Using SMDB database to reveal the sulfate reduction metabolic characteristics of microorganisms in the subtropical mangrove ecosystem

Shuming Mo [1, †], Jinhui Li [1, †], Bin Li [2], Ran Yu [1], Shiqing Nie [1], Zufan Zhang [1], Jianping Liao [3], Qiong Jiang [1], Bing Yan [2, *], and Chengjian Jiang [1, 2 *]

[1] State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, Guangxi Research Center for Microbial and Enzyme Engineering Technology, College of Life Science and Technology, Guangxi University, Nanning 530004, China.

[2] Guangxi Key Lab of Mangrove Conservation and Utilization, Guangxi Mangrove Research Center, Guangxi Academy of Sciences, Beihai 536000, China.

[3] School of Computer and Information Engineering, Nanning Normal University, Nanning 530299, China.

† These authors contributed equally to this work.

*: Corresponding Author:

Tel: +86-771-3270736; Fax: +86-771-3237873

Email: jiangcj0520@vip.163.com (CJ); gxybing@tom.com (BY)

Abstract

Motivation

Sulfate reduction is an important process in the sulfur cycle. However, the relationship between this process and the genotype of microorganisms involved in subtropical mangrove ecosystems is poorly understood. A professional and efficient gene integration database of sulfur metabolism has not been established yet.

Results

In this study, a sulfur metabolism gene integrative database (SMDB) had been constructed successfully. The database achieved high coverage, fast retrieval, and low false positives. Then the sulfate reduction by microorganisms in subtropical mangroves ecosystem had been evaluated quickly and accurately with SMDB database. Relative enrichment of sulfate adenylyltransferase indicated that environmental factors select for a partial dissimilatory sulfate reduction process. Furthermore, the sulfate reduction community compensates by producing certain sulfate-reduction genes in response to high-sulfur environments. Taxonomic assignment of dissimilatory sulfate-reduction genes revealed that Crenarchaeota and Halobacterota are completely responsible for this process. Sulfite reductase can help the community cope with the toxic sulfite produced by these Archaea phyla. Collectively, these findings demonstrated that Halobacterota and Crenarchaeota play essential roles in dissimilatory sulfate reduction.

Availability and implementation

SMDB is freely available under http://smdb.gxu.edu.cn/ and https://github.com/taylor19891213/sulfur-metabolism-gene-database.

40 **Supplementary information**

41 Supplementary data are available at Bioinformatics online.

42 **Keywords: sulfate-reduction gene families, mangrove sediment, sulfate-reduction**

43 **genotype, sulfur metabolism gene integrative database**

1 Introduction

Dissimilatory sulfate reduction, which results in the conversion of sulfate (SO_4^{2-}) to sulfide (S^{2-}), is an important reaction in the sulfur cycle (Wenk et al., 2018). The study of dissimilatory sulfate reduction can reveal the occurrence of all dissimilatory sulfate-reduction genes in a community. However, sulfate reduction lacks a complete pathway, a condition that might be a common occurrence (Bukhtiyarova et al., 2019).

The high occurrence of partial sulfate reduction implies that environmental conditions can affect gene occurrence. Gene families, including ATP sulfurylase (*sat*), ATP sulfurylase (APS) reductase (*aprAB*), and dissimilatory sulfite reductase (*dsrABC*), are involved in the canonical dissimilatory sulfate-reduction pathway (Jochum et al., 2018, Anantharaman et al., 2018). Recently, the gene family of dissimilatory sulfite reductase beta subunit (*dsrB)* has been applied to study the diversity of sulfate-reducing bacteria (SRB) (Vavourakis et al., 2019). Dissimilatory sulfate reduction is primarily driven by SRB, and a complete absence of oxygen (O_2) is vital for SRB to gain energy (Wu et al., 2017).

To the best of our knowledge, a professional sulfur database has not been constructed yet. Different databases, such as Clusters of Orthologous Groups (COG) (Galperin et al., 2015), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2016), Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups (eggNOG) (Huerta-Cepas et al., 2016), SEED subsystems (Overbeek et al., 2005), and M5nr (Wilke et al., 2012), are widely used for functional assignment of predicted genes. However, the application of these databases in analyzing sulfur cycle via shotgun

metagenome sequencing data has several obstacles. First, few of these databases contain a complete database of gene families involved in sulfur cycle. Second, the use of these databases obtains a wrong relative abundance of gene families because of similar functional domains of the genes. For example, certain gene families, such as polysulfide reductase chain A (*psrA)* and thiosulfate reductase (*phsA)*, share high sequence similarity (Stoffels et al., 2012). This condition often leads to the mismatch interpretation of potential ecological processes. Finally, searching for sulfur cycle genes in large databases is time consuming. Therefore, a specific database of genes related to the sulfur cycle pathway should be established to address these problems. This novel database can also be applied to shotgun metagenomics research.

Mangrove sediments are usually characterized as anoxic with high levels of sulfur and salt and rich in nutrients (Ferreira et al., 2010). Sulfate reduction is one of the most active processes occurring in mangrove ecosystem (Lin et al., 2019). Dissimilatory sulfate reduction drives the formation of enormous quantities of reducing sulfide. The study of sulfate reduction in mangroves is interesting and important, but this endeavor has several limitations. First, although the diversity of sulfate-reducing bacteria has been studied in mangrove ecosystems, an understanding of sulfate reduction in these ecosystems remains insufficient (Wu et al., 2019). Second, researchers have discovered culturable sulfate reduction of single strains via genomic analysis in other ecological environments (Spring et al., 2019), but similar studies have been rarely conducted in mangrove ecosystems. Finally, the relationship between sulfate reduction and the genotype of microorganisms involved in this process in mangroves are also poorly

understood. Furthermore, the environmental conditions that select for dissimilatory sulfate-reduction gene families for a more frequent reliance on sulfate reduction remain unclear.

Therefore, in this study, we created a manually managed database that primarily gathers most of the publicly available sulfur cycle gene families to address the limitations of available public databases in sulfur cycle analysis. An integrated database, namely, sulfur metabolism gene integrative database (SMDB), covering sulfur cycle gene families was established. SMDB was applied to samples from the subtropical mangrove ecosystem of Beibu Gulf in China to generate functional profiles. And then the sulfate reduction in the mangrove ecosystem had been evaluated. This study presents unique microbial features as a consequence of adapting to distinct environmental conditions to dissimilatory sulfate-reduction process.

2 Materials and methods

2.1 Sulfur metabolism gene integrative database development

A manually created integrated database was constructed to profile sulfur cycle genes from shotgun metagenomes, as the described with Tu et al. (2019), with slight modifications. The database was expected to bear the following characteristics: all sulfur gene families should be accurate and false positives should be minimized in database searching. The framework is shown in Supplementary Fig. S1.

2.1.1 Sulfur metabolism gene integrative database sources

The universal protein (UniProt) database (http://www.uniprot.org/) was utilized to

retrieve core database sequences of sulfur cycle gene families by using keywords. Other databases were used for retrieving nontarget homologous sequences. The other databases included COG (ftp://ftp.ncbi.nih.gov/pub/COG/COG2014/data), eggNOG (http://eggnogdb.embl.de/download/eggnog_4.5/), KEGG (http://www.genome.jp/kegg/), and M5nr (ftp://ftp.metagenomics.anl.gov/data/M5nr/current/M5nr.gz). In addition, the homologous sequences from RefSeq non-redundant proteins (NR) database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/) were also added in the core constructed database. These databases were selected because they are widely used in metagenome research.

2.1.2 Core database for constructing sulfur cycle genes

The KEGG and MetaCyc databases (Caspi et al., 2016) were referenced to manually retrieve gene families involved in sulfur cycle. Only those genes that had been experimentally confirmed to be involved in sulfur metabolism were collected through an extensive search of the literature. This step was followed by searching the Uniprot database with their keywords to identify the corresponding annotated sulfur metabolism gene families. For gene families with vague definitions (e.g., *cysQ* and *MET17*), the full protein name was added in the keywords to exclude sequences containing vague annotations. These datasets were retained as the core database for sulfur cycle genes.

2.1.3 Full sulfur metabolism gene integrative database construction

To reduce false positives in database searching and increase the comprehensiveness of the core database, COG, eggNOG, KEGG, M5nr, and NR databases were used to retrieve nontarget homologous sequences to construct a full database. The full sulfur

metabolism gene integrative database included the core database of sulfur cycle gene families and their nontarget homologous gene families. The sequence files of this integrative database were clustered using the CD-HIT program (version 4.6) (Fu et al., 2012) at 95% and 100% identity cutoffs. Representative sequences from this integrative database were then selected to construct SMDB. SMDB database had been deposited in https://github.com/taylor19891213/sulfur-metabolism-gene-database on January 8th, 2020. The SMDB website (http://smdb.gxu.edu.cn/) had been online since June 22th, 2020.

2.2 Case study

2.2.1 Sampling sites and sediment collection

The ShanKou mangrove sediment in Beihai City, China (21°29′25.74″N, 109°45′49.43″E) was selected as the sampling location (Supplementary Fig. S4). Rhizosphere samples (RS) from near-root were designated as RS1, RS2, and RS3. Nonrhizosphere samples (NRS) from nonrhizosphere were labelled as NRS1, NRS2, and NRS3. Nonmangrove samples (MS) were assigned as NMS1, NMS2, and NMS3. Mangrove samples (MS) included groups of RS and NRS. Several sediment properties, such as total organic carbon (TOC), total organic nitrogen (TN) content, nutrients (NH_4^+, NO_3^-), and available sulfur (AS), were determined. Redox potential (ORP) was measured in situ by using a portable ORP meter (BPH-220, Bell, China). Soil suspension was obtained by centrifugation to determine pH by using a pH meter (PHS-2C; Sanxin, China) and salinity by using a salinity meter (PAL-06S; Atago, Japan). TOC, TN, TP, AS, NH_4^+, sulfide, Fe, and NO_3^- were determined

following Chinese standard methods (State Bureau of Technical Supervision, 1994).

2.2.2 DNA extraction and high-throughput sequencing

DNA was extracted from the sediments by using the FastDNA SPIN kit for soil (MP Biomedicals, USA) according to the manufacturer's instructions. At least 6 μg DNA from each sample was submitted to the Shanghai Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China) for sequencing on an Illumina platform. The data output from each DNA sample was over 10 Gb. The metagenomic sequencing data were deposited to the NCBI SRA database under the BioProject PRJCA002311.

2.2.3 Shotgun metagenomic sequence processing and analysis

Initial quality assurance/quality control, including trimming sequencing adapters and bar codes from sequence reads, was performed. Adapter sequences were removed using SeqPrep v1.33 (https://github.com/jstjohn/SeqPrep). Moreover, sequences less than 100 bp, sequences with quality of less than 20, and reads containing an N base were removed using Sickle v1.2 (Joshi NA, 2011). Finally, clean reads were created.

For functional annotation, the clean reads were aligned against the online SMDB database website (http://smdb.gxu.edu.cn/). The blast software of the SMDB database is DIAMOND v0.9.14 (Buchfink et al., 2015) with parameters set as an e-value cutoff of 1×10^{-5} and sequence identity cutoff of 60%. The online website of the SMDB database output was converted to m8 blast format. Then, the reads were assigned as being from a particular gene if the aligned region was >25 amino acids with an identity of >60%. Gene families of dissimilatory sulfate reduction were screened out. A correlation heat map was used to visualize the composition of dissimilatory sulfate

reduction across all nine samples.

2.2.4 16S read extraction and characterization

Reads from the 16S rRNA encoding region of bacterial genomes were extracted using SortMeRNA (Kopylova et al., 2012) by running the software in filtering mode (https://github.com/biocore/sortmerna/). The reference database was compiled by Brian Bushnell from Silva database (https://www.arb-silva.de/no_cache/download/archive/release_138/Exports/). The reads extracted using SortMeRNA were clustered with vsearch (Rognes et al., 2016) with an id = 0.97. The clusters were then matched using vsearch to the Silva SSUref NR99 dataset to assign their taxonomy.

3 Results

3.1 Characteristics of sulfur metabolism gene integrative database

A total of 175 gene (sub)families covering 11 sulfur metabolism pathways, including assimilatory sulfate reduction, thiosulfate disproportionation, sulfide oxidation, dissimilatory sulfate reduction, sulfite oxidation, sulfur oxidation, sulfur reduction, tetrathionate oxidation, tetrathionate reduction, thiosulfate oxidation, and organic degradation/synthesis, were recruited in SMDB (Supplementary Table S1). SMDB obtained 115,321 and 395,737 representative sequences at 95% and 100% identity cutoffs, respectively.

To demonstrate the necessity of building a manually created sulfur metabolism gene database, we first compared the coverage and accuracy of sulfur metabolism gene

198 (sub)families in SMDB with those of existing public databases. Among the 175 gene
199 (sub)families we recruited in SMDB, only 118, 157, 144, 169, and 172 could be found
200 in these gene (sub)families in COG, eggNOG, KEGG, M5nr, and NR databases,
201 respectively (Fig. 1). The accuracy of SMDB containing sulfur gene (sub)families
202 exceeded that of public databases (Supplementary Fig. S2). We characterized sulfur
203 metabolism gene (sub)families by searching nine shotgun metagenomic datasets
204 obtained by sampling the mangrove environment against SMDB, COG, eggNOG,
205 KEGG, M5nr, and NR databases. Subsequently, the number of sulfur metabolism gene
206 families detected by searching against SMDB far exceeded that of the five other
207 databases (Supplementary Fig. S3).

208 **3.2 Abundance and diversity of sulfur (sub) gene families**

209 A total of 165 different sulfur gene (sub)families were annotated. The sulfur gene
210 (sub)families in each sample ranged from 165 (sample RS3) to 160 (sample RS3)
211 (Supplementary Table S2). The abundance of pathways showed that the organic sulfur
212 degradation/synthesis pathway in these samples was the highest (Supplementary Fig.
213 S6). The pathways of sulfur cycle gene (sub) families showed that dissimilatory sulfate
214 reduction and sulfide oxidation were significantly high in mangrove sediments ($p <$
215 0.05, Supplementary Fig. S8). However, the pathways of assimilatory sulfate reduction,
216 tetrathionate reduction, tetrathionate oxidation, sulfide oxidation, and sulfite oxidation
217 were significantly high in RS ($p < 0.05$, Supplementary Fig. S8). Hence, we focused on
218 the analysis of dissimilatory sulfate reduction in mangrove ecosystems.

219 **3.3 Bacterial community composition**

220 To study the distribution of dissimilatory sulfate reduction in microbial communities,
221 we used SortMeRNA (Kopylova et al., 2012) to extract 16S rRNA reads. The
222 taxonomic assignments of the reads assigned to 16S rRNA genes indicated that
223 members of Proteobacteria and Desulfobacterota were equally dominant, with
224 compositions ranging from 31% to 39%, across all reads from each sample (Figs. 2A
225 and 2B). Other phyla, such as Crenarchaeota and Chloroflexi, comprised approximately
226 5% of the total reads. Halobacterota accounted for approximately 1.2% of the total reads
227 (Figs. 2A and 2B).

228 **3.4 Genes for dissimilatory sulfate reduction**

229 The prevalence of dissimilatory sulfate-reduction genes at the sites was initially
230 analyzed by considering the complete dissimilatory sulfate-reduction pathway.
231 Accordingly, reads coming from genes encoding sulfate adenylyl transferase (*sat*),
232 adenylyl sulfate reductase subunit A (*aprA*), and dissimilatory sulfite reductase (*dsrA*
233 *and dsrB*) were obtained. In the MS and NMS groups, adenylyl sulfate reductase-
234 associated electron transfer protein (*qmoA*) and *dsr* subunit C (*dsrC*) were significantly
235 high in mangrove samples ($p < 0.05$, Fig. 3A). However, sulfite reductase (NADPH),
236 flavoprotein alpha-component (*cysJ*), phosphoadenosine phosphosulfate reductase
237 (*cysH*), sulfite reductase (NADPH) hemoprotein beta-component (*cysI*), sulfite
238 reductase (*sir*), and sulfite reduction-associated complex DsrMKJOP proteins *dsrP* and
239 *dsrO* were significantly high in rhizosphere samples, whereas anaerobic sulfite

240 reductase (*asrB* and *asrC*) and *dsrB* were significantly low in rhizosphere samples ($p <$
241 0.05, Fig. 3B).
242 Reads obtained from *sat* and *aprA* were more frequent than reads from *qmoA*, and *cysI*
243 (Figs. 2C and 2D). The largest number of reads was assigned to *sat*, accounting for
244 10%–12% of the total dissimilatory sulfate-reduction reads in all samples. Reads
245 assigned to *qmoA* did not exceed 1% of the total dissimilatory sulfate-reduction reads
246 (Supplementary Fig. S7).

247 **3.5 Taxonomic assignment of dissimilatory sulfate-reduction genes**

248 Proteobacteria were the predominant sources of 16S rRNA reads. Therefore, taxonomic
249 assignments were done at the phylum level to obtain good process of dissimilatory
250 sulfate-reduction genes families (Table 1). The numbers of dissimilatory sulfate-
251 reduction genes of Crenarchaeota and Halobacterota exceeded those of other taxonomic
252 assignments, whereas the remaining genes were assigned to Firmicutes, Proteobacteria,
253 and Desulfobacterota. This finding was somewhat unexpected because the 16S rRNA
254 gene analysis showed that these samples were as prevalent as Proteobacteria (Figs. 2A
255 and 2B). Proteobacteria were the sources of reads assigned to *sat,*
256 dimethylsulfoniopropionate lyase *(cysQ)*, *dsrA,* and *dsrB* (Table 1). The high
257 abundance of *aprA* was obtained from Crenarchaeota and Halobacterota (Table 1). *Sat*
258 and *dsrA* were mainly obtained from Halobacterota. In addition, *asr* and *cys* gene
259 families were obtained from Crenarchaeota, Halobacterota, and Proteobacteria (Table
260 1).

3.6 Sediment properties

To determine the environmental factors that likely shape the structure and composition of dissimilatory sulfate-reduction genes in mangrove sediments, we generated a correlation heat map (Supplementary Fig. S9). Certain sediment properties, such as pH, TOC, AS, ORP, NH_4^+, NO_3^-, TN, TP, Fe, salinity, and sulfide content, were determined (Supplementary Fig. S5). The concentrations of AS, Fe, TOC, and TN were significantly higher in the mangrove samples than those in the nonmangrove samples. Moreover, the concentrations of TOC, TN, and TP were significantly higher in the rhizosphere samples than those in the nonrhizosphere samples (Supplementary Tables S5 and S6, $p < 0.05$).

Result showed significant correlations between sediment properties and dissimilatory sulfate-reduction gene (sub) families ($p < 0.01$, Supplementary Fig. S9). pH was negatively correlated with *qmoA*. Sulfide content was negatively correlated with adenylylsulfate reductase subunit B (*aprB)*. Results of Pearson correlation analysis of the pathways are summarized in Supplementary Table S4. Results showed no significant correlation between environmental factors and the dissimilatory sulfate-reduction pathway. AS was negatively correlated with the organic degradation/synthesis pathway (Supplementary Table S4).

4 Discussion

Sulfur cycle is one of the most important biogeochemical cycles in Earth ecosystems. In this study, we developed a manually created database for the fast and accurate profiling of sulfur metabolism genes from shotgun metagenome sequencing data. The

developed SMDB addresses the bottleneck encountered in the previous public databases. First, SMDB has a precise definition for sulfur metabolism gene families. Unlike other database generations (Huerta-Cepas et al., 2016), gene families in SMDB were manually retrieved by using keywords combined with sequence similarity. Typical examples include the *phsA* and *psrA* genes that encode particle thiosulfate reductase and polysulfide reductase, respectively, but these genes share high sequence similarity (Stoffels et al., 2012). Hence, distinguishing the activities of these two enzymes during genome annotation is difficult. Second, SMDB presents a comprehensive coverage of gene families in sulfur metabolism. We aimed to provide the latest knowledge and research progress in sulfur metabolism studies, such as aerobic dimethyl sulfide (DMS) degradation (Koch and Dahl, 2018). The gene families *TST/MPST/sseA* for K01011 display considerable sequence similarity in KEGG, but they are distinct in two sulfur transferase subfamilies. Finally, SMDB considers the problem of false positives. Herein, this issue was addressed by including homologous gene families from multiple public databases. Most importantly, the small size of SMDB reduces the computational cost required to output the functional profiling of files of sulfur metabolism gene families.

We applied SMDB in characterizing sulfur metabolism gene (sub)families in mangrove sediment environment to demonstrate its potential applications in microbial ecology studies. The metagenomics survey described herein provided strong evidence that in neutral pH, anoxic, high-sulfur soil, the environment in the Beibu Gulf of China favored the enrichment of dissimilatory sulfate-reduction genotypes. AS concentrations were consistently higher in the mangrove zones than in the nonmangrove zones ($p < 0.05$, Supplementary Table S5), suggesting that the concentrations of available sulfur determine the abundance of dissimilatory sulfate-reduction reads. Spearman correlation

309 coefficients of abiotic factors and dissimilatory sulfate-reduction genes were calculated
310 by SPSS (George and Mallery, 2013). We found that sulfide content limited the
311 abundance of *aprB* (Supplementary Table S3), suggesting that sulfate reduction is
312 inhibited by sulfide content when it exceeds a certain threshold. However, our current
313 understanding of dissimilatory sulfate reduction suggests that AS contents are
314 insufficient to explain gene trends, and available O_2 is also an important factor.
315 Mangrove sediments are usually characterized as anoxic (Ferreira et al., 2010).
316 Mangrove conditions favoring dissimilatory sulfate reduction were probably of
317 sufficient duration to contribute to the patterns observed herein, but a more detailed
318 study of sediment properties should be conducted for confirmation.

319 This pattern supported the hypothesis that environmental conditions act on individual
320 genes in the pathway and favor the occurrence of partial dissimilatory sulfate reduction.
321 In all samples, the reads from *sat* occurred at higher levels than that from dissimilatory
322 sulfate-reduction genes (Figs. 2C and 2D). This finding was unexpected because
323 previous analysis of the occurrence of inorganic sulfur cycle genes in sequence
324 genomes found that *aprA* occurs at higher frequencies than the other gene-encoding
325 proteins required for dissimilatory sulfate-reduction genes (Vavourakis et al., 2019).
326 This condition is probably due to the simultaneous sulfate reduction and sulfate
327 oxidation processes (Watanabe et al., 2013). The protein product of *sat* is responsible
328 for the activation of sulfate to adenosine 5′-phosphosulfate in the first step of
329 dissimilatory sulfate reduction. Owing to the unique ecological environments of
330 subtropical mangrove ecosystems, the class Proteobacteria of Betaproteobacteria is the
331 dominant SRB group in mangrove (Varon-Lopez et al., 2014). The high abundance of

332 *sat* and *cysQ* gene families were derived from Proteobacteria possibly because special
333 environmental conditions in mangrove sediments choose to convert sulfate to APS
334 (Table 1, Figs. 2A–2D), indicating that the first step of dissimilatory sulfate reduction
335 is mainly performed by Proteobacteria. In addition, Crenarchaeota and Halobacterota
336 had relatively complete gene families of dissimilatory sulfate reduction in this study
337 (Table 1). The reads of *aprA* in mangrove zones were basically obtained from
338 Crenarchaeota and Halobacterota, suggesting that the function of aprA in mangrove
339 ecosystems may be mainly performed by these Archaea phyla. Other studies have
340 shown that Crenarchaeota and Halobacterota were involved in sulfur metabolism
341 (López-López et al., 2015, Cho et al., 2018). These findings showed that unique
342 ecological environments selected these microbial communities for dissimilatory sulfate
343 reduction.
344 Notably, sulfite is cytotoxic above a certain threshold if not rapidly metabolized and
345 can wreak havoc at the cellular and whole-plant levels. By providing the gene families
346 of *asrC, asrA*, *asrB, cysI, cysJ*, and *cysH* that reduces sulfite levels to protect the
347 members of the local community (Bender et al., 2019). These activities are required for
348 the biosynthesis of L-cysteine from sulfate (Zeghouf and Chemistry, 2000). If sulfite
349 reductase is absent, most organisms in the community would be inhibited and grow
350 slowly. The sediment properties (e.g., OPR, Fe) results in the release of sulfides from
351 sediments (Shi et al., 2016). Other studies have shown that sulfides in mangrove
352 sediments are eventually converted to L-cysteine (Lin et al., 2019).
353 We depicted the process of sulfur metabolism in mangrove sediments. Sulfur

354 metabolism in mangrove sediments includes biological and abiotic processes (Fig. 4).
355 For abiotic processes, Fe(III) is buried and acts as an oxidant for sulfide in deeper
356 sediment layers where it partly binds the produced sulfide as iron sulfide (FeS) and
357 pyrite (FeS_2) (Wasmund et al., 2017). In this study, the concentration of Fe was
358 consistently higher in the mangrove zones than that in the nonmangrove zones ($p < 0.05$,
359 Supplementary Table S5). FeS can be oxidized into various oxidation states of sulfur.
360 For biological processes, the role of microorganisms includes both oxidation and
361 reduction of sulfur compounds. Dissimilatory sulfate reduction involves the reduction
362 of sulfate to sulfide where the electron acceptor is elemental sulfur. By contrast, the
363 reduced sulfide is assimilated in proteins and amino acids in assimilatory sulfate
364 reduction (Fuentes-Lara et al., 2019). Sulfide oxidation by O_2, NO_3^-, Fe^{3+}, or Mn^{4+} as
365 electron acceptors produces sulfur. In addition, sulfur undergoes coupled oxidation and
366 reduction of sulfur compounds (thiosulfate, sulfite, and sulfur) to sulfate and sulfide.
367 DMS was derived from the degradation of organic sulfur (e.g.,
368 dimethylsulfoniopropionate).
369 The mechanism of dissimilatory sulfate-reduction metabolic pattern in mangrove
370 ecosystem was constructed (Fig. 5). The figure shows the relative abundances of gene-
371 encoding enzymes involved in dissimilatory sulfate reduction. The reduced sulfur
372 compounds were likely oxidized to sulfite by reverse dissimilatory sulfite reductase
373 (encoded by *rdsrAB*) and thiosulfate sulfur transferase (Fig. 5A). EC 2.7.7.4 (sulfate
374 adenylyltransferase) had the highest abundance among the enzymes. EC 1.8.99.2
375 (adenylyl-sulfate reductase) and EC 1.8.99.5 (dissimilatory sulfite reductase) had

higher abundance in mangrove samples than in nonmangrove samples. Hence, the results indicated that the majority of sulfate was converted into APS in dissimilatory sulfate reduction pathway owing to the special environment of the mangrove ecosystem. The model of dissimilatory sulfite reductase also supports the hypothesis that mangrove samples prefer dissimilatory sulfate reduction. In rhizosphere samples, the abundance of EC 1.8.99.5 was low, whereas that of EC 1.8.1.2 (assimilatory sulfite reductase) was high (Fig. 5B). This finding indicated that rhizosphere microorganisms are conducive to the assimilation of sulfate to L-cysteine. Moreover, this result highlighted the importance of mangrove plants in sulfur cycle.

In sum, a high-coverage database, namely, SMDB, was successfully constructed. This integrative database contains 175 gene (sub)families and covers 11 sulfur metabolism processes. SMDB was able to quickly and accurately analyze sulfur metabolism. The subtropical mangrove ecosystem favored dissimilation sulfate reduction. Proteobacteria convert sulfate to APS owing to environmental conditions in mangrove sediments. However, Crenarchaeota and Halobacterota were found to be responsible for dissimilatory sulfate reduction. This study provides a theoretical basis for the mechanism of sulfur cycle metabolism in subtropical mangrove wetland ecosystems.

Acknowledgements

The authors would like to thank the editors and the reviewers for the valuable suggestions to improve the manuscript.

Funding

This research was supported by the Science and Technology Basic Resources Investigation Program of China (Grant No. 2017FY100704), the Natural Science Fund for Distinguished Young Scholars of Guangxi Zhuang Autonomous Region of China (Grant No. 2019GXNSFFA245011), the Natural Science Foundation of Guangxi Zhuang Autonomous Region of China (Grant No. 2018GXNSFAA050090, 2017GXNSFBA198101).

Conflicts of Interest: The authors declare no conflict of interest.

References

Anantharaman, K. et al. (2018) Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. ISME J, 12, 1715-1728.

Bender, D. et al. (2019) Impaired mitochondrial maturation of sulfite oxidase in a patient with severe sulfite oxidase deficiency. Hum Mol Genet, 28, 2885-2899.

Buchfink, B. et al. (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods, 12, 59-60.

Bukhtiyarova, P.A. et al. (2019) Isolation, characterization, and genome insights into an anaerobic sulfidogenic Tissierella bacterium from Cu-bearing coins. Anaerobe, 56, 66-77.

Caspi, R., et al. (2016) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases, Nucleic Acids Res, 44(D1), D471-D480.

421 State Bureau of Technical Supervision. (1996) Standardization Administration of the
422 People's Republic of China. GB/T 4459.6-1996. (in Chinese).
423 Cho, E.S. et al. (2018) *Haloplanus rallus sp.* nov., a halophilic archaeon isolated from
424 crude solar salt. Int J Syst Evol Microbiol, 68, 3226-3231.
425 Ferreira, T.O. et al. (2010) Spatial patterns of soil attributes and components in a
426 mangrove system in Southeast Brazil (Sao Paulo). J Soils Sediments, 10, 995-1006.
427 Fu, L. et al. (2012) CD-HIT: accelerated for clustering the next-generation sequencing
428 data. Bioinformatics, 28, 3150-3152.
429 Fuentes-Lara, L.O. et al. (2019) From Elemental Sulfur to Hydrogen Sulfide in
430 Agricultural Soils and Plants. Molecules, 24, 2282.
431 Galperin, M.Y. et al. (2015) Expanded microbial genome coverage and improved
432 protein family annotation in the COG database. Nucleic Acids Res, 43, D261-269.
433 George, D. and Mallery, P. (2013) IBM SPSS Statistics 21 Step by Step: A Simple Guide
434 and Reference. Pearson.
435 Hausmann, B. et al. (2019) Draft Genome Sequence of *Desulfosporosinus sp.* Strain
436 Sb-LF, Isolated from an Acidic Peatland in Germany. Microbiol Resour Announc, 8,
437 e00428-19.
438 Huerta-Cepas, J. et al. (2016) eggNOG 4.5: a hierarchical orthology framework with
439 improved functional annotations for eukaryotic, prokaryotic and viral sequences.
440 Nucleic Acids Res, 44, D286-D293.
441 Jochum, L.M. et al. (2018) Single-Cell Genomics Reveals a Diverse Metabolic
442 Potential of Uncultivated Desulfatiglans-Related Deltaproteobacteria Widely

443 Distributed in Marine Sediment. Front Microbiol, 9, 2038.

444 Jorgensen, B.B. et al. (2019) The Biogeochemical Sulfur Cycle of Marine Sediments.
445 Front Microbiol, 10, 849.

446 Joshi NA, F.J. (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool
447 for FastQ files [Software]. Available at https://github.com/najoshi/sickle.

448 Kanehisa, M. et al. (2016) KEGG as a reference resource for gene and protein
449 annotation. Nucleic Acids Res, 44, D457-D462.

450 Koch, T. and Dahl, C. (2018) A novel bacterial sulfur oxidation pathway provides a new
451 link between the cycles of organic and inorganic sulfur compounds. ISME J, 12, 2479-
452 2491.

453 Kopylova, E. et al. (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs
454 in metatranscriptomic data. Bioinformatics, 28, 3211-3217.

455 Lin, X.L. et al. (2019) Mangrove Sediment Microbiome: Adaptive Microbial
456 Assemblages and Their Routed Biogeochemical Processes in Yunxiao Mangrove
457 National Nature Reserve, China. Microb Ecol, 78, 57-69.

458 López-López, O. et al. (2015) Metagenomics of an Alkaline Hot Spring in Galicia
459 (Spain): Microbial Diversity Analysis and Screening for Novel Lipolytic Enzymes.
460 Front Microbiol, 6, 1291.

461 Overbeek, R. et al. (2005) The subsystems approach to genome annotation and its use
462 in the project to annotate 1000 genomes. Nucleic Acids Research, 33, 5691-5702.

463 Rognes, T. et al. (2016) VSEARCH: a versatile open source tool for metagenomics.
464 PeerJ, 4, e2584.

465 Santos, A.A. et al. (2015) A protein trisulfide couples dissimilatory sulfate reduction to
466 energy conservation. Science, 350, 1541-1545.
467 Shi, J.C. et al. (2016) The Variation Characteristic of Sulfides and VOSc in a Source
468 Water Reservoir and Its Control Using a Water-Lifting Aerator. Int J Environ Res Public
469 Health, 13, 427.
470 Spring, S. et al. (2019) Sulfate-Reducing Bacteria That Produce Exopolymers Thrive
471 in the Calcifying Zone of a Hypersaline Cyanobacterial Mat. Front Microbiol, 10, 862.
472 Stoffels, L. et al. (2012) Thiosulfate Reduction in Salmonella enterica Is Driven by the
473 Proton Motive Force. J Bacteriol, 194, 475-485.
474 Tu, Q.C. et al. (2019) NCycDB: a curated integrative database for fast and accurate
475 metagenomic profiling of nitrogen cycling genes. Bioinformatics, 35, 1040-1048.
476 Varon-Lopez, M. et al. (2014) Sulphur-oxidizing and sulphate-reducing communities
477 in Brazilian mangrove sediments. Environ Microbiol, 16, 845-855.
478 Vavourakis, C.D. et al. (2019) Metagenomes and metatranscriptomes shed new light on
479 the microbial-mediated sulfur cycle in a Siberian soda lake. BMC Biol, 17:69.
480 Wasmund, K. et al. (2017) The life sulfuric: microbial ecology of sulfur cycling in
481 marine sediments. Environ Microbiol Rep, 9, 323-344.
482 Watanabe, T. et al. (2013) Diversity of sulfur-cycle prokaryotes in freshwater lake
483 sediments investigated using *aprA* as the functional marker gene. Syst Appl Microbiol,
484 36, 436-443.
485 Wenk, C.B. et al. (2018) Electron carriers in microbial sulfate reduction inferred from
486 experimental and environmental sulfur isotope fractionations. ISME J, 12, 495-507.

Wilke, A. et al. (2012) The M5nr: a novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. BMC Bioinformatics 13, 141.

Wu, J. et al. (2017) Oxygen Reduction Reaction Affected by Sulfate-Reducing Bacteria: Different Roles of Bacterial Cells and Metabolites. Indian J Microbiol, 57, 344-350.

Wu, S. et al. (2019) Depth-related change of sulfate-reducing bacteria community in mangrove sediments: The influence of heavy metal contamination. Mar Pollut Bull, 140, 443-450.

Chemistry, M. et al. (2000) A Simplifed Functional Version of the *Escherichia coli* Sulfite Reductase. J Biol Chem, 275, 37651-37656.

Table 1. Taxonomic assignments of dissimilatory sulfate-reduction genes.

Phylum	RS1	RS2	RS3	NRS1	NRS2	NRS3	NMS1	NMS2	NMS3
Crenarchaeota	*aprA;asrA ; dsrK; dsrN*	*aprB; cysJ; cysN; dsrB; nrnA*	*dsrN; nrnA*	*aprA;dsrB; dsrK*	*cysJ; nrnA*	*cysN; cysQ*	*asrC; cysH; cysN; nrnA*	*asrB; asrC; cysD; cysH; cysQ; dsrK; nrnA; sat*	*asrB; cysI; nrnA*
Halobacterota	*cysJ*	*nrnA; asrA; asrC; cysJ; dsrA; dsrM; sat*	*aprA;aprB; cysD;dsrA; dsrB; sat*	*aprA;cysQ; dsrA;dsrB; sat*	*aprA; asrA; cysD; cysH*	*dsrA;dsrM; nrnA*	*asrC; cysD; cysN;cysQ; dsrK; dsrM; dsrO; nrnA; sat*	*asrA; asrB; asrC; cysI; cysN; cysQ; dsrA; dsrB; dsrM; dsrP; nrnA; sat*	*dsrP*
Amorphea	NA	*cysN*	cysD	NA	*cysN; cysNC*	NA	NA	*aprA; dsrA*	NA
Firmicutes	NA	*cysN; dsrB*	NA	NA	*cysN*	NA	NA	*cysH; cysNC*	NA
Chloroflexi	NA	NA	*dsrN*	NA	NA	NA	*dsrA*	NA	NA
Iainarchaeota	NA	NA	NA	*asrA*	NA	NA	NA	NA	NA
Proteobacteria	NA	NA	NA	*cysQ*	*sat*	NA	*asrC*	*dsrA; dsrB; dsrM*	NA
Verrucomicrobiota	NA	NA	NA	NA	*nrnA*	NA	NA	NA	NA
Thermoplasmatota	NA	NA	NA	NA	NA	NA	*sat*	*dsrN*	NA
Actinobacteriota	NA	NA	NA	NA	NA	NA	NA	*asrC; cysD; cysQ; dsrB; nrnA*	NA
Bacteroidota	NA	NA	NA	NA	NA	NA	NA	*aprA; asrA; sir*	NA
Deferrisomatota	NA	NA	NA	NA	NA	NA	NA	*aprA*	NA
Desulfobacterota	NA	NA	NA	NA	NA	NA	NA	*cysQ; sat; sir*	NA
Euryarchaeota	NA	NA	NA	NA	NA	NA	NA	*aprA; dsrK*	NA
Planctomycetota	NA	NA	NA	NA	NA	NA	NA	*aprB*	NA
Spirochaetota	NA	NA	NA	NA	NA	NA	NA	*asrB*	*dsrM*
Zixibacteria	NA	NA	NA	NA	NA	NA	NA	*dsrJ*	NA
Gemmatimonadota	NA	NA	NA	NA	NA	NA	NA	NA	*asrB*

499 **Figure Legends**

500 Figure 1. Sulfur cycle gene (sub)families were detected using SMDB, NR, M5nr,
501 KEGG, eggNOG, COG databases.

502 Figure 2. Comparison of taxonomic assignments of 16S reads extracted from the
503 metagenomes according to A) mangrove and nonmangrove or B) rhizosphere and
504 nonrhizosphere groups. Comparison of read numbers of genes associated with
505 dissimilatory sulfate reduction grouped into C) mangroves and nonmangrove and D)
506 rhizosphere and nonrhizosphere. Rates are expressed as the relative frequency of reads
507 for each gene multiplied by the average of the total read counts for all samples. Error
508 bars represent 95% confidence interval, and statistically significant groupings are
509 indicated by different letters above the bars ($p < 0.05$).

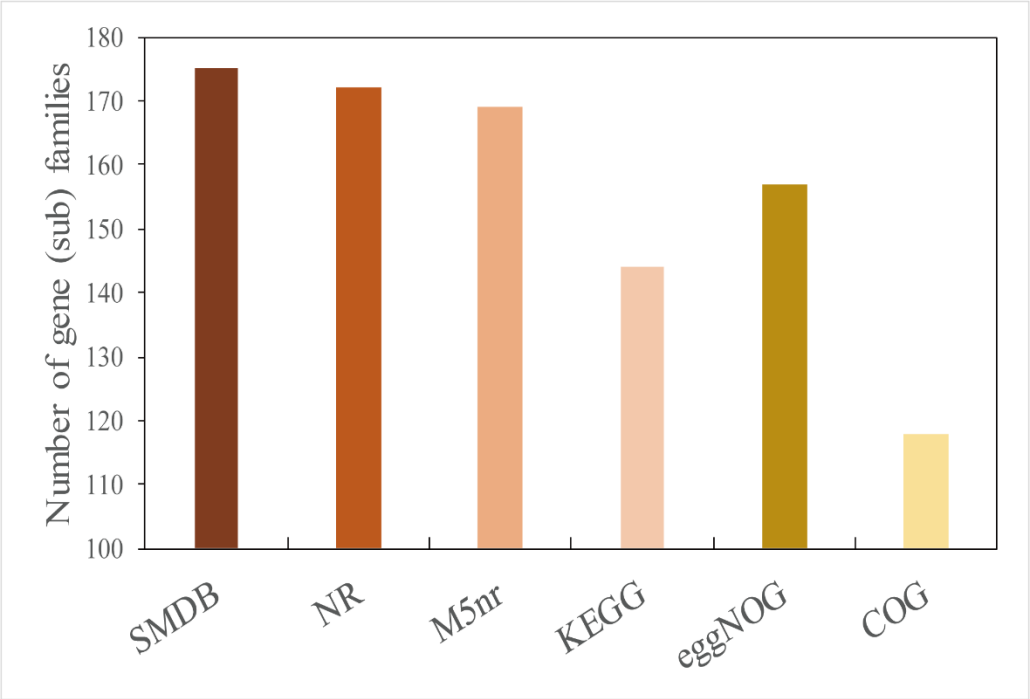
510 Figure 3. Significant differences in sulfur (sub) gene families between mangrove and
511 nonmangrove sediments (A) and rhizosphere and nonrhizosphere sediments (B).

512 Figure 4. Conceptual depiction of sulfur metabolism in mangrove ecosystem, including
513 reactions of inorganic and organic sulfur compounds and transformations of sulfur
514 compounds of intermediate oxidation states (sulfur cycle intermediates, SCI). Inorganic
515 sulfur compounds include oxidation and reduction of sulfur compounds. Black lines
516 depict biological sulfur transformations by microorganisms. Green lines depict abiotic
517 reaction-mediated sulfur transformations to pyrite (FeS_2). Sulfur compounds are
518 depicted within yellow eclipses.

519 Figure 5. Model for the pathway of dissimilatory sulfate reduction. A) Relative
520 abundances of the gene-encoding enzymes involved in dissimilatory sulfate reduction.

521 The EC numbers of enzymes are boxed. The relative abundances of enzymes in the two
522 samples are shown in the nearby color bar, in which the two segments from left to right
523 represent MS and NMS, respectively. B) Relative abundances of the gene-encoding
524 enzymes involved in dissimilatory sulfate reduction. The EC numbers of enzymes are
525 boxed. The relative abundances of genes in the two samples are shown in the nearby
526 color bar, in which the two segments from left to right represent RS and NRS,
527 respectively.

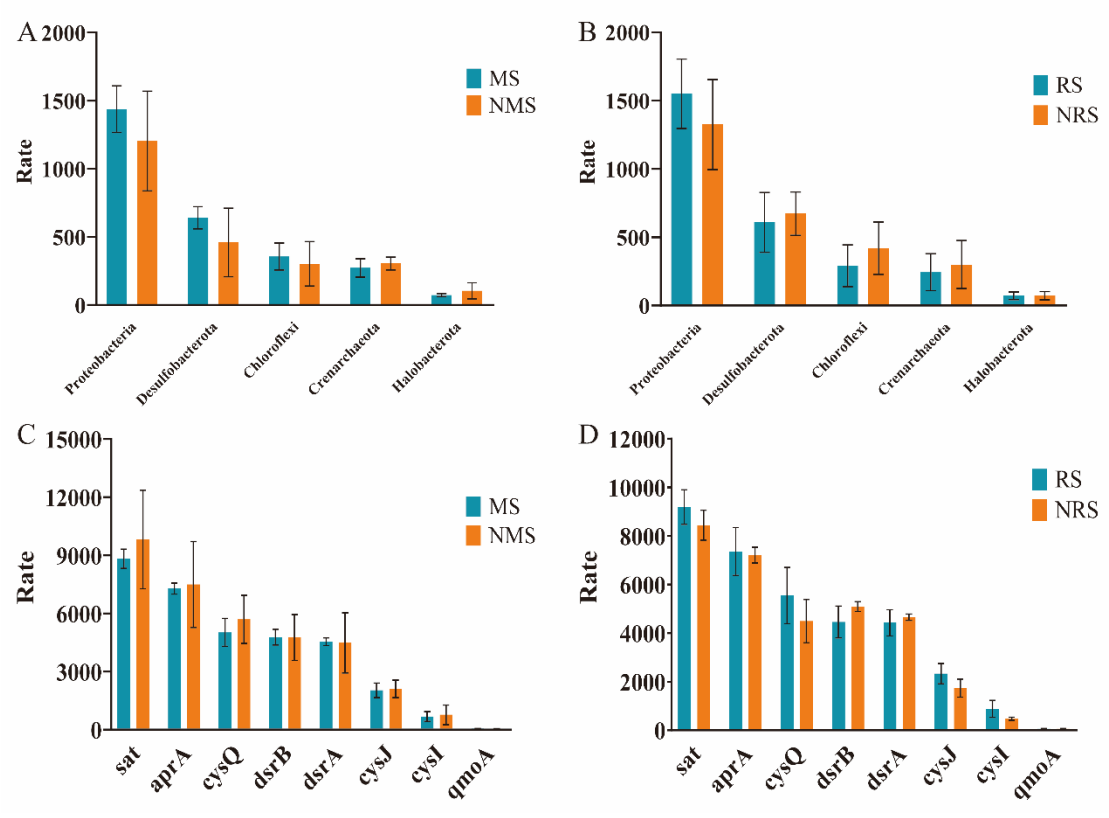
528 **Figure 1**



529

530

Figure 2

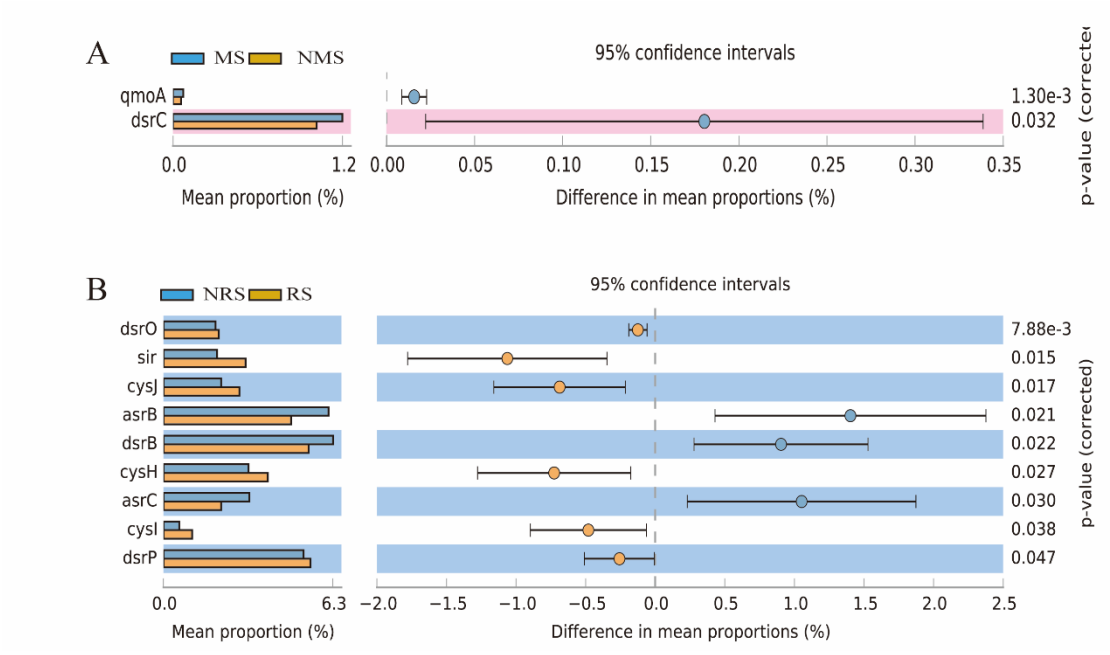


531

532

533

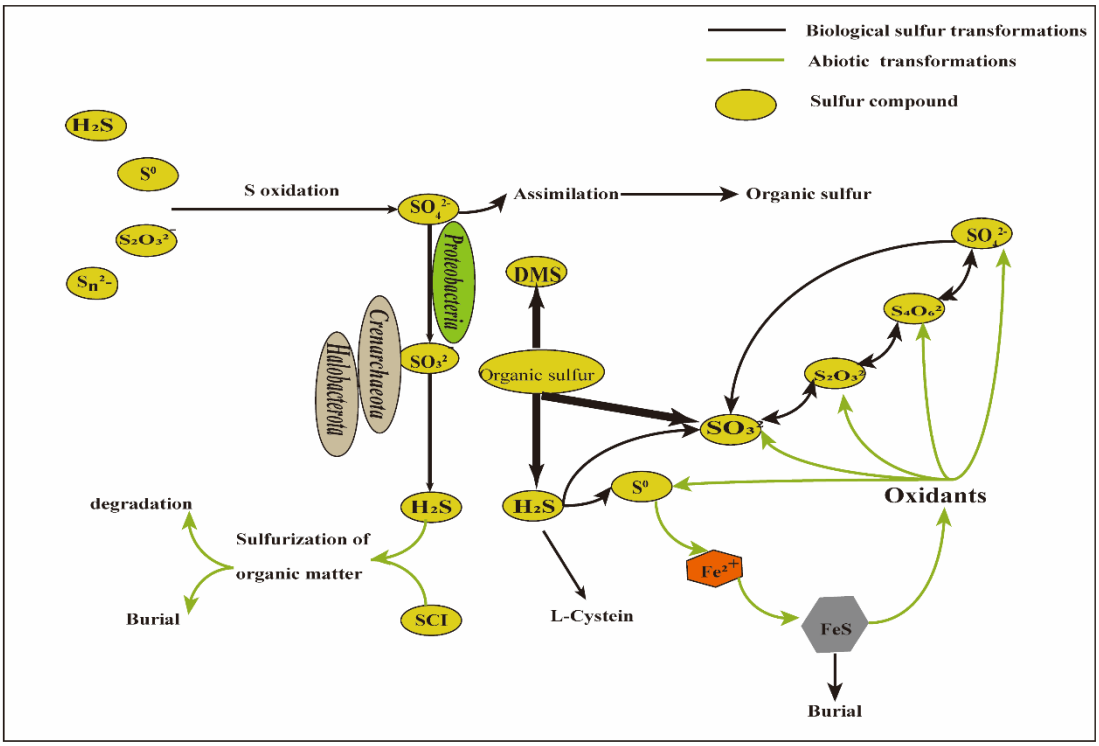
Figure 3



534

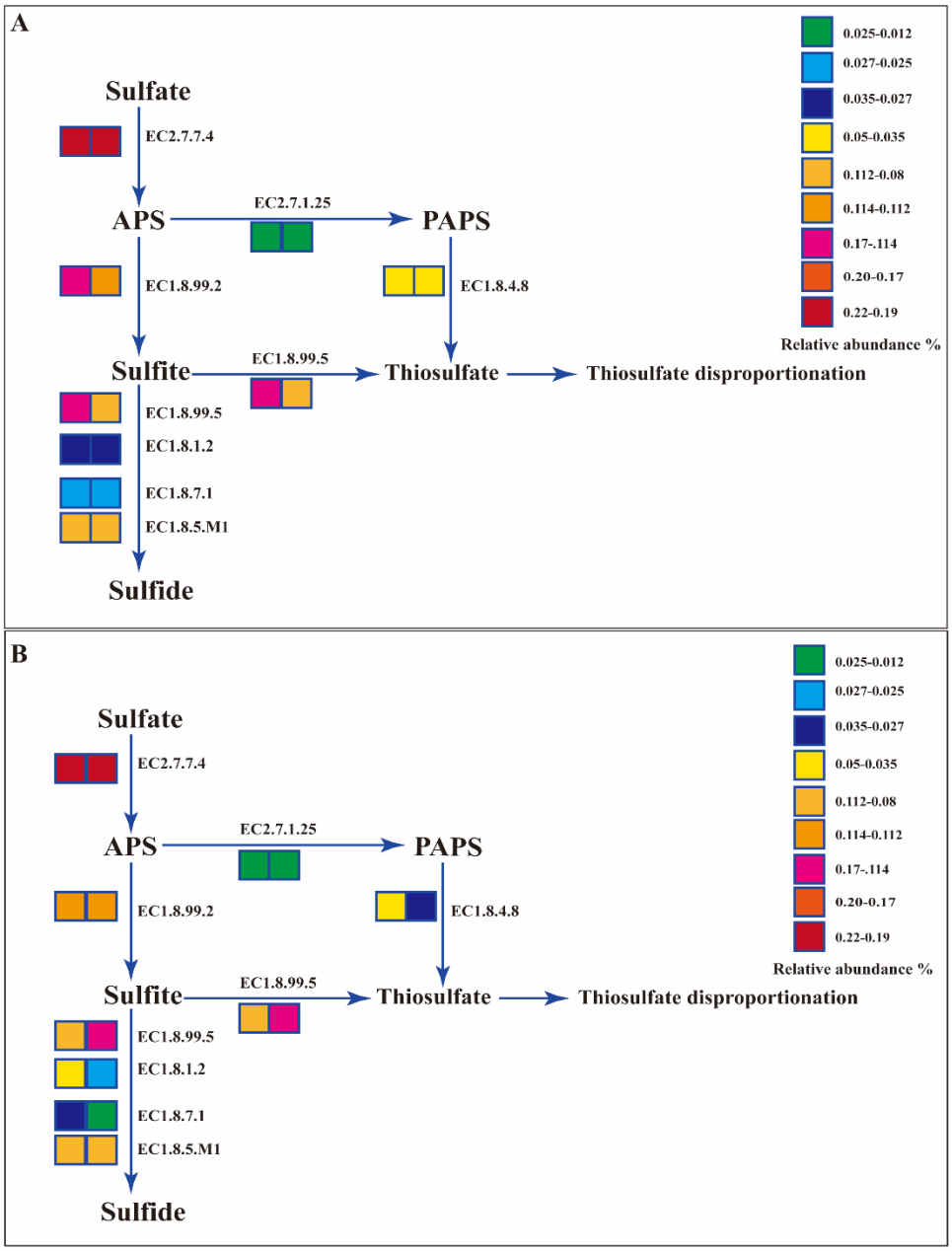
535

Figure 4


Biological sulfur transformations
Abiotic transformations
Sulfur compound
H2S
S0
S2O3 2-
Sn 2-
S oxidation
SO4 2-
Assimilation
Organic sulfur
Proteobacteria
Crenarchaeota
Halobacteroia
SO3 2-
DMS
Organic sulfur
SO4 2-
S4O6 2-
S2O5 2-
SO3 2-
Oxidants
degradation
H2S
H2S
S0
Sulfurization of
organic matter
Burial
SCI
L-Cystein
Fe2+
FeS
Burial

539

Figure 5



540