Metagenomic assembly unravel microbial response to redox fluctuation in acid sulfate soil

Jian-Qiang Su a,1, Yu Xia c,1, Huai-Ying Yao a,d,1, Ya-Ying Li a,d, Xin-Li An a, Brajesh K. Singh e,f, Tong Zhang c,*, Yong-Guan Zhu a,b,*

a Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, 1799 Jimei Road, Xiamen, 361021, China
b State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China
c Environmental Biotechnology Laboratory, The University of Hong Kong, Hong Kong
d Key Lab of Urban Environmental Processes and Pollution Control, Ningbo Urban Environment Observation and Research Station-NUEORS, Chinese Academy of Sciences, Ningbo, 315800, China
e Hawkesbury Institute for the Environment, Western Sydney University, Richmond, NSW 2751, Australia
f Global Centre for Land-Based Innovation, Western Sydney University, Richmond, NSW 2751, Australia

A R T I C L E   I N F O

Article history:
Received 5 August 2016
Received in revised form 28 November 2016
Accepted 29 November 2016
Available online 9 December 2016

Keywords:
Acid sulfate soil
Microbial community
Metagenomic
Redox fluctuation
Sulfur cycling gene

A B S T R A C T

Acid sulfate soil (ASS) is sensitive to redox fluctuations induced by climate change and human activities. Oxidation of sulfur and sulfide in ASS leads to the release of acid and consequently metals, posing severe hazards to coastal environment, while microbial contribution and response to oxidation is poorly understood. Here we used metagenomic sequencing to delineate the shift in microbial community structures and functional genes in ASS upon exposure to aerobic conditions. Aerobic incubation resulted in significant shifts in microbial communities in both topsoil and parent material. Archaea decreased significantly in the parent material after aerobic incubation. The relative abundance of sulfur cycling genes in the parent material layer was significantly higher than those in the topsoil, and multiple sulfide oxidation genes increased after aerobic oxidation. Metagenomic assembly enabled construction of eight key draft genomes from ASS. Three of them (GS3, GS6 and GT3) are novel strains of Thermoplasmatales, Acidothermales (Acidothermus) and Acidimicrobiales, respectively. Functional gene annotation of these population genomes revealed a dominance of sulfur cycling genes and acid tolerant genes. These findings highlight the microbial response to environmental change and identify the ecological adaptation and survival strategies of microbial communities in acid sulfate soils.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Acid sulfate soils (ASS) occupy about 1 million Km² worldwide (White et al., 2007), containing iron sulfides (predominately as pyrite) and the products of sulfide oxidation (White and Melville, 1996), and are mainly located in the delta plain and coastal lowlands in tropical and subtropical regions (Dent, 1992). The potential environmental hazard associated with ASS often results from the acid released by oxidation of sulfide minerals. These sulfide minerals are generally stable when kept submerged. However, due to global climate change (e.g. drought) and human activities (agricultural cultivation), these soils are frequently exposed to air, resulting in the oxidation of reduced sulfur compounds. This can negatively impact agricultural productivity, acidify groundwater and water bodies, poison aquatic organisms and degrade concrete and steel structures. Redox fluctuation will also affect ionic composition including Fe and other mineral elements that coupled with sulfur reduction and oxidation. In coastal regions, climate change and soil management may subject the ASS to frequent alternate draining/wetting cycles. Rewetting of ASS can release both acid and metals to surface water and to coastal environments (Stroud et al., 2014).

A variety of redox reactions involving inorganic and organic compounds of sulfur, mediated by microorganisms, can drive sulfur cycling in ASS (Biderre-Petit et al., 2011). Microbial processes,
including sulfur reduction and oxidation, are considered to promote the formation of acid soil (Bronswijk et al., 1993; Ward et al., 2002). Oxygen (O$_2$) and ferric ion (Fe$^{3+}$) are the main pyrite oxidizing agents. Diverse soil microbes, including sulfur-oxidizing bacteria and iron-oxidizing bacteria, are key contributors in sulfur transformation. However, little is known about the role of structural and functional microbial communities in sulfur transformation and soil acidification. Few studies have characterized soil microbial communities relevant to sulfur cycling targeting functional genes that code for key enzymes, such as dissimilatory sulfate reductase and sulfur dehydrogenase (Wu et al., 2013). Thus our understanding of microbial-mediated processes in ASS remains extremely limited.

Dominant microbes are considered to play significant roles in the maintenance and regulation of key processes because they are adapted to particular soil conditions. Given the vast majority of microbes are uncultivable (Singh et al., 2009), genome-centric metagenomics and draft genome assembly of uncultivated microbes are needed to understand their diverse role in ecology, ecosystem functions and adaptation to environmental conditions (Albertsen et al., 2013; Prosser, 2015). A variety of microbial genomes were obtained using single-cell genomics (Rinke et al., 2013), and binned genomes have been retrieved from different environments (Narasimarao et al., 2012; McLean et al., 2013; Yelton et al., 2013; Hug et al., 2016), including soil (Mondav et al., 2014; Hultman et al., 2015; Lin et al., 2015). However, the application of metagenomic sequencing to construct genomes of uncultivable microbes from soil is still challenging due to the higher diversity of soil microbial communities and the high relative abundance of key species is rare (Howe et al., 2014). Here we used deep metagenomic sequencing to determine the effect of oxidation on microbial phylogenetic composition and functional genes, and relate these data to the dynamics of sulfur transformation and soil acidification in ASS. Further, we constructed draft genomes of eight key microbial species from these samples and determined potential metabolic pathways involved in sulfur cycling.

2. Materials and methods

2.1. Soil sampling and incubation

Soil samples were collected from a typical area with ASS in Taishan county, Guangdong Province, China (22.11N, 112.77E). This area is located in the western part of the Pearl River Delta and faced South China Sea. Subtropical monsoon climates prevail here with an average temperature 21.8 °C and annual rainfall 2350 mm. Double rice cropping has been practiced in this site since 2003, and soil samples were collected at November 2012, during the ripening stage of rice plants. Soil samples with four replicates from four plots, were taken from depth of 0–15 cm (topsoil) and 80–100 cm (parent material layer), respectively. To evaluate the effect of land use on genetic diversity, neighboring native wasteland (topsoil) was also chosen as a study site for DNA extraction and metagenomic analysis. The samples from parent material layer were collected using a gouge auger, and the outer layer of each core was removed to prevent contamination. The parent material layer samples were placed into anaerobic plastic bags and flushed with nitrogen, thoroughly mixed, and then put into car-refrigerator for transport. The samples from the top layer (T) and parent material layer (S) were separated into three portions. The first portion (approximately 100 g of fresh soil) was passed through a 2.0 mm sieve, and then incubated at 25 °C and 90% relative air humidity for two weeks in aerobic condition. The soil samples were adjusted to a moisture content of 60% water holding capacity and distilled water was added to maintain initial moisture during the incubation. The second portion was stored at 4 °C for chemical analyses, and the third portion was stored at −80 °C for DNA extraction. Soil pH, water-soluble sulfur, total oxidizable sulfur and other chemical characteristics (Table S1) were measured, soil DNA were extracted and sequenced before and after incubation, respectively.

2.2. Soil chemical analysis

Soil pH was measured using a glass electrode (soil: water, 1:2.5). The inorganic N was extracted with 1 M KCl (soil: solution, 1:10), and determined by colorimetrically using a micro-plate reader (Shand et al., 2008). The total C, N, and S content in soil were determined by macro elemental analyzer (Vario MAX). Soil available phosphorus and potassium were determined by NaHCO$_3$ and NH$_4$OAc extraction methods, respectively (Bao, 1999). Water-soluble sulfate ions were extracted by adding 50 mL of distilled water purged with helium to 2.5 g of soil. All the samples were shaken for 1 h on a reciprocal shaker. The suspensions were filtered through a 0.22 µm membrane, and then immediately analyzed using an ion-exchange chromatograph. Total oxidizable sulfur (TOS) were estimated and calculated following the peroxy oxidation combined acidity and sulfate method (Ahern et al., 1998). According the change of water-soluble sulfate concentrations, the proton production was calculated using the software ECOSAT (Keizer and Riemsdijk, 1998).

2.3. Soil DNA extraction, sequencing and quality control

We tried DNA extraction by using commercial kits with 0.25–0.5 g soil, however, we did not obtain sufficient DNA for metagenomic sequencing due to low biomass in our samples, particularly in parent material. Therefore, we performed DNA extraction from a large volume of soil samples (5 g for topsoil and 50 g for parent material) by the freeze-grinding, SDS-based methods. Extracted DNA was purified using a low melting agarose gel followed by phenol extraction (Zhou et al., 1996). All DNA samples were stored under −20 °C before Illumina sequencing. Soil DNA was sequenced on Illumina Hiseq2000 platform with PE100 strategy at Personal Biotechnology Co., Ltd. (Shanghai, China). The derived raw reads were quality filtered with Prinseq (Schmieder and Edwards, 2011). First, reads with more than one ambiguous character “N” or with average quality score lower than 20 were filtered out. Next, reads showed more than 10 exact duplications were removed as artifacts.

2.4. Phylogenetic and functional annotation

To resolve the phylogenetic structure of the soil community, rRNA reads were identified from the dataset by searching against Silva SSU database version 11.5 (Quast et al., 2013) using BLASTN (Camacho et al., 2009) with e-value cut-off of 1E-20. The rRNA BLASTN result was parsed in Megan 5 (Huson et al., 2011) with default lowest common ancestor (LCA) algorithm to assign taxonomy. To facilitate the annotation speed, nucleotide sequences of targeted KEGG Orthologies (KO) involve in sulfur, nitrogen, methane metabolic pathways were extracted from the KEGG database and used as the subject database for BLASTN search. The pass-QC reads were searched against the extracted database with e-value cut-off of 1E-5, after that, in order to prevent false positive hit resulted from searching against subset database, strict similarity cutoff allowing only three mismatches over >80% of the read length were used to filter the BLASTN result before statistical analysis.
2.5. Sequence-based multi-dimensional genome binning

A sequence-based multi-dimensional binning approach combining de novo assembly and different coverage clustering was applied to retrieve genome bins from the soil metagenome. Dataset from the topsoil (T) and parent material (S) were individually assembled with CLCbio Generic Workbench 6.0.4 (CLCbio Denmark). Next, to evaluate the coverage of assembled scaffolds in different metagenomic datasets, reads in each dataset were mapped back to the assembled scaffolds with similarity cutoff of 98% over 90% of the read length by CLCbio Generic Workbench 6.0.4. Only scaffolds longer than 2 kb were used for the subsequent coverage binning. With the assumption that scaffolds belonging to the same genome shall share similar coverage across different metagenomes, genome bins were retrieved from coverage plot in R by recruiting gathered scaffolds clusters with protocol previously established (Albersen et al., 2013). Finally, tetranucleotide frequency (TNF) was used as genomic features to further purify the retrieved genome bins by removing scaffolds having conflicting TNF to the rest of the bin. The completeness and potential pollution of the retrieved genome bins were estimated by the Libani and 90 essential single copy genes (ESCGS) shared in >95% of the bacterial and archaeal populations, respectively, as provided by the Genome Properties at TIGRFAM (Haft et al., 2013), and further checked by an alternative method based on 35 conserved COGs (Raes et al., 2007). Average nucleotide identity (ANI) was calculated as previously reported (Konstantinidis and Tiedje, 2005).

2.6. Statistical analysis

Analysis of variance (ANOVA) and multivariate analyses were conducted using Genstat16th (NAG Ltd., Oxford, UK). Means, least significant differences at 5% level were calculated by one-way ANOVA. The data on functional genes were expressed as relative abundance of canonical variate analysis (CVA), after first reducing the dimensionality by principal component analysis. Heatmap were performed in R3.1.0 (RCoreTeam, 2014) with pheatmap package (Kolde, 2013). Principal coordinate analyses (PCoA) of both nucleotide identity (ANI) higher than 94% (Konstantinidis and Tiedje, 2005).

2.7. Sequence accession numbers

All metagenomic sequences have been deposited to the MG-RAST database under the accession number mgp13205 (http://metagenomics.anl.gov/linkin.cgi?project=mgp13205). The retrieved genome bins were submitted to IMG for genome annotation (see Table 1 for the IMG accession).

3. Results

3.1. Soil chemical properties

The total sulfur content in the surface soil were nearly twice that of nitrogen, and the parent material layer had a much higher total sulfur content compared to the surface layer, suggesting the sulfur deposition and accumulation during the formation of ASS (Table S1). The oxidizable sulfur was highly correlated with total sulfur, accounting for approximately 60% of the total sulfur. To obtain the effects of incubation on soil biogeochemical properties, we incubated soils in aerobic conditions which resulted in marked increase in water-soluble sulfate during incubation, which was accompanied by a significant (P < 0.05) decrease in soil pH (Fig. 1).

3.2. Microbial phylogeny and functional genes

We determined the microbial community composition for both topsoil and parent materials. The metagenome data revealed significant shifts in phylogenetic structures in both topsoil and parent material after aerobic incubation (Fig. S1). Eurarchaeota, Firmicutes and Proteobacteria showed significant differences after aerobic incubation. Actinobacteria increased in all the samples after incubation, while Archaea decreased significantly in the parent material layer after incubation (Fig. S2).

The composition of functional genes involved in C, N, P, and S cycling also showed significant difference between samples before and after incubation (Fig. S3a). The overall structures of the genes related to soil sulfur cycling in the topsoil were distinct from those in the parent material layer, and had a lower normalized abundance compared to the parent material layer (Fig. 2). Sulfur oxidation and reduction gene abundance shifted significantly after incubation. We found that the abundances of dsrC, dsrE, soxX and soxB1 genes increased after incubation in all soils, whereas some reduction genes in the parent material layer, e.g., cysJ, cysU, cysP, were found to be in lower abundance after incubation. Canonical variate (CV) analysis emphasized significant differences in the composition of sulfur cycling genes between surface and parent material layers, and further before and after incubation (Figs. S3c and d). Multivariate analysis of sulfur reduction genes revealed that the soils

Table 1

<table>
<thead>
<tr>
<th>Name</th>
<th>Phylum</th>
<th>Total length Mb</th>
<th>Scaffolds number</th>
<th>GC %</th>
<th>Completeness</th>
<th>Redundancy</th>
<th>Abundance in soil samples, %</th>
<th>IMG accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS1*</td>
<td>Proteobacteria</td>
<td>5.0</td>
<td>35</td>
<td>60</td>
<td>100%</td>
<td>3.7%</td>
<td>11.514</td>
<td>29.722</td>
</tr>
<tr>
<td>GS2*</td>
<td>Proteobacteria</td>
<td>6.0</td>
<td>57</td>
<td>67</td>
<td>100%</td>
<td>1.9%</td>
<td>18.779</td>
<td>10.613</td>
</tr>
<tr>
<td>GS3</td>
<td>Euryarchaeota</td>
<td>2.3</td>
<td>22</td>
<td>69</td>
<td>90.0%</td>
<td>1.6%</td>
<td>9.509</td>
<td>0.867</td>
</tr>
<tr>
<td>GS4*</td>
<td>Proteobacteria</td>
<td>6.2</td>
<td>234</td>
<td>68</td>
<td>98.1%</td>
<td>1.0%</td>
<td>2.515</td>
<td>3.672</td>
</tr>
<tr>
<td>GS6</td>
<td>Actinobacteria</td>
<td>5.2</td>
<td>748</td>
<td>72</td>
<td>96.3%</td>
<td>2.9%</td>
<td>4.031</td>
<td>4.150</td>
</tr>
<tr>
<td>GT1</td>
<td>Acidobacteria</td>
<td>4.5</td>
<td>316</td>
<td>61</td>
<td>90.7%</td>
<td>1.0%</td>
<td>0.014</td>
<td>0.003</td>
</tr>
<tr>
<td>GT2</td>
<td>Acidobacteria</td>
<td>3.8</td>
<td>110</td>
<td>57</td>
<td>96.3%</td>
<td>1.9%</td>
<td>0.016</td>
<td>0.011</td>
</tr>
<tr>
<td>GT3</td>
<td>Actinobacteria</td>
<td>3.8</td>
<td>418</td>
<td>67</td>
<td>96.3%</td>
<td>8.7%</td>
<td>0.021</td>
<td>0.013</td>
</tr>
</tbody>
</table>

a GS1, GS2 and GS4 are respectively the same species of Serratia marcescens WW4, Delftia acidovorans SPH-1, Achromobacter xylosoxidans NBRC15126 as indicated by average nucleotide identity (ANI) higher than 94% (Konstantinidis and Tiedje, 2005).

b Estimated abundance of the genome bin in each soil sample was evaluated as the number of reads aligned to the genome bin in percentage of the total reads number in the corresponding metagenome.

c Sample name abbreviation: SB, SA: parent material (S) soil before (B) and after (A) incubation; TB, TA: topsoil (T) before (B) and after (A) incubation; TN: native wasteland soil.
3.2. Environmentally relevant genes in retrieved population genomes

After incubation had higher ordinates on CV1. Correlation analysis of the loadings on CV1 indicated that cysJ, cysC, cysP genes were negatively correlated and cysD, cysA were positively correlated with CV1. Furthermore, Procrustes analyses showed that both phylogenetic community composition and functional genes clustered by sampling site and aerobic incubation, consistently displaying highly significant goodness-of-fit measures ($P < 0.001$, $M^2 = 0.179$) (Fig. S3b).

3.3. Metagenomic assembly summary and phylogeny of retrieved bins

To retrieve draft population genomes from the soil, the metagenomes of topsoil and parent materials were assembled and the phylogeny of each scaffold was annotated at phylum level (Fig. S4). A total of 34.3 Gb and 41.6 Gb Illumina reads (Table S2) from the topsoil (T) and parent material (S) were de novo assembled, resulting in 82 and 382 Mb contigs $>1$ kb, respectively (Table S3). About 77.3% of parent material and 52.4% of topsoil metagenomic reads were utilized in the assembly (Table S3), similar to an adequate genomic recovery from soil metagenomes (Howe et al., 2014). Eight near-complete (estimated completeness $>90\%$) genomes were retrieved from the assembly (Table S4). These genomes accounted for the microbial majority (on average 45%) for the parent material layer and showed prevalence (52% relative abundance) in topsoil community (Table 1). Among them, GT1 and GT2 showed phylogenetic affiliation to an unculturabl subgroup2 lineage within Acidobacteria (class). The draft genome of GS3 comprised 2.3 Mb of sequence in 22 scaffolds; genome annotation showed a complete sulfide oxidation pathway in GS3 (Fig. S5 and Fig. 3). GS3, GS6 and GT3 were novel strains of Thermoplasmatales, Acidothermales (Acidothermus) and Acidimicrobiales, respectively.

Fig. 3. While, Proteobacteria GS1, GS2 and GS4 were identified as the same species (ANI > 94%) (Konstantinidis and Tiedje, 2005) as Serratia marcescens WW4, Delftia acidovorans SPH-1 and Achromobacter xylosoxidans NBRC15126, respectively (Fig. 3).

3.4. Abundance and functions of retrieved bins

In the parent material layer, Proteobacteria GS1, GS2, and Eur-yarchaeota GS3 were the most prevalent populations, accounting for 40% of the community. After aerobic incubation, the abundance of GS1 almost tripled, while both muGS2 and GS3 showed a clear decline. Aerobic incubation of topsoil resulted in the enrichment of GS1, Acidobacteria GT1, and Actinobacteria GT3, especially GT1, which had relative abundance of 5.2% in the community of abandoned soil (TN) which had been subjected to long term acidification (Table 1).

Sox gene, which is responsible for the complete oxidation from sulfide to sulfate, was absent from genome bins recovered. Instead, genes (sqr) encoding sulfide-quinone reductase which can oxidize sulfide to sulfur, were commonly identified, suggesting the metabolic capacity to initiate sulfide oxidation in the ASS community (Fig. 3). Additionally, key genes (dsr and aprA) potentially enabling the reverse sulfite dissipatory reduction and potentially involved in the oxidation of sulfide to sulfate (Hipp et al., 1997), were identified in the genomes of Proteobacteria GS1 and Eur-yarchaeota GS3 but not in Proteobacteria GS2 (Fig. 3). Genes related to this pathway (sqr) clustered in a previously unreported operon in GS3 (Fig. 4). Noteworthy, we identified genes for APS reductase (Apr) and ATP sulfurylase (ATPS) in GS3 and Actinobacteria GS6 which can convert sulfite to sulfate, indicating potential interspecies metabolic cooperation with the dominant sulfide oxidizer (Proteobacteria GS1) to complete the sulfide oxidation by converting sulfite to sulfate (Fig. 5).

None of the genomes showed metabolic potential for nitrogen fixation or ammonia oxidation (Fig. 3). Abilities for complete
Fig. 3. 16S rRNA gene phylogeny and key energy conservation/adaptation pathways. Neighbor-joining phylogenetic tree (bootstrap: 1000) showing the affiliation of retrieved genomes, is produced with neighbor-joining analysis based on ClustalW alignment of 16S rRNA genes recovered in the draft genomes and representative 16S rRNA gene sequences downloaded from Silva SSU115, IMG and NCBI Ref_seq databases. Branches are colored according to corresponding phyla, namely Actinobacteria (purple), Proteobacteria (red), Acidobacteria (green) and Euryarchaeota (blue). The relative abundance of each genome in different soil samples is shown as bar chart in the middle. The following right-hand columns indicate the presence of specific genes related to metabolic pathways of sulfur oxidation (orange), sulfur reduction (yellow), nitrogen cycling (green), low pH adaptation (blue), and carbon fixation (red). Genes involved in low pH adaptation are further classified according to their proton resistant mechanism: (1) proton impermeable membrane, (2) proton buffer molecules and (3) regulatory acid resistance factors. Genes with pink asterisk are absent from all the retrieved genomes. Names and abbreviation of the key genes listed: sox enzymes (sox), reverse dissimilatory sulfite reductase enzymes (dsr), APS reductase (apr), ATP sulfurylase (ATPS), sulfide:quinone oxidoreductase (sqr); ammonium monooxygenase (amo), nitrate, nitrite, nitric oxide, and nitrous oxide reductase; urase (ure); squalene-hopene cyclase (shc), potassium transport system (kdp), phosphate transport system permease protein (pst) and glutamate, arginine and lysine decarboxylation (gadAB, adi, speA, cadA); ATP-dependent protease (cpl-PX), DNA-binding protein HU-beta (hupB), starvation-inducible DNA-binding protein (dps) and transcriptional repressor LexA (lexA); ribulose-1,5-bisphosphate carboxylase-oxygenase large subunit (rbcL) and small subunit (rbcS), CO dehydrogenase (cox), and acetyl-CoA synthase (acs). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
denitrification were commonly identified in these populations as revealed by the nitrate/nitrite reductase encoded in the genome bins, especially in *Proteobacteria* GS4, where nitric oxide (NO) and nitrous oxide (N2O) reductase were also identified in multiple copies (Table S5), suggesting its role in scavenging NO and N2O intermediates to complete denitrification. Additionally, *Proteobacteria* GS2 had multiple-copy genes of urease (*ure*), suggesting its potential ability to manipulate the pH of its environment as ammonia, a basic molecule, is released during the hydrolysis of urease (Zimmer, 2000) (Table S5).

Given the low pH of ASS, which is below 4.0 (before incubation) and 3.0 (after incubation) (Fig. 1), respectively, all retrieved genomes represented metabolic tolerance towards a low pH environment with different adaptation machineries. Proton impermeable membrane associated genes was identified in *Actinobacteria* GS6, *Acidobacteria* GT1, GT2 and *Actinobacteria* GT3, while *Proteobacteria* GS1, GS2 and GS4 relied more on the regulatory mechanism (*clpPX* and *hupB*) for proton resistance, but both pathways were absent in *Euryarchaeota* GS3 based on the analysis of binned genome (Fig. 3). Similar to other obligate chemolithotrophic sulfur-oxidizing bacteria (SOB) in Beta- and Gammaproteobacteria (Ghosh and Dam, 2009), key genes for carbon fixation via the Wood–Ljungdahl pathway (*cso* and *acs*) and reverse citric acid (TCA) cycle were present in all retrieved genomes except *Proteobacteria* GS4 (Fig. 3). In addition to this reductive acetyl-CoA pathway, *Actinobacteria* GS6 also encompassed genes encoding the key enzyme (*rbc*, also known as RuBisCo) and a full suite of genes for the Calvin–Benson–Bassham (CBB) cycle in which carbon dioxide is fixed by the energy of light, suggesting its carbon fixation potential and chemoautotrophic life style (Fig. 3).

4. Discussion

The metagenome data revealed that there was a significant difference in phylogenetic community composition between the topsoil and parent material layer. The parent material layer had a much simpler microbial community. *Thermoplasmatales* and uncultured *Actinobacteria* had higher relative abundances in the parent material layer, which could attribute to the highly acidic conditions (Fig. 1). Canonical variate analysis of the 16S rRNA genes showed that the community composition shifted significantly after incubation, and the difference was more pronounced in the parent material than in the topsoil (Fig. S1). Archaea in the parent material decreased in relative abundance during incubation (Fig. S2). Since most archaea can survive in extreme environments, such as low pH and high salinity, the aerobic environment during incubation could be an important reason for the shift of archaeal community. Our results may suggest that bacteria have a better niche than archaea under aerobic condition in ASS. A similar result was found when drought conditions exhibited significant negative effects on soil archaeal community abundance (Tian et al., 2015).

The metagenome sequencing provided the sensitivity necessary to characterize subtle changes in the abundance of functional genes. Similar to phylogenetic microbial composition, significant differences in functional genes were observed between the topsoil and parent material. The results are likely related to unique conditions found in the parent material, including high acidity (pH < 3.3) and anaerobic conditions. We also found that the functional gene composition shifted rapidly during incubation. The difference in functional gene structure between soil samples before and after incubation was much stronger than that between the topsoil and the parent material (Fig. S3). The Bray–Curtis distances calculated from taxon abundances and functional gene abundances were significantly correlated
(Fig. 5Sb), suggesting that the overall changes in functional genes were closely associated with those in phylogenetic compositions of these communities. This observation is supported by previous studies (Gilbert et al., 2010; Bryant et al., 2012), demonstrating strong correlations between metagenomic functional composition and taxonomic composition in spite of functional redundancy at the genomic level. Although individual functional genes may not necessarily be correlated with community structure, the overall functional attributes of soil microbial communities appear to be predictable across broad gradients in a restricted environment such as ASS based on the phylogenetic composition of the communities (Fierer et al., 2012). The significant correlation between functional genes and taxonomic composition may partially due to the lack of redundancy in the highly acidic soil.

The potential environmental hazard associated with ASS often results from the strong acidity, which is derived from sulfur oxidation. Oxygen was reported to be the rate-limiting factor in the oxidation of pyrite under acidic environment (Ward et al., 2004). As expected, aerobic incubation resulted in a drastic decrease in soil pH and increase in water-soluble sulfate (Fig. 1). The abundance and composition of sulfur oxidation and reduction genes shifted during this short-term incubation (Fig. 2 and Figs. S3c and 4). These results suggest that the activities of microbial communities capable of catalyzing sulfolithotrophy were induced by aerobic conditions. The distinct microbial communities in four treatments have similar sulfur transformation functions, suggesting some redundancy of functional genes and the diversity of sulfur oxidation pathways (Mackelprang et al., 2011). Functional genes related to sulfur cycling in the parent material layer and topsoil were distinct from each other (Fig. 2), and the higher gene abundances in parent material could be explained by the high sulfur content. The obligate chemolithotrophic bacteria for sulfur oxidation belonged predominantly to Beta- and Gammaproteobacteria (Ghosh and Dam, 2009). The normalized abundance of sulfur cycling genes were higher in parent material (Fig. 2), implying that sulfur cycling microbes may play an important role and dominate in parent material. Assembled genomes Proteobacteria GS1, GS2 were prevalent in parent material, with identified sulfur cycling genes in their genomes.

Genome reconstruction revealed wide-spread sulfooxidizing and the proton resistant ability within the community bins, implying that the high concentration of oxidizable sulfide (in the form of pyrite, FeS2) and decrease in pH resulted from its oxidation were the major factors shaping the microbial community structure of ASS. Consequently, via aerobic incubation which lowered the environmental pH to an extreme of below pH 3.0, we observed significant enrichment (almost tripled in relative abundance) of Proteobacteria GS1 (strain of Serratia marcescens, which is commonly found in soil) displaying concurrent reverse dissimilatory sulfate reduction potential (rDsr) and versatile low pH adaptation capacity. Moreover, the complete sulfide oxidation pathway was identified in a novel species of Thermoplasmatales (GS3). However, compared with its bacterial counterpart, this sulfur-oxidizing archaea (SOA) showed genetic incompetence of regulatory proton resistance factors of clpPX, hupP and lexA, all of which were recently reported to help host cells to survive under severe acid stress (Guazzaroni et al., 2013), causing its apparent decline in the population during aerobic incubation. Since SOB showed relatively stronger resistance towards a low pH environment, suggesting that SOB might make a larger contribution in microbial sulfooxidation in this acid sulfate soil than SOA. The low abundance of sulfur genes identified in the total community (Fig. 2) in conjunction with the rDsr/sqo-empowered sulfide oxidation pathway identified in major SOB suggests that the sulfide/sulfite oxidation in the system was facilitated by the reverse-acting sulfate reduction other than the extensively studied soil system. Previous genome interpretation of isolated strains showed sox enzymes as the central pathway in SOB of Alphaproteobacteria, while the mechanisms of sulfur oxidation in Beta- and Gammaproteobacteria have not been well established (Ghosh and Dam, 2009). Results from this study indicated that the molecular sulfide oxidation mechanism of the major Betaproteobacteria SOB from this acid sulfate soil microbiome was initialized by the reverse dissimilatory sulfate reduction pathway and facilitated by interspecies cooperation with SOA Thermoplasmatales sp. (GS3) for subsequent sulfite oxidation to complete the cycle. The near-complete genome of novel SOA species of Thermoplasmatales provided genomic evidence for the role of this lineage in sulfide oxidation, which was support by previous report of sulfite/organo-sulfonate reduction by a soil Thermoplasmata sp. (Lin et al., 2015).

In the sulfur lori of GS3, we observed co-location of sulfide:quinone oxidoreductase (sqr) immediately after a TusA-Dsre-Dsre-TusA arrangement (Fig. 4) which mediates the thiosulfate transfer in acidothermophilic sulfur and tetrathionate (S4O62−)-oxidizing Methanosphaera cuprina Ar-4 (Liu et al., 2014). The subsequent clustering of genes for membrane associated proteins and sulfite transporter ( TauE) may encode the membrane protein enabling the inner and outer cellular sulfite transportation. These observations may provide insights into the metabolism of chemolithotrophic SOA phenotypes.

Key genes involved in sulfur oxidation were identified in Proteobacteria GS1 and Actinobacteria GT3, but not in the dominant Acidobacteria GT1, indicating that, in contrast to the parent material, metabolic advantage towards sulfide/sulfite oxidation was not as evident in the topsoil community. Such discrepancy was probably determined by the lower oxidizable sulfur content in the topsoil (Table S1). Furthermore, the fragmented assembly caused by the increased community complexity of the topsoil (Table S3) may also hinder our de novo genome recovery, or on the other hand, unknown sulfur oxidation genes inaccessible by current metagenomic annotators may exist. To further resolve the possible pathways of GT1, isolation experiments were carried out using the topsoil as inoculum, however GT1 was unfortunately not among the nineteen isolates we obtained (data not shown).

The comparison of metagenomic, draft genomes and process data provided a detailed response of soil microbial communities to redox changes. It should be noted that the ex situ incubation would change the soil physical status, such as the disturbance of soil porosity thus changing the kinetics of air penetration and microbial composition. Nevertheless, the availability of oxygen should be one of the key driver here, particularly in parent material, which is reflected by the shift of the microbial communities. The shift of bacterial phylogeny and functional genes revealed the microbes potentially contributing to the acidification of ASS in response to oxidation. This study implies the potential linkage between metagenomic data and dominant biogeochemical fluxes. Results highlighted that dominant microbes in ASS have developed a unique adaptation strategy for life under harsh acidic conditions. Additionally, this study provides draft genomes of eight key species from ASS and determines their metabolic pathways of sulfur cycling. These include three novel lineages (affiliated to Thermoplasmatales, Acidothermales and Acidimicrobiales) which provide an insight into the ecology, evolution and adaptation of soil microbial communities under extreme environmental conditions.

Conflict of interest

We declare no conflict of interest.
Acknowledgements

This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences, Grant NO. XDB15020302, XDB15020301 and XDB15020402, and the National Natural Science Foundation of China (41430858 and 41525002).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.soilbio.2016.11.027.

References


Ostrobothnian acid sulfate soil. FEMS Microbiology Ecology 84, 555–563.

