

1 **Chemolithoautotrophic diazotrophs dominate dark nitrogen fixation**
2 **in mangrove sediments**

3 Shasha Wang^{1#}, Lijing Jiang^{1#*}, Zhuoming Zhao¹, Zhen Chen¹, Jun Wang¹, Karine
4 Alain³, Liang Cui¹, Yangsheng Zhong¹, Yongyi Peng¹, Qiliang Lai¹, Xiyang Dong^{1,2*}
5 and Zongze Shao^{1,2*}

6 ¹ Key Laboratory of Marine Genetic Resources, Third Institute of Oceanography,
7 Ministry of Natural Resources of China; State Key Laboratory Breeding Base of
8 Marine Genetic Resources; Fujian Key Laboratory of Marine Genetic Resources;
9 Sino-French Laboratory of Deep-Sea Microbiology (MicrobSea), Xiamen 361005, PR
10 China

11 ² Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai
12 519000, PR China

13 ³ Univ Brest, CNRS, Ifremer, EMR6002 BIOMEX, Biologie Interactions et
14 adaptations des Organismes en Milieu EXtrême, IRP 1211 MicrobSea, F-29280
15 Plouzané, France

16
17 # These authors contributed equally to this work.

18 * Corresponding authors: Zongze Shao, Key Laboratory of Marine Genetic Resources,
19 Third Institute of Oceanography, Ministry of Natural Resources of China, 186 Daxue
20 Road, Fujian 361005, China. Email: shaozz@163.com, Xiyang Dong, Key Laboratory
21 of Marine Genetic Resources, Third Institute of Oceanography, Ministry of Natural
22 Resources of China, 186 Daxue Road, Fujian 361005, China. Email:
23 dongxiyang@tio.org.cn, and Lijing Jiang, Key Laboratory of Marine Genetic
24 Resources, Third Institute of Oceanography, Ministry of Natural Resources of China,
25 186 Daxue Road, Fujian 361005, China. Email: jianglijing@tio.org.cn

26 Running title: Chemolithoautotrophic diazotrophs in mangrove sediments

Supplementary Figures

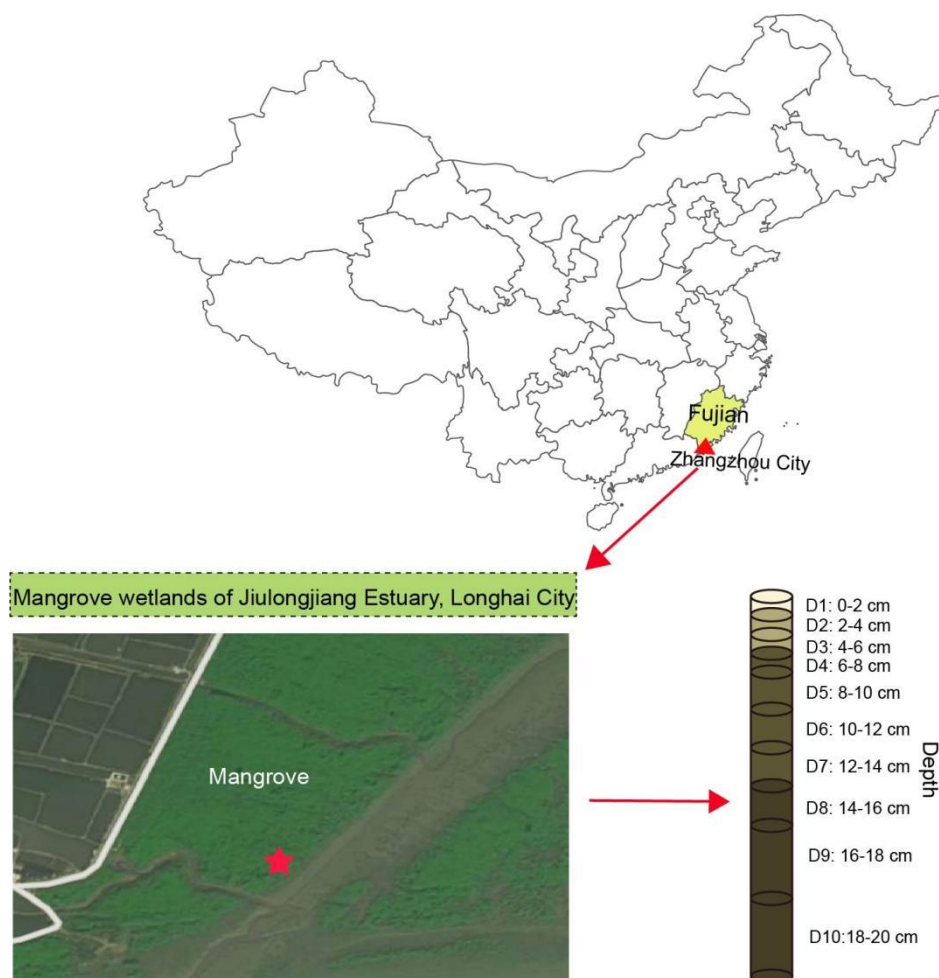


Fig. S1. Location of the sampling habitats at the Zhangzhou Mangrove wetland, China, which is dominated by *Kandelia obovata*. In this study site, one sediment core was collected using a custom-made sampler with a 20-cm depth and a 11-cm diameter. The sediment core was subdivided into 10-depths: 0-2, 2-4, 4-6, 6-8, 8-10, 10-12, 12-14, 14-16, 16-18, and 18-20 cm.

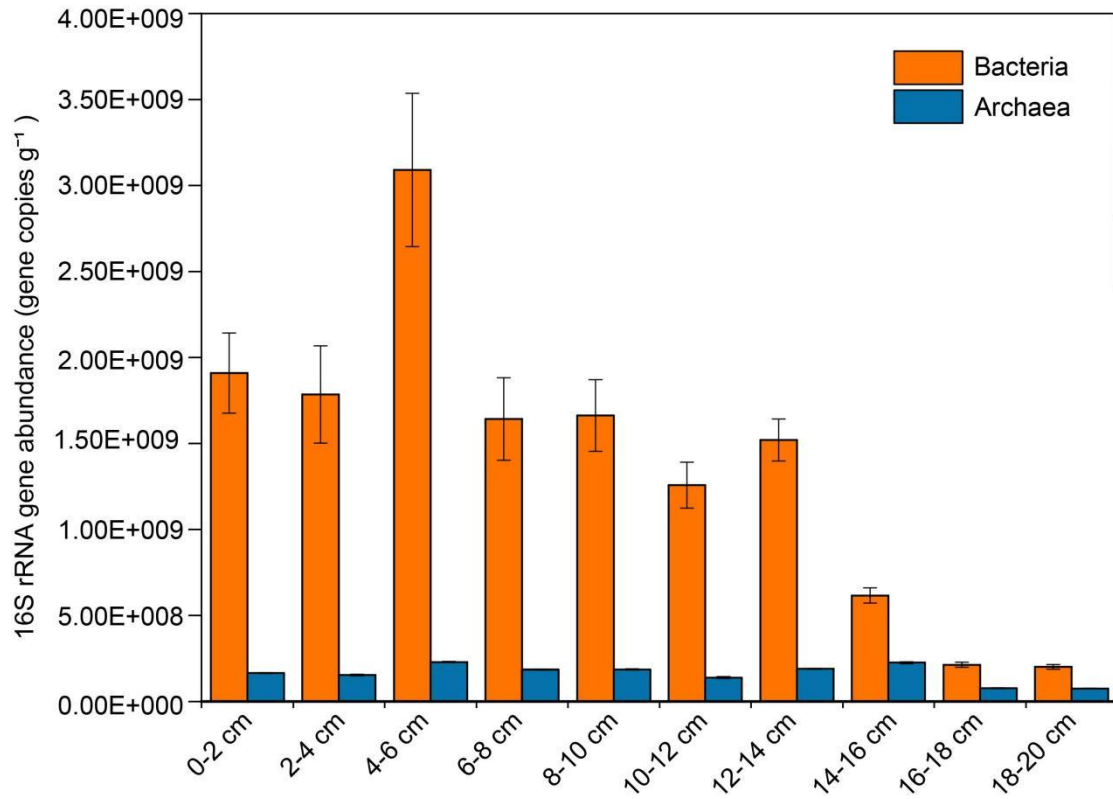


Fig. S2. Quantitative results of the abundance of bacterial and archaeal 16S rRNA gene from ten mangrove sediment samples. Standard deviations are indicated by error bars.

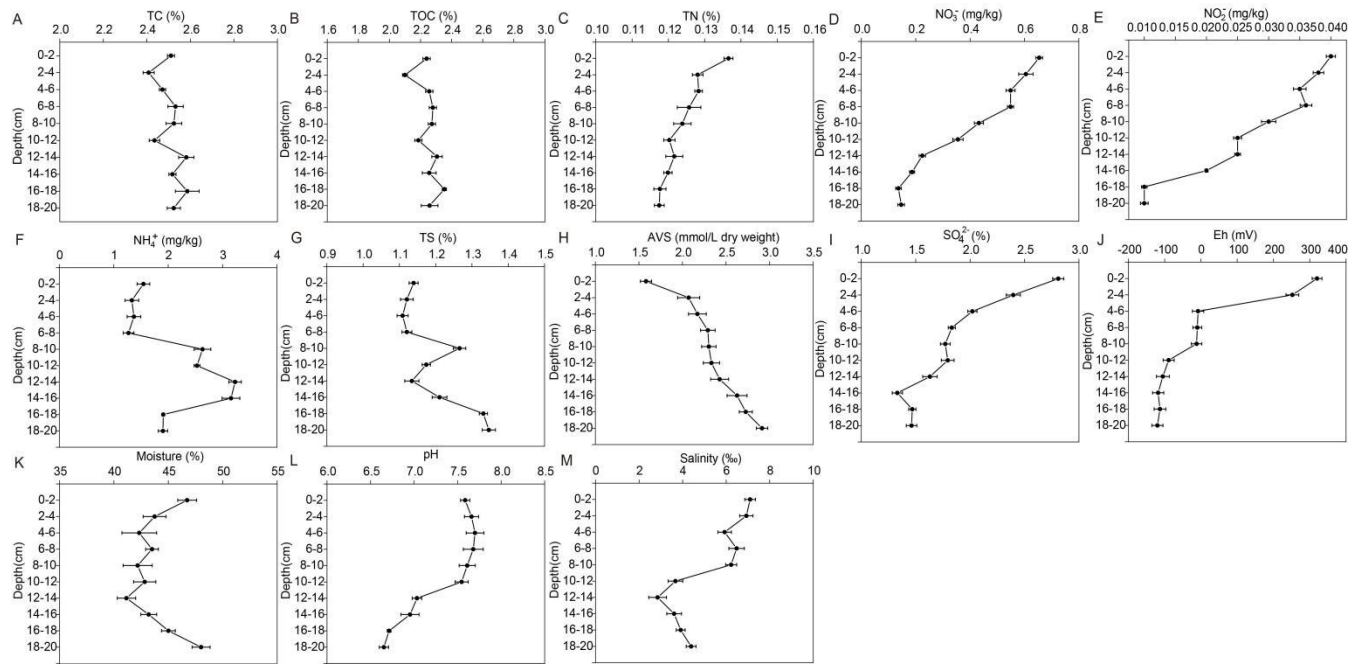


Fig. S3. Vertical distribution of various physicochemical characteristics across the depths of mangrove sediments. Data include the concentrations of TC (A), TOC (B), TN (C), NO_3^- (D), NO_2^- (E), NH_4^+ (F), TS (G), AVS (H), SO_4^{2-} (I), Eh (J), moisture (K), pH (L), and salinity (M). Error bars represent standard deviation.

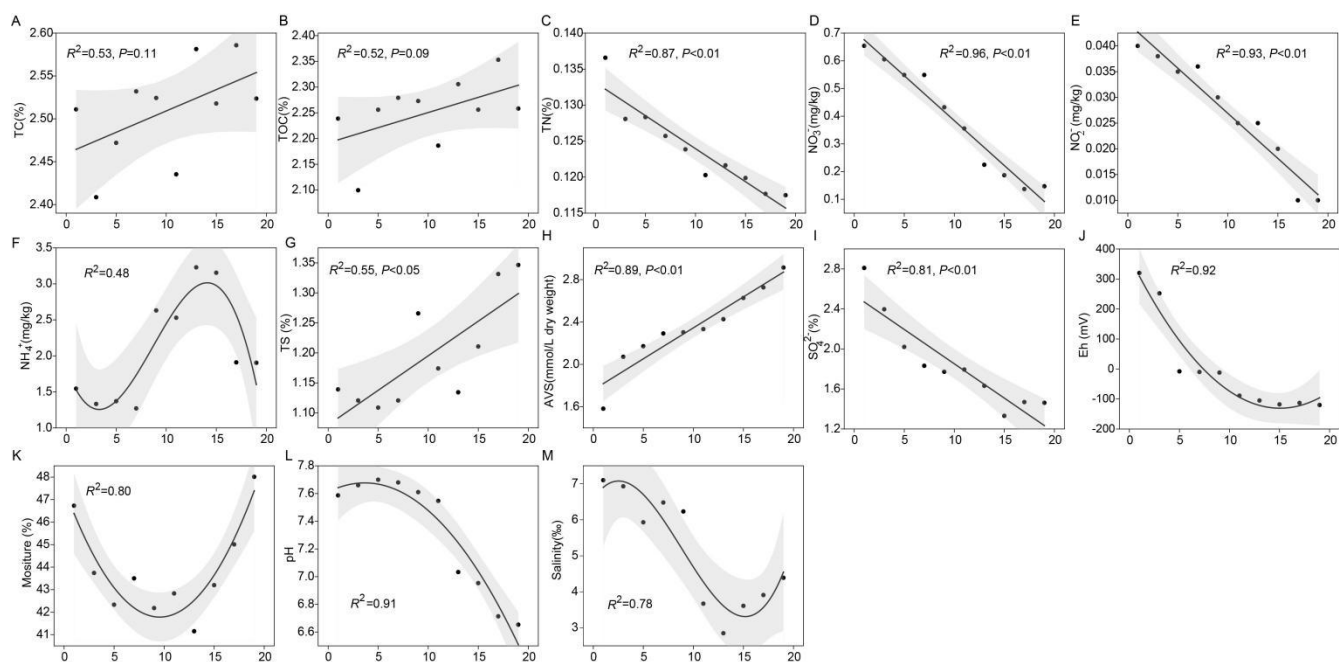


Fig. S4. Relationships between physicochemical properties and depth of mangrove sediments. Curve regression analysis of sediments properties TC (A), TOC (B), TN (C), NO_3^- (D), NO_2^- (E), NH_4^+ (F), TS (G), AVS (H), SO_4^{2-} (I), Eh (J), moisture (K), pH (L), and salinity (M) and depths. Black lines and gray shaded areas represent linear regressions and 95% confidence intervals, respectively. R^2 was obtained by linear regression analysis and P was obtained by Pearson's correlation analysis.

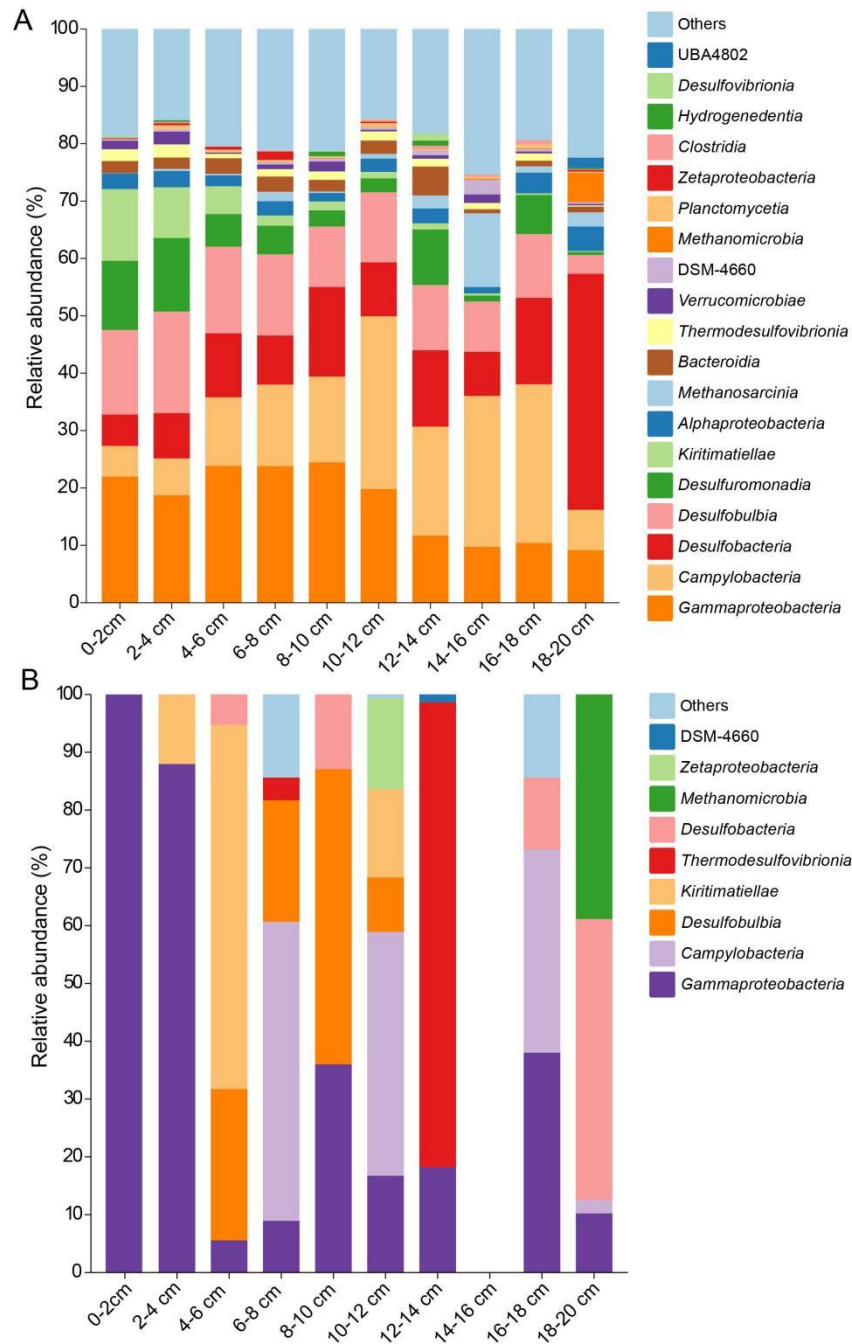


Fig. S5. Relative abundances of *nifH* genes at the class level at the metagenomics (A) and metatranscriptomics (B) from different mangrove sediment samples.

Different taxonomic groups are represented by different colors. “Other” represents the unknown classification of the communities.

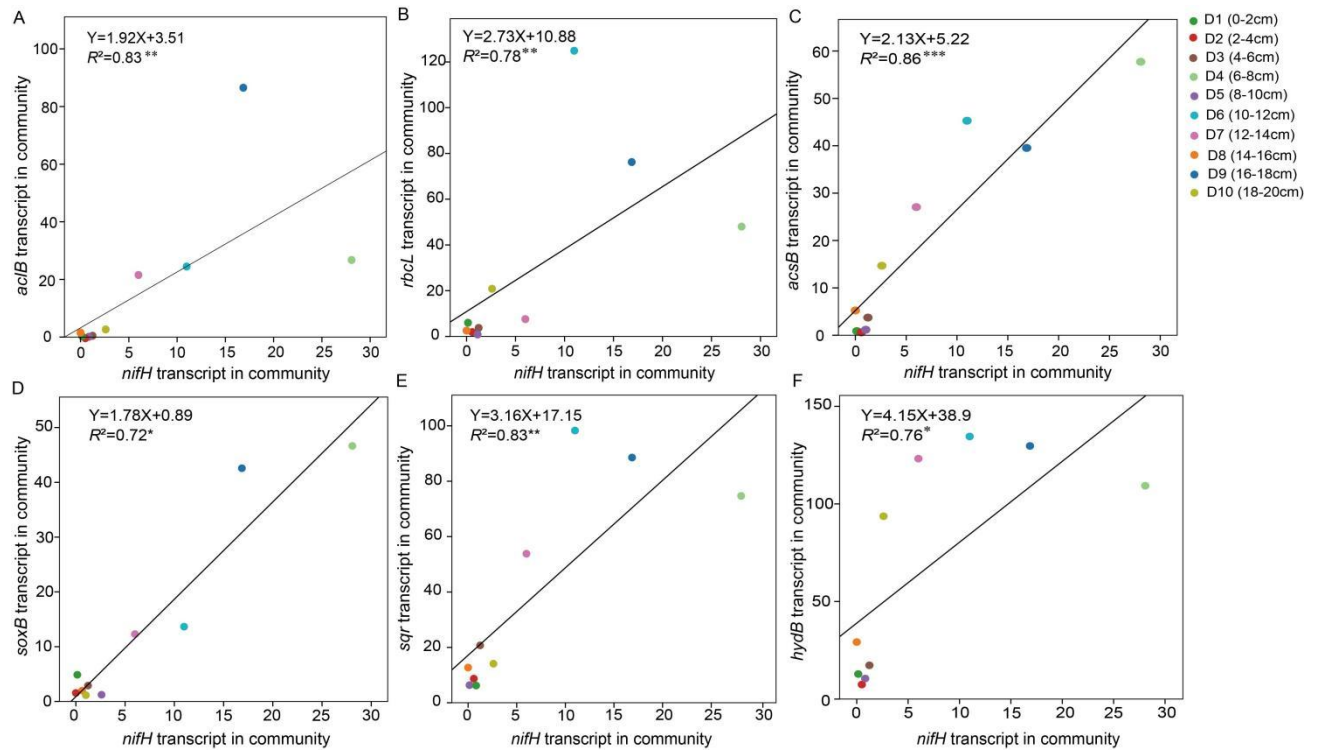


Fig. S6. Correlations between the transcriptional expression of *nifH* gene and other functional genes including *aclB*, *rbcL*, *acsB*, *sqr*, *soxB*, and *hydB* in mangrove sediments. Each point represents an individual sample. The solid line indicates the linear regression of the *nifH* gene and the functional genes. R^2 was obtained by linear regression analysis and significance levels are denoted with * ($0.01 < P < 0.05$), ** ($0.001 < P < 0.01$), and *** ($P < 0.001$).

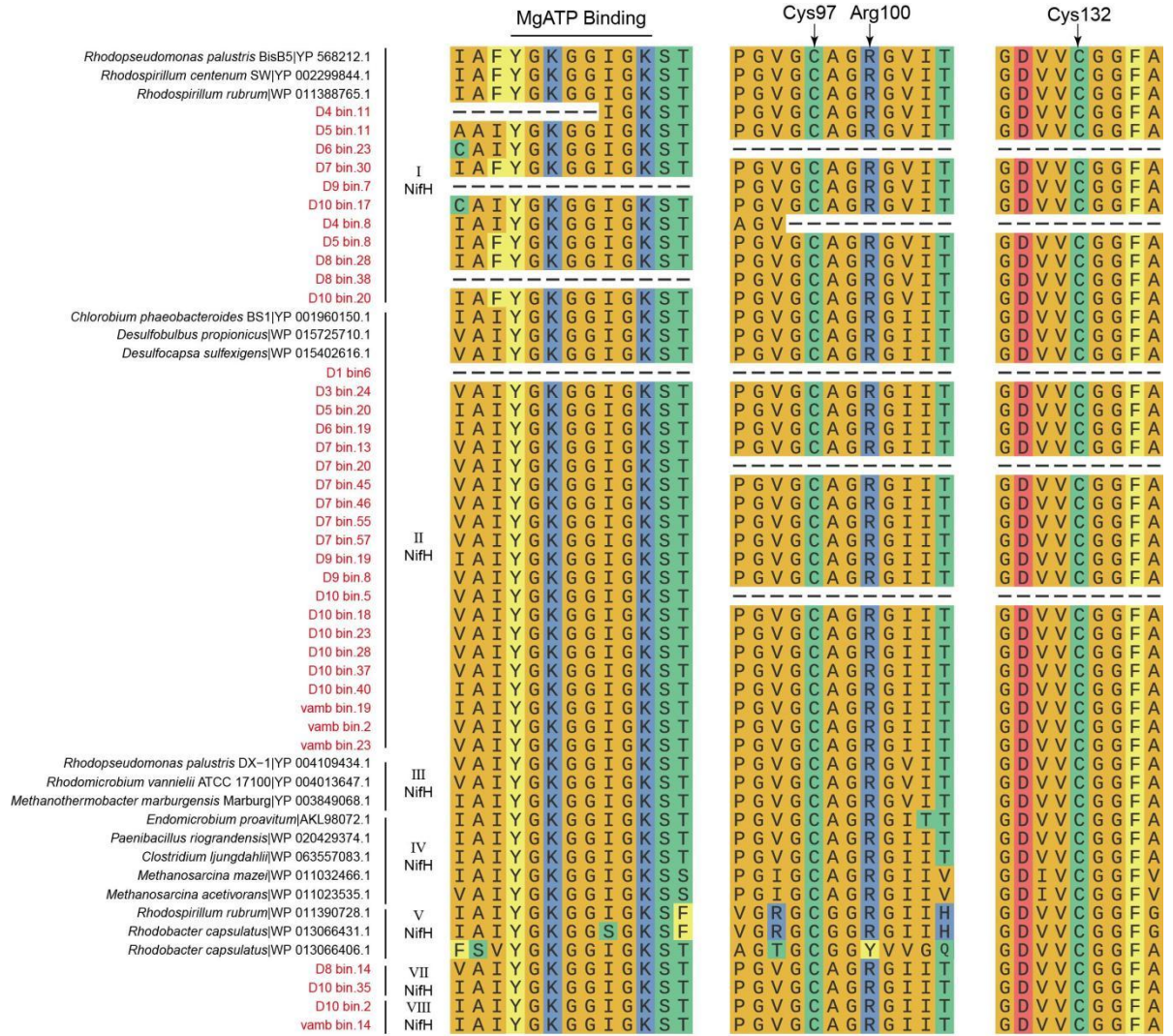


Fig. S7. NifH superfamily amino acid alignment. Multiple alignments of NifH superfamily sequences are shown in the region of active site residues responsible for MgATP binding and hydrolysis Fe₄-S₄ iron sulfur cluster binding.

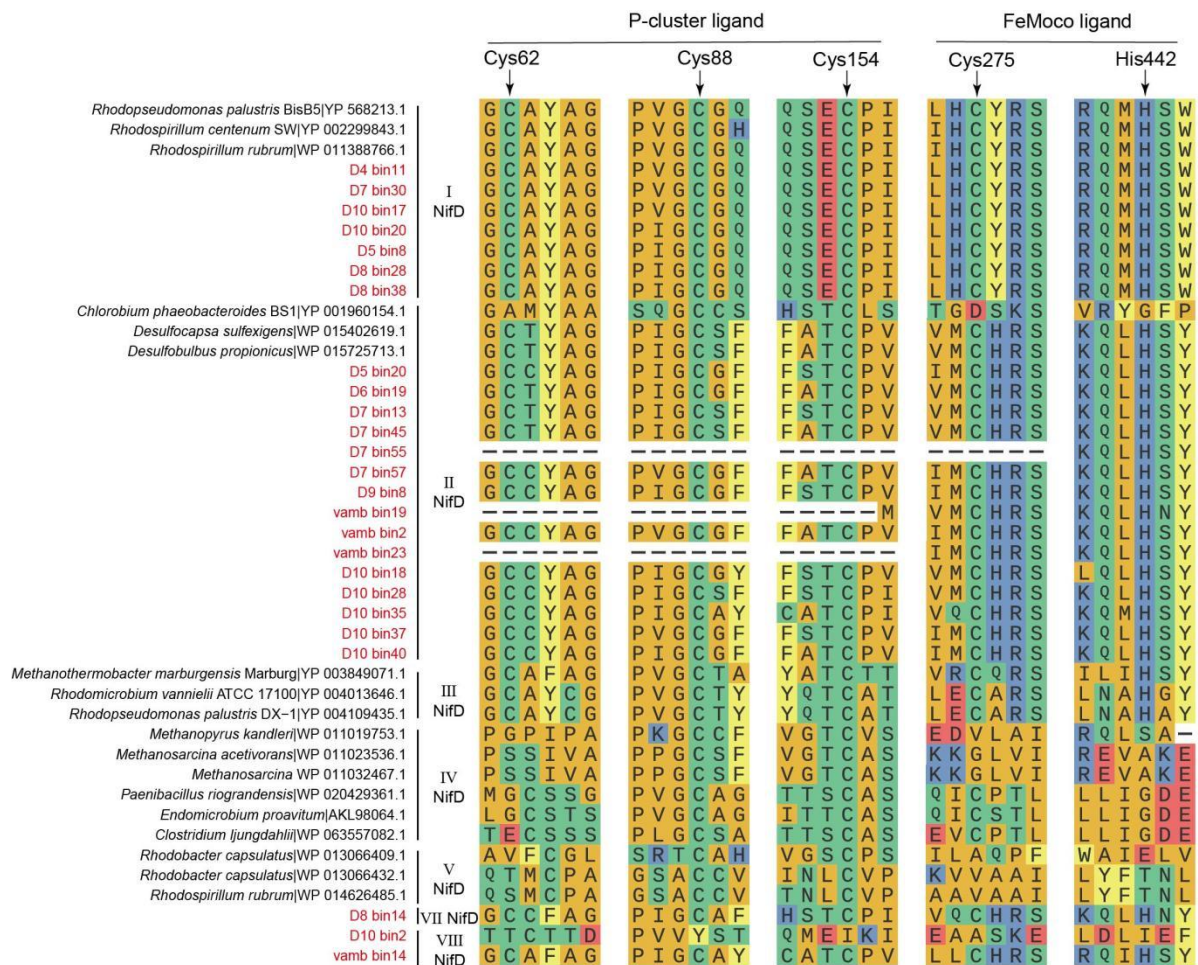


Fig. S8. NifD superfamily amino acid alignment. Multiple alignments of NifD superfamily sequences are shown in the region of active site residues responsible for coordination of the P-cluster and FeMo-cofactor within the molybdenum nitrogenase subunit NifD, and substrate bound to the FeMo-cofactor.

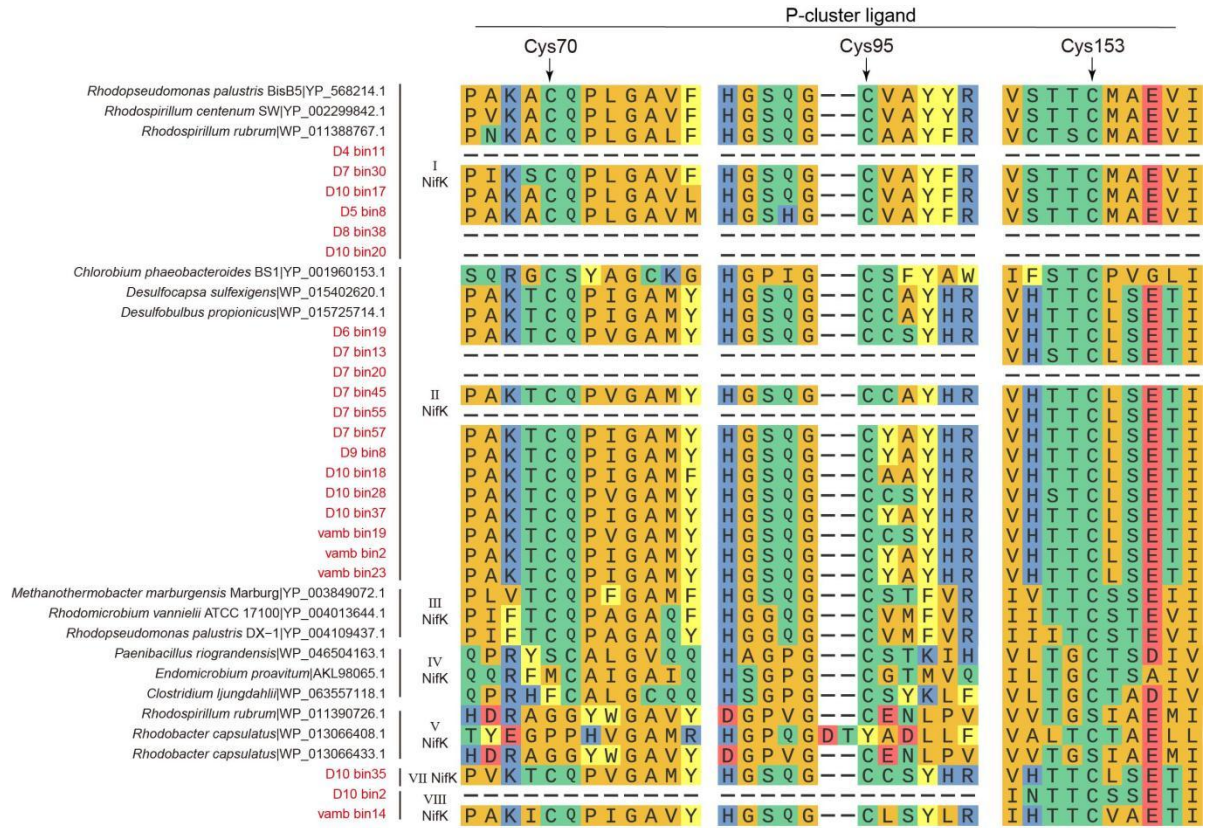


Fig. S9. NifK superfamily amino acid alignment. Multiple alignments of NifK superfamily sequences are shown in the region of active site residues responsible for coordination of the P-cluster within the molybdenum nitrogenase subunit NifK.

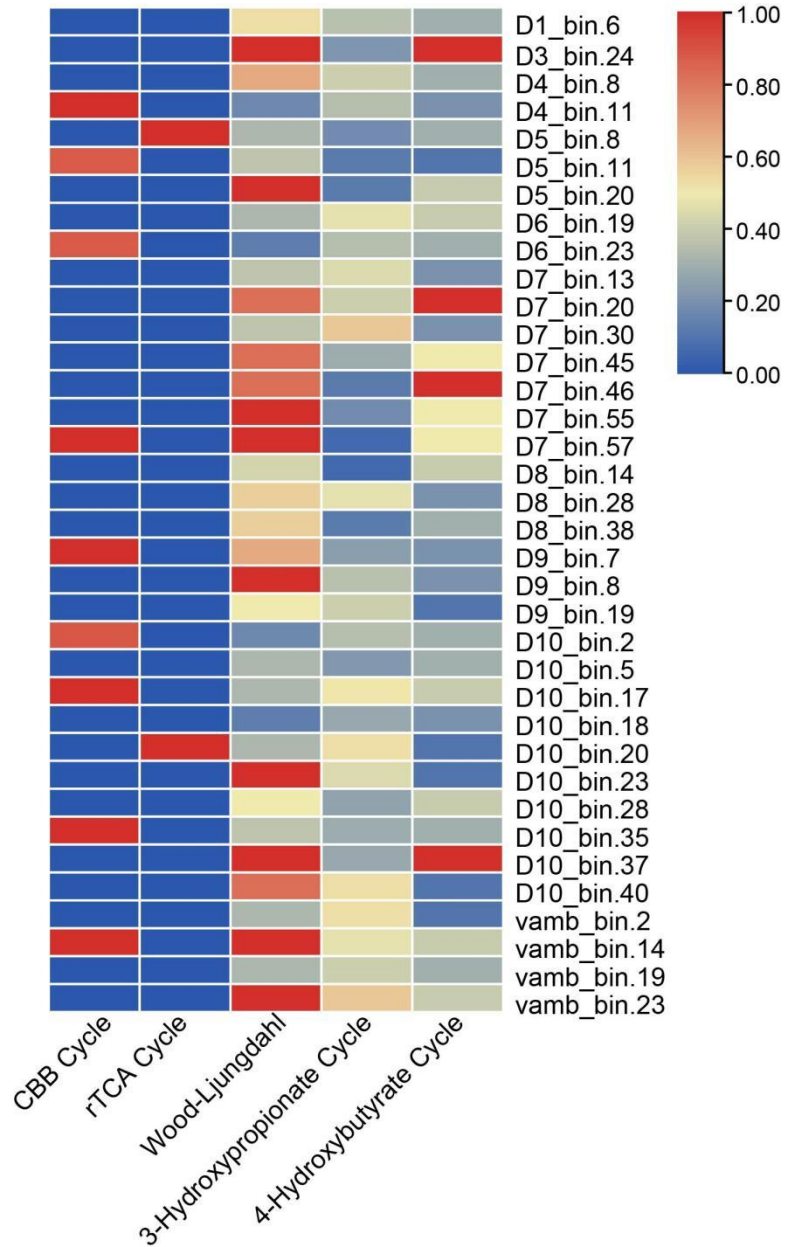


Fig. S10. The completeness of carbon fixation pathways by 36 putative nitrogen-fixating MAGs calculated by KEGG-Decoder.

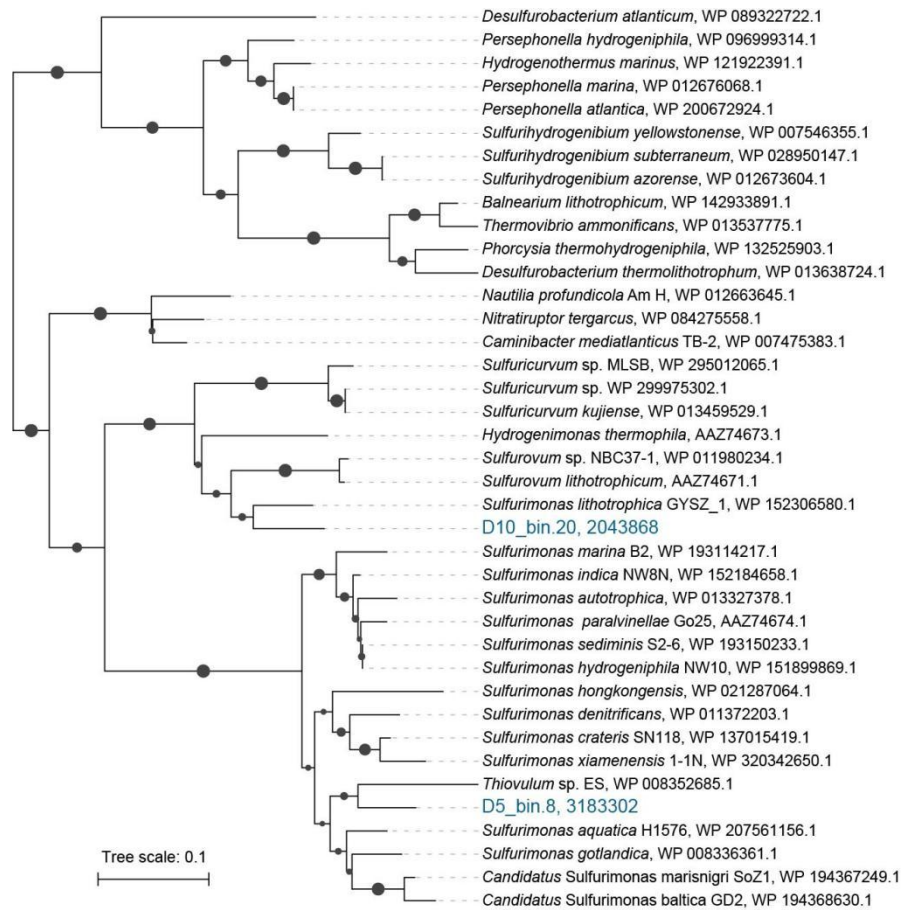
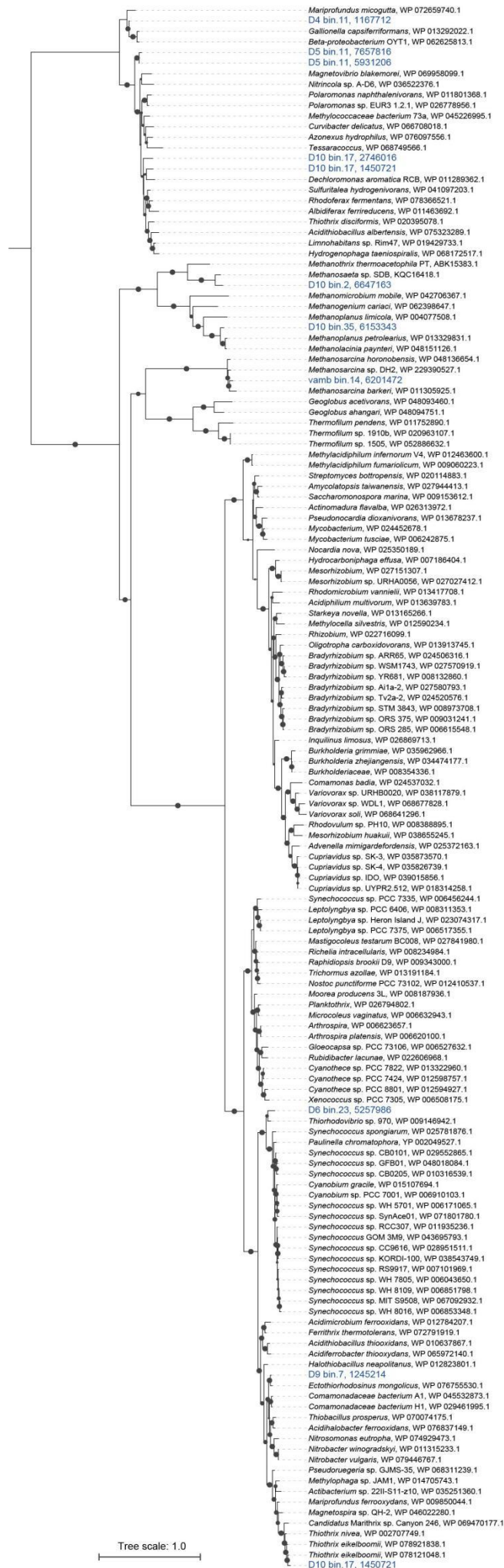


Fig. S11. Maximum-likelihood tree of amino acid sequences of AclB, a marker for carbon fixation. The tree shows sequences from mangrove sediments metagenome-assembled genomes obtained in this study (blue) alongside representative reference sequences (black). The subgroup of each reference sequence is denoted. Only bootstrap values $\geq 50\%$ (based on 1000 bootstrap replicates) are shown at nodes as closed circles.



Tree scale: 1.0

Fig. S12. Maximum-likelihood tree of amino acid sequences of RbcL, a marker for carbon fixation. The tree shows sequences from mangrove sediments metagenome-assembled genomes obtained in this study (blue) alongside representative reference sequences (black). Only bootstrap values $\geq 50\%$ (based on 1000 bootstrap replicates) are shown at nodes as closed circles.

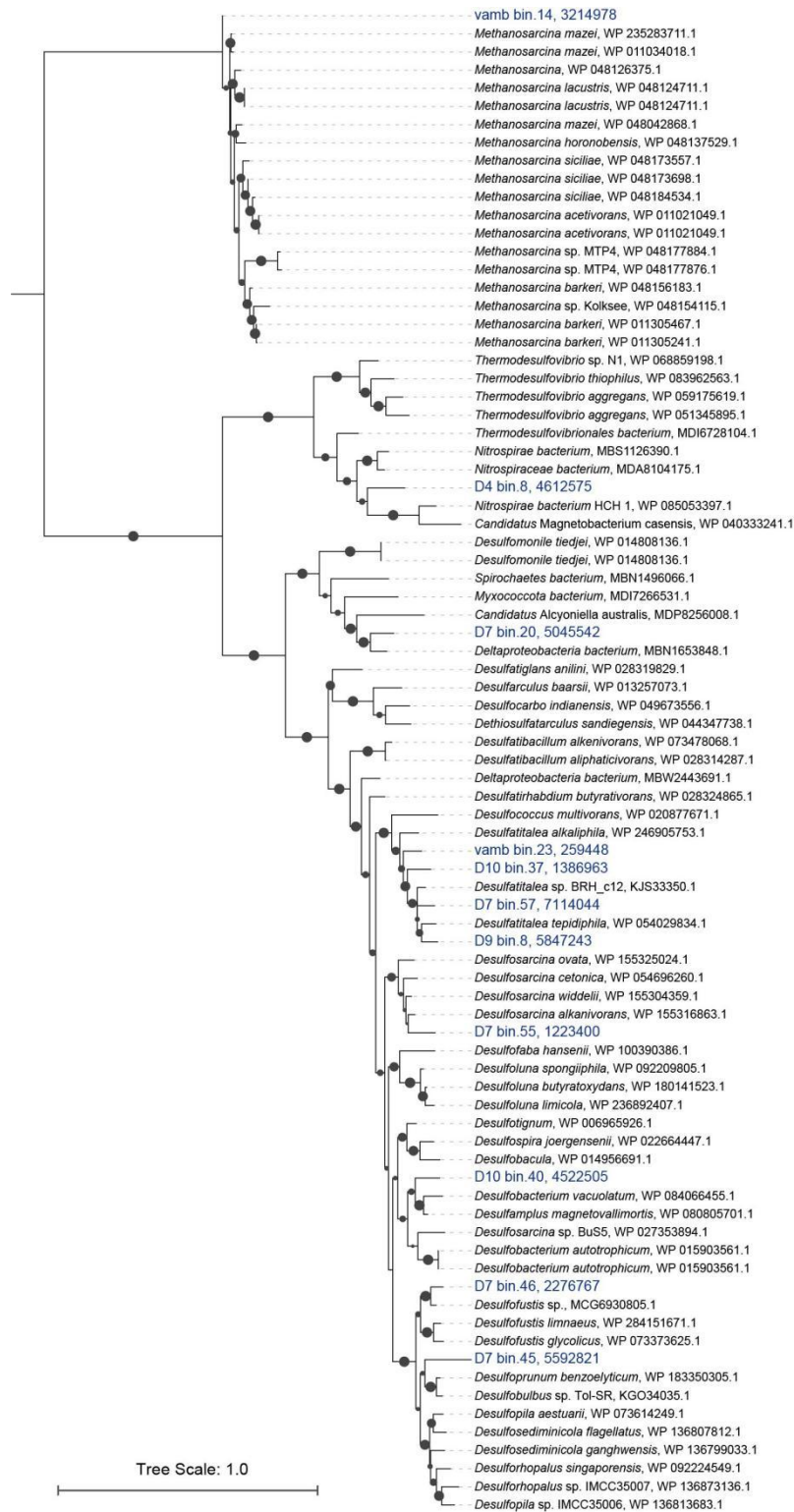


Fig. S13. Maximum-likelihood tree of amino acid sequences of AcsB, a marker for carbon fixation. The tree shows sequences from mangrove sediments metagenome-assembled genomes obtained in this study (blue) alongside representative reference sequences (black). Only bootstrap values $\geq 50\%$ (based on 1000 bootstrap replicates) are shown at nodes as closed circles.

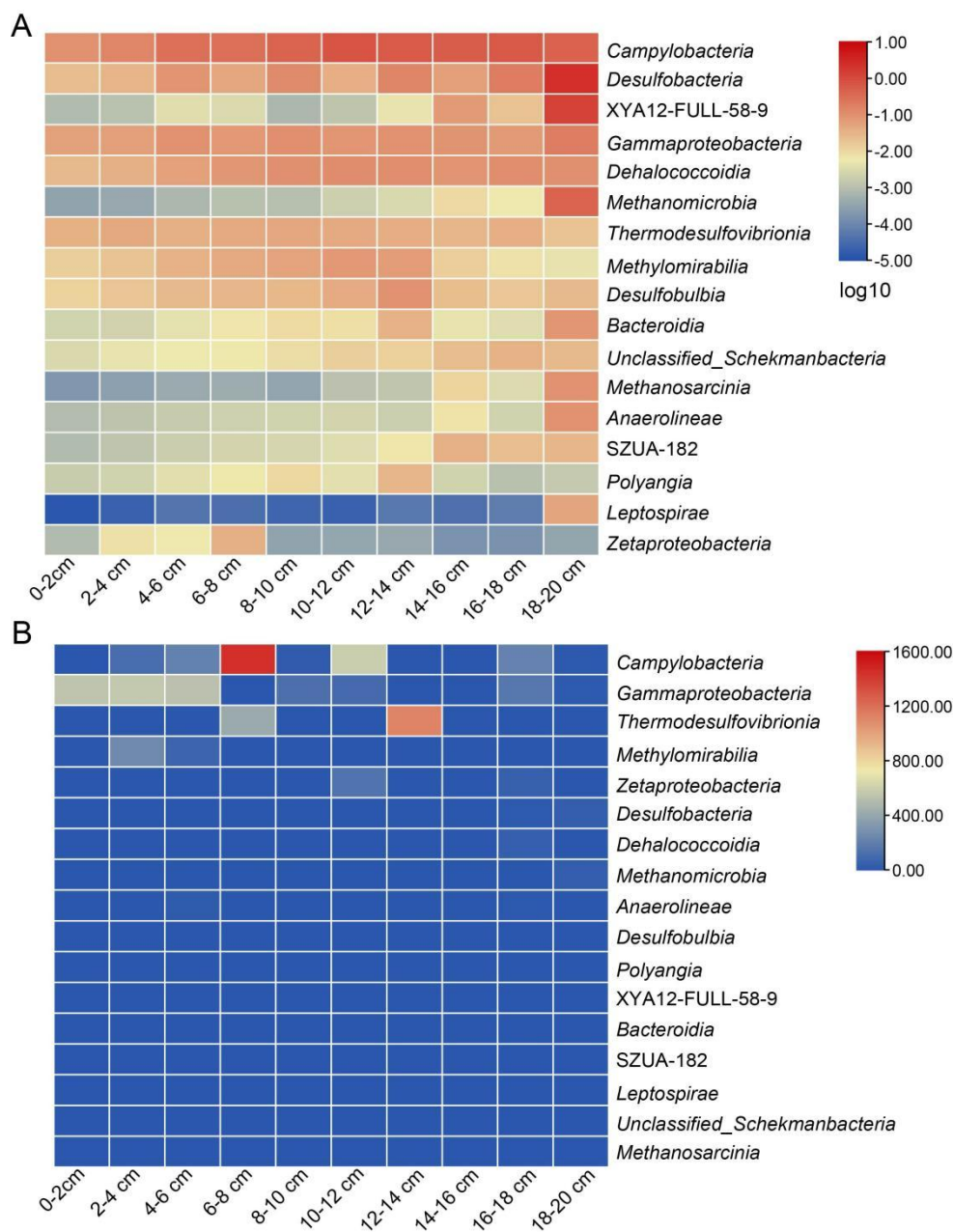


Fig. S14. Relative abundance (A) and transcript expression (B) of *nifH*-carrying MAGs at the class level in different mangrove sediment samples. Different taxonomic groups are represented by different colors.

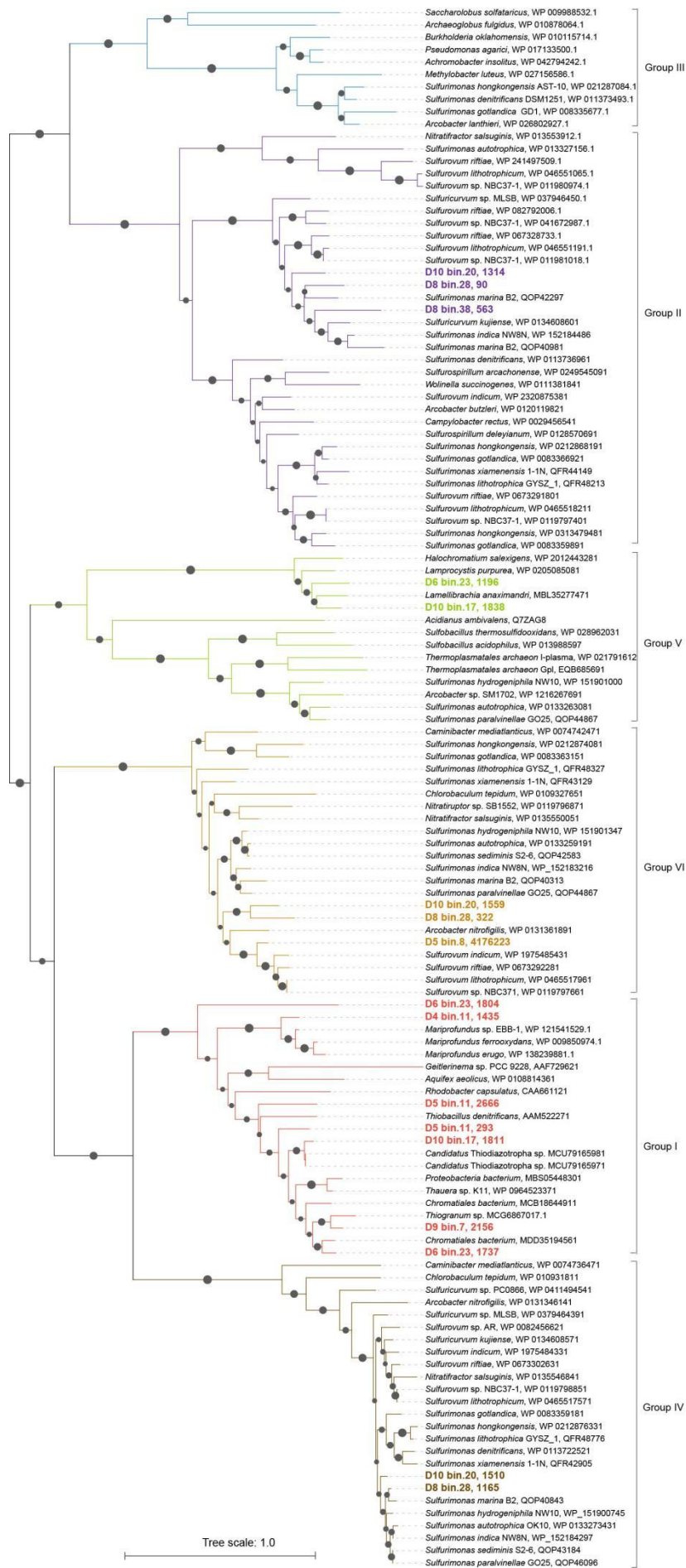


Fig. S15. Maximum-likelihood tree of amino acid sequences of Sqr, a marker for sulfide oxidation. The tree shows sequences from mangrove sediments metagenome-assembled genomes obtained in this study (different colors for group I-VI) alongside representative reference sequences (black). Only bootstrap values \geq 50% (based on 1000 bootstrap replicates) are shown at nodes as closed circles.

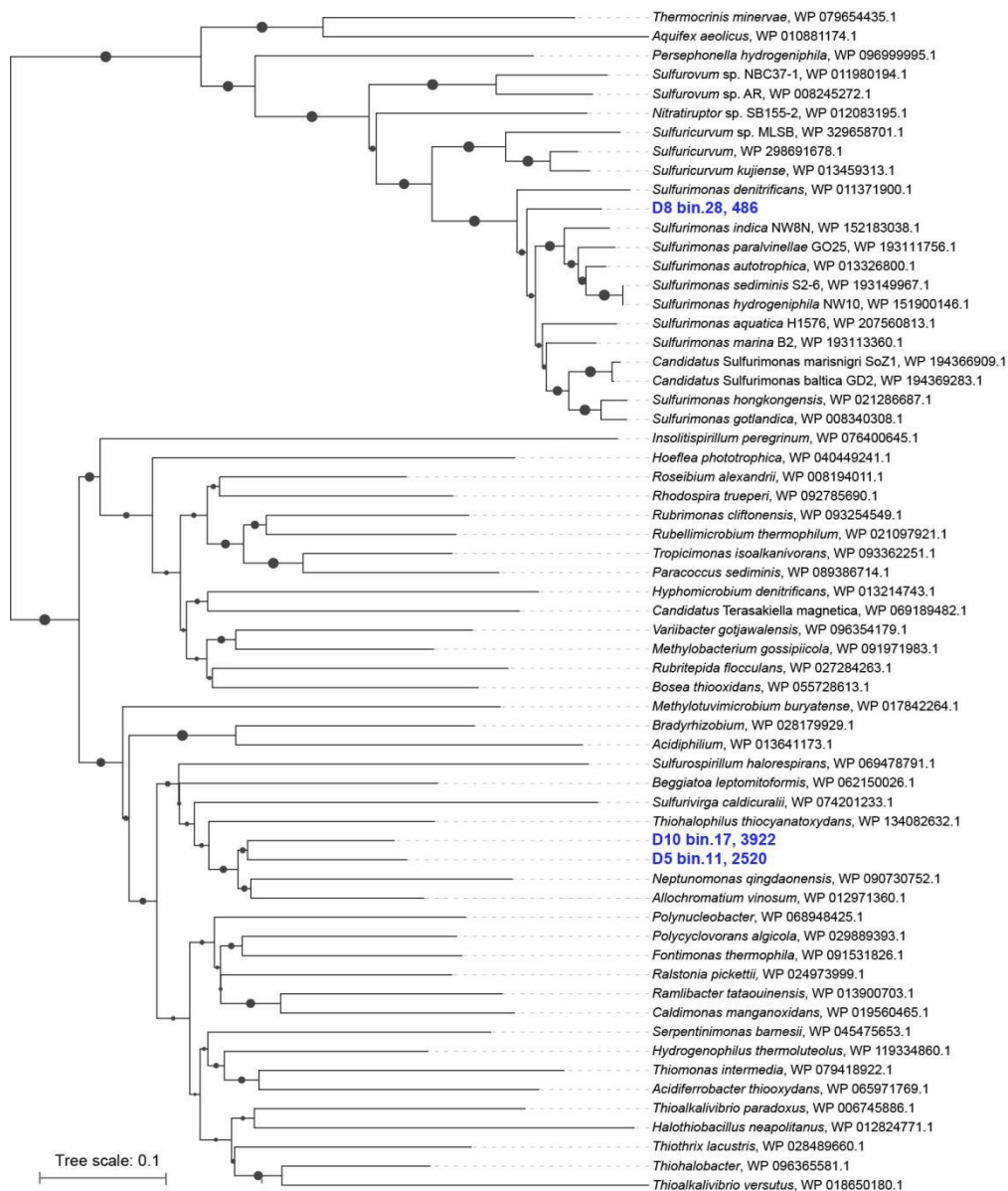


Fig. S16. Maximum-likelihood tree of amino acid sequences of SoxB, a marker for thiosulfate oxidation. The tree shows sequences from mangrove sediments metagenome-assembled genomes obtained in this study (blue) alongside representative reference sequences (black). Only bootstrap values $\geq 50\%$ (based on 1000 bootstrap replicates) are shown at nodes as closed circles.

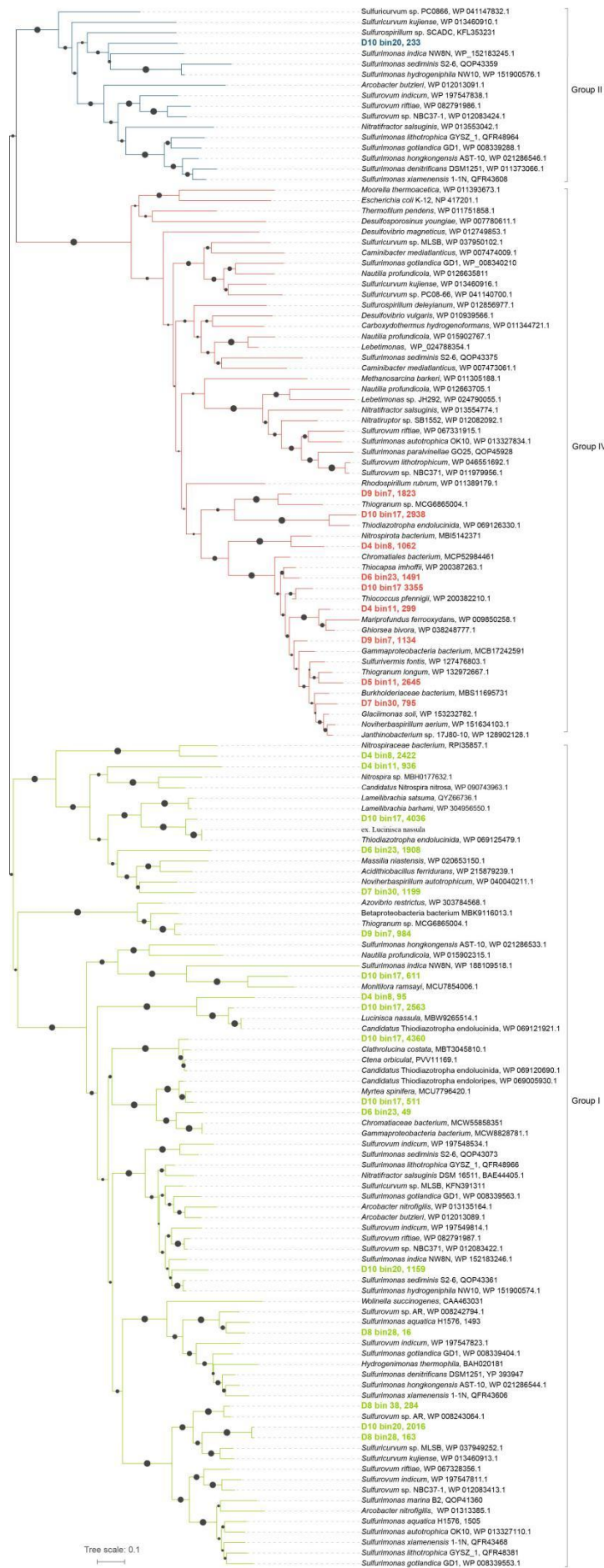


Fig. S17. Maximum-likelihood tree of amino acid sequences of [NiFe]-hydrogenase large subunit, which are markers for hydrogen oxidation.

The tree shows sequences from mangrove sediments metagenome-assembled genomes obtained in this study (different colors for group I, II, and IV) alongside representative reference sequences (black). Only bootstrap values $\geq 50\%$ (based on 1000 bootstrap replicates) are shown at nodes as closed circles.

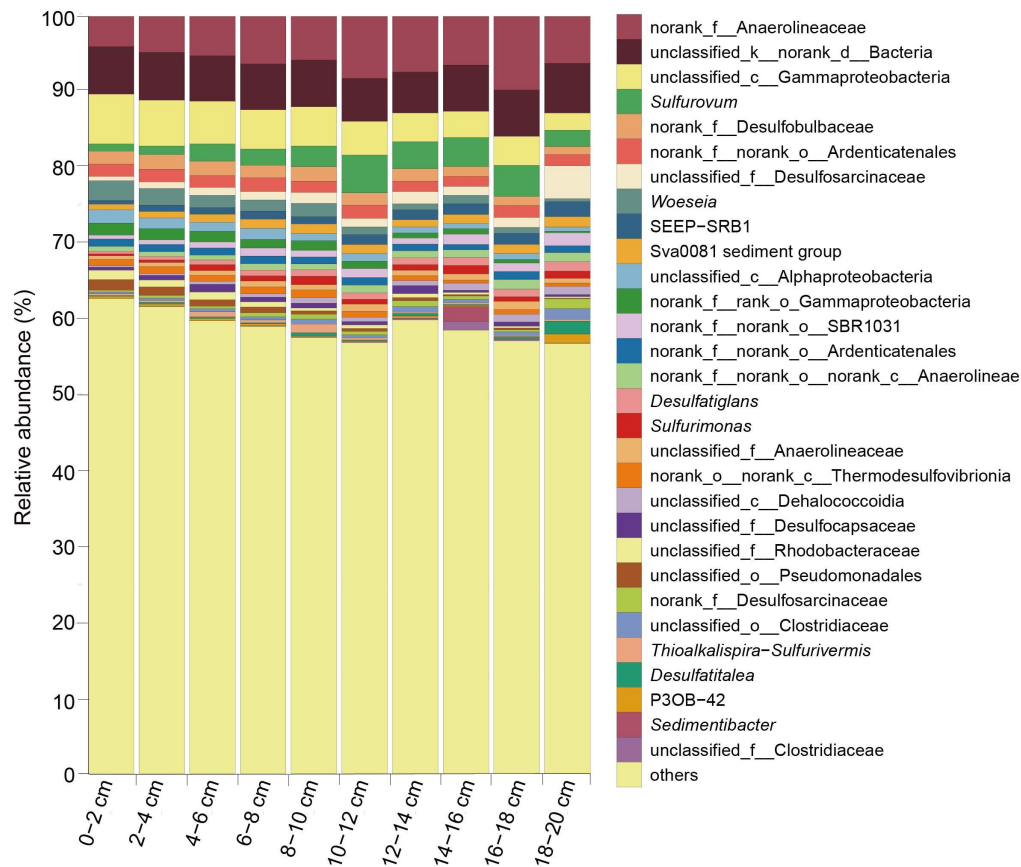


Fig. S18. Taxonomic compositions of microbial communities at the genus level based on 16S rRNA gene fragments extracted from metagenomes using the phyloFlash pipeline. For each plot, all taxa present in a relative abundance of $\geq 1.0\%$ in at least one of the samples are shown, whereas those present in $< 1.0\%$ relative abundance are grouped in the “others” category.

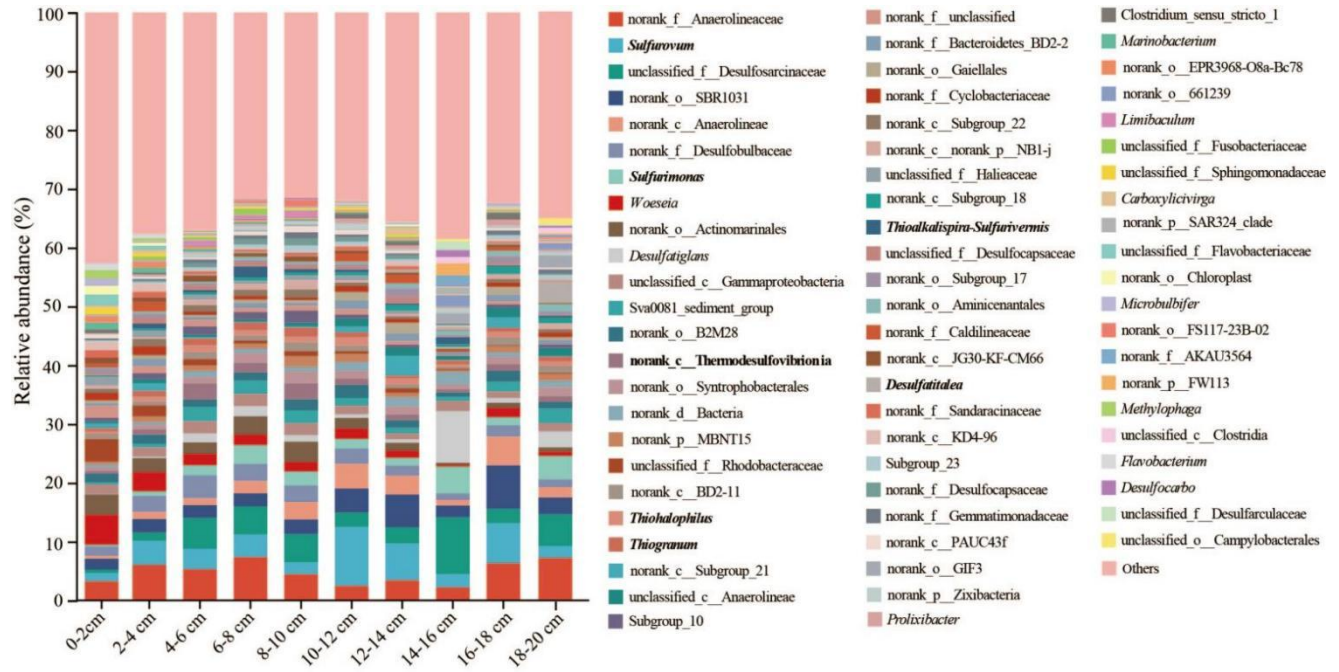


Fig. S19. Taxonomic distribution produced by 16S rRNA gene amplicon sequencing among the ten mangrove sediments at the genus level. For each plot, all taxa present in a relative abundance of $\geq 1.0\%$ in at least one of the samples are shown, whereas those present in $< 1.0\%$ relative abundance are grouped with unassigned sequences in the “others” category. The bold fonts represent the chemolithoautotrophic taxa.

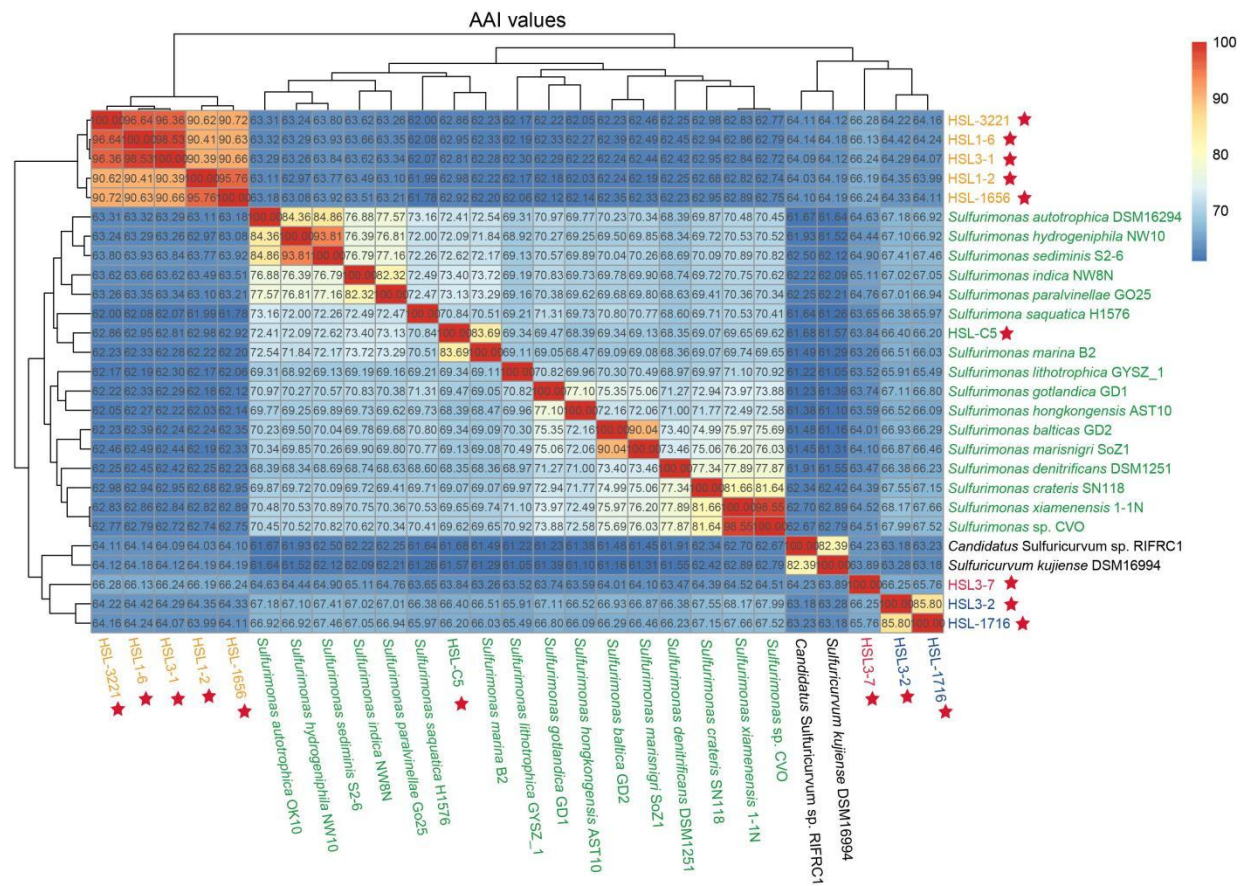


Fig. S20. AAI values between genomes of ten isolates and genomes of type strains of *Sulfurimonas* in the phylum *Campylobacterota*. The asterisk represents the strains isolated in this study.

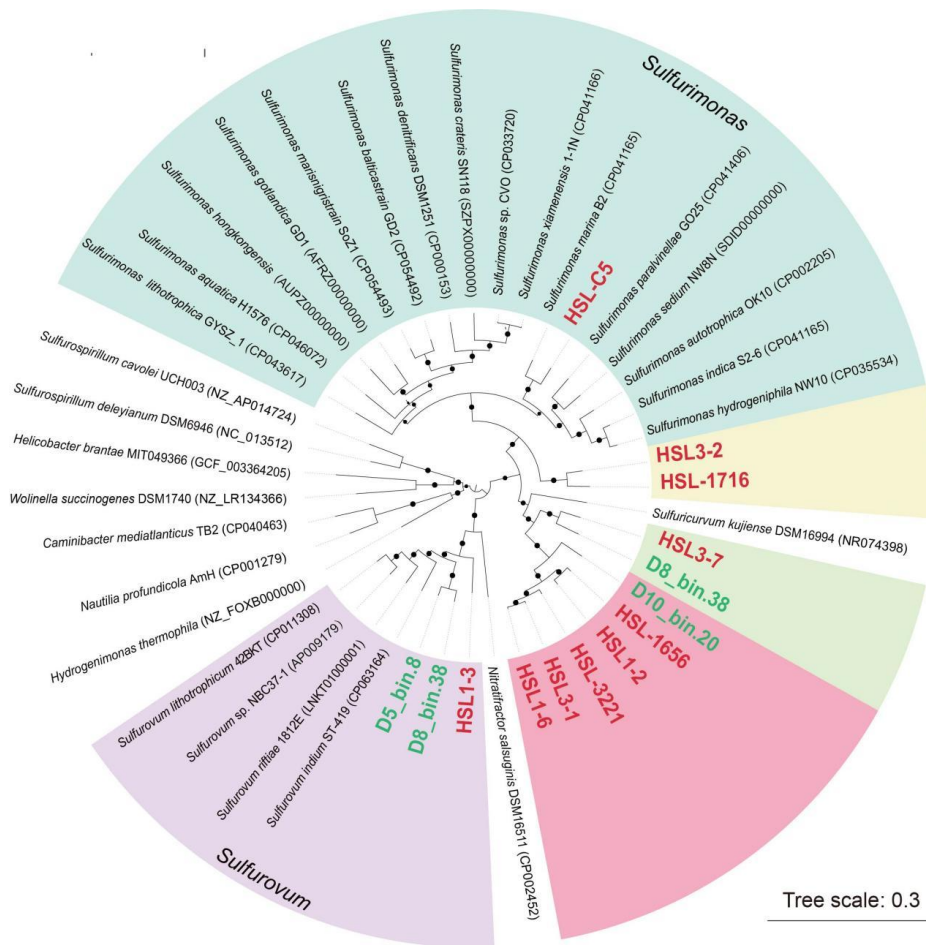


Fig. S21. The maximum-likelihood phylogenetic tree of ten pure cultures from *Campylobacterota* and their closely related taxa, based on the whole-genome sequence data. The red font indicates the isolates in this study, and the green font indicates the *Campylobacterota* MAGs from this study. The background colors of red, yellow, and light green represent three putative new genera. The bar represents 0.3 substitutions per position.

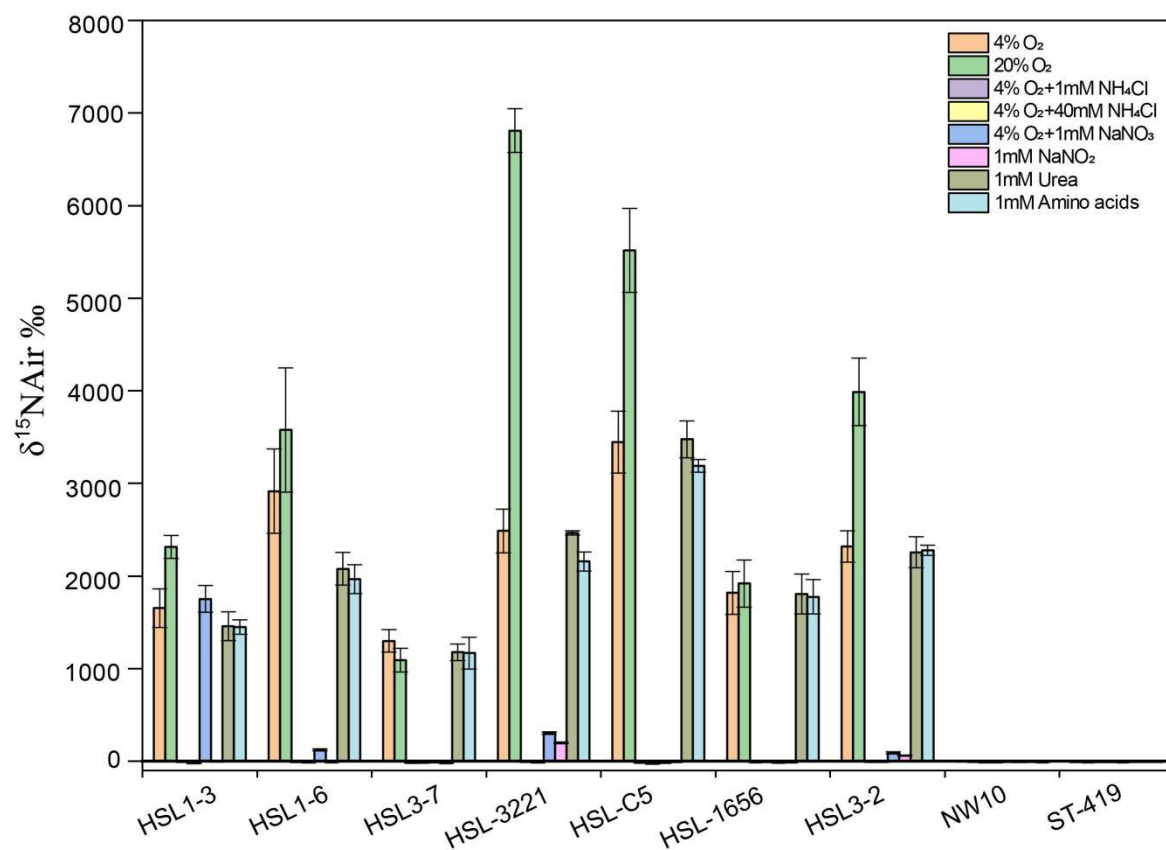


Fig. S22. Nitrogen fixation activities from representative strains in the presence of different oxygen concentration and different inorganic or organic nitrogenous compounds. Standard deviations are indicated by error bars.