. Crispim JS, Dias RS, Vidigal PMP, de Sousa MP, da Silva CC, Santana MF, et al. Screening and characterization of prophages in *Desulfovibrio* genomes. *Sci Rep.* 2018;8:9273.
 62. Šimoliūnas E, Kaliniene L, Truncaitė L, Zajančkauskaitė A, Staniulis J, Kaupinis A, et al. *Klebsiella* phage vB_KleM-RaK2 – a giant singleton of the family *Myoviridae*. *PloS One.* 2013;8(4):e60717.
 63. Silva IR, Jers C, Meyer AS, Mikkelsen JD. Rhamnogalacturonan I modifying enzymes: an update. *New Biotechnol.* 2016;33(1):41–54.
 64. Cohen-Kupiec R, Chet I. The molecular biology of chitin digestion. *Curr Opin Biotechnol.* 1998;9(3):270–7.
 65. Upadhyay AA, Fleetwood AD, Adebali O, Finn RD, Zhulin IB. Cache domains that are homologous to, but different from PAS domains comprise the largest superfamily of extracellular sensors in

- prokaryotes. PLoS Comput Biol. 2016;12(4):e1004862.
66. Colleran E, Finnegan S, Lens P. Anaerobic treatment of sulphate-containing waste streams. Anton Leeuw Int J G. 1995;67(1):29–46.
67. Vignais PM, Billoud B. Occurrence, classification, and biological function of hydrogenases: an overview. Chem Rev. 2007;107(10):4206–72.
68. Morris RL, Schmidt TM. Shallow breathing: bacterial life at low O₂. Nat Rev Microbiol. 2013;11(3):205–12.
69. Parikka KJ, Le Romancer M, Wauters N, Jacquet S. Deciphering the virus-to-prokaryote ratio (VPR): insights into virus-host relationships in a variety of ecosystems. Biol Rev. 2017;92(2):1081–100.
70. Anantharaman K, Duhaime MB, Breier JA, Wendt KA, Toner BM, Dick GJ. Sulfur oxidation genes in diverse deep-sea viruses. Science. 2014;344(6185):757–60.
71. Roux S, Hawley AK, Beltran MT, Scofield M, Schwientek P, Stepanauskas R, et al. Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell- and meta-genomics. Elife. 2014;3:e03125.
72. Jin M, Guo X, Zhang R, Qu W, Gao BL, Zeng RY. Diversities and potential biogeochemical impacts of mangrove soil viruses. Microbiome. 2019;7:58.
73. Howard MB, Ekborg NA, Weiner RM, Hutcheson SW. Detection and characterization of chitinases and other chitin-modifying enzymes. J Ind Microbiol Biot. 2003;30(11):627–35.
74. Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, et al. Functional metagenomic profiling of nine biomes. Nature. 2008;452(7187):629–32.
75. Wang XX, Kim Y, Ma Q, Hong SH, Pokusaeva K, Sturino JM, et al. Cryptic prophages help bacteria cope with adverse environments. Nat Commun. 2010;1:147.
76. Sharon I, Battchikova N, Aro EM, Giglione C, Meinel T, Glaser F, et al. Comparative metagenomics of microbial traits within oceanic viral communities. ISME J. 2011;5(7):1178–90.

Figures

Figure 1

Analysis of dissimilatory sulfate reduction genes on microbial genomes (bins) retrieved from the revegetated mine wasteland. Bins with a completeness > 50% and contamination < 10% are shown. Functions of the bins in sulfate reduction were predicted according to the presence and/or absence of those key genes for the pathway. Representative organization of dissimilatory sulfate reduction genes on the bins belonging to Acidobacteria and Deltaproteobacteria is displayed, respectively. More details were provided in Table S1 and Table S2.

Figure 1

Analysis of dissimilatory sulfate reduction genes on microbial genomes (bins) retrieved from the revegetated mine wasteland. Bins with a completeness > 50% and contamination < 10% are shown. Functions of the bins in sulfate reduction were predicted according to the presence and/or absence of those key genes for the pathway. Representative organization of dissimilatory sulfate reduction genes on the bins belonging to Acidobacteria and Deltaproteobacteria is displayed, respectively. More details were provided in Table S1 and Table S2.

Figure 2

Comparative analysis of selected metabolic potentials of sulfate-reducing microorganisms (SRMs). The wasteland-borne genomes (whose names are in blue) with a completeness > 90% and contamination < 10% were included in the analysis. For comparison, the complete genomes of *Desulfotomaculum vulgare* Hildenborough (an oxygen-tolerant model cultured SRM) and *Desulfococcus multivorans* DSM 2059 (an oxygen-sensitive model cultured SRM) were chosen. The top 10 most abundant glycoside hydrolase families across the genomes are shown. Those hydrogenase subgroups that are known to be involved in sulfate reduction or to be oxygen-tolerant are listed. MCP, methyl-accepting chemotaxis protein; BCP, bacterioferritin comigratory protein; Cbo, cytochrome bd oxygen reductase; Cco, cytochrome c oxidase; CcPx, cytochrome c peroxidase; SOD, superoxide dismutase; CysK, cysteine synthase; GPX, glutathione peroxidase; NPX, NADH peroxidases; ROO, rubredoxin-oxygen oxidoreductase; SOR, superoxide reductase; TPX, thioredoxin peroxidase. More details were presented in Tables S3-S8.

Figure 2

Comparative analysis of selected metabolic potentials of sulfate-reducing microorganisms (SRMs). The wasteland-borne genomes (whose names are in blue) with a completeness > 90% and contamination < 10% were included in the analysis. For comparison, the complete genomes of *Desulfococcus vulgaris* Hildenborough (an oxygen-tolerant model cultured SRM) and *Desulfococcus multivorans* DSM 2059 (an oxygen-sensitive model cultured SRM) were chosen. The top 10 most abundant glycoside hydrolase

families across the genomes are shown. Those hydrogenase subgroups that are known to be involved in sulfate reduction or to be oxygen-tolerant are listed. MCP, methyl-accepting chemotaxis protein; BCP, bacterioferritin comigratory protein; Cbo, cytochrome bd oxygen reductase; Cco, cytochrome c oxidase; CcPx, cytochrome c peroxidase; SOD, superoxide dismutase; CysK, cysteine synthase; GPX, glutathione peroxidase; NPX, NADH peroxidases; ROO, rubredoxin-oxygen oxireductase; SOR, superoxide reductase; TPX, thioredoxin peroxidase. More details were presented in Tables S3-S8.

Figure 3

Predicted linkages between SRMs and their viruses. The wasteland-borne genomes with a completeness > 50% and contamination < 10% were included in the analysis. For comparison, the complete genomes of *D. vulgaris* and *D. multivorans* were chosen. Cell shapes of SRMs were drawn based on information from the literature and from our analysis (i.e. the presence or absence of core genes encoding flagellum). Cell widths of SRMs were approximately proportional to their predicted genome sizes. Increasing abundances of wasteland-borne SRMs are indicated by darker green colors. Whenever applicable, viral shapes were drawn based on their taxonomy information. Otherwise, viruses are indicated by hexagons. The hexagons associated with a small rectangle represent tailed viruses. The widths of hexagons were approximately proportional to the predicted viral genome sizes, using a size scale smaller than that for their hosts. Those hexagons filled with the same number were affiliated to the same viral cluster. Singletons represent novel viruses. Increasing abundances of the acidobacterial viruses are indicated by darker red colors. Due to the lack of abundance information, the two model cultured SRMs and their viruses are shown in blue

and yellow, respectively. *, calculated as the normalized mean coverage depth. More details were presented in Table S9.

Figure 3

Predicted linkages between SRMs and their viruses. The wasteland-borne genomes with a completeness > 50% and contamination < 10% were included in the analysis. For comparison, the complete genomes of *D. vulgaris* and *D. multivorans* were chosen. Cell shapes of SRMs were drawn based on information from the literature and from our analysis (i.e. the presence or absence of core genes encoding flagellum). Cell widths of SRMs were approximately proportional to their predicted genome sizes. Increasing abundances of wasteland-borne SRMs are indicated by darker green colors. Whenever applicable, viral shapes were drawn based on their taxonomy information. Otherwise, viruses are indicated by hexagons. The hexagons associated with a small rectangle represent tailed viruses. The widths of hexagons were approximately proportional to the predicted viral genome sizes, using a size scale smaller than that for their hosts. Those hexagons filled with the same number were affiliated to the same viral cluster. Singletons represent novel viruses. Increasing abundances of the acidobacterial viruses are indicated by darker red colors. Due to the lack of abundance information, the two model cultured SRMs and their viruses are shown in blue and yellow, respectively. *, calculated as the normalized mean coverage depth. More details were presented in Table S9.

Gene types	AMGs encoding glycoside hydrolases		AMGs related to chemotaxis		AMGs related to antioxidation
Viruses	ULRT3_2.bins.110.1	<i>D. valgaris</i> .2 <i>D. valgaris</i> .5	<i>D. valgaris</i> .1	<i>D. valgaris</i> .1 <i>D. valgaris</i> .2	<i>D. multivorans</i> .2
Proteins	UGL (EC 3.2.1.172)	Endochit (EC 3.2.1.14)	MCP 1a (Cluster I)	MCP 1a (Cluster II)	Ni-SOD
Putative functions

Figure 4

Putative functions of proteins encoded by auxiliary metabolic genes (AMGs) in viruses infecting SRMs. UGL, D-4,5-unsaturated β -glucuronyl hydrolase; Endochit, endochitinase; MCP, methyl-accepting chemotaxis protein; Ni-SOD, nickel-containing superoxide dismutase; LBD, ligand binding domain; SD, signaling domain; TM, transmembrane; HAMP, histidine kinase, adenyl cyclase, methyl-accepting chemotaxis protein and phosphatase. Representative substrates of glycoside hydrolases (GHs) encoded by AMGs and the associated simplest products are shown. Cleavage points of the substrates are indicated by red arrows. The computational protein models of GHs, MCPs and Ni-SOD are displayed. Representative ligands of MCPs and the active site of Ni-SOD are shown. *, if the products of GHs or representative ligands of MCPs can be utilized directly by SRMs to reduce sulfate, their names are in blue. More details were presented in Tables S10 and S11.

Gene types	AMGs encoding glycoside hydrolases		AMGs related to chemotaxis		AMGs related to antioxidation
Viruses	ULRT3_2.bins.110.1	<i>D. valgaris.2</i> <i>D. valgaris.5</i>	<i>D. valgaris.1</i>	<i>D. valgaris.1</i> <i>D. valgaris.2</i>	<i>D. multivorans.2</i>
Proteins	UGL (EC 3.2.1.172)	Endochit (EC 3.2.1.14)	MCP 1a (Cluster I)	MCP 1a (Cluster II)	Ni-SOD
Putative functions

Figure 4

Putative functions of proteins encoded by auxiliary metabolic genes (AMGs) in viruses infecting SRMs. UGL, D-4,5-unsaturated β -glucuronyl hydrolase; Endochit, endochitinase; MCP, methyl-accepting chemotaxis protein; Ni-SOD, nickel-containing superoxide dismutase; LBD, ligand binding domain; SD, signaling domain; TM, transmembrane; HAMP, histidine kinase, adenylyl cyclase, methyl-accepting chemotaxis protein and phosphatase. Representative substrates of glycoside hydrolases (GHs) encoded by AMGs and the associated simplest products are shown. Cleavage points of the substrates are indicated by red arrows. The computational protein models of GHs, MCPs and Ni-SOD are displayed. Representative ligands of MCPs and the active site of Ni-SOD are shown. *, if the products of GHs or representative ligands of MCPs can be utilized directly by SRMs to reduce sulfate, their names are in blue. More details were presented in Tables S10 and S11.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables1.xlsx](#)
- [SupplementaryTables1.xlsx](#)
- [Supplementaryinformation.docx](#)
- [Supplementaryinformation.docx](#)