

**Supplementary Information: Genomic potential for mercury biotransformation
in marine sediments across marginal slope to hadal zone**

Author List:

Zhuobo Li¹, Taoshu Wei², Lisheng He², Haifeng Qian³, Yong-Guan Zhu⁴, Yong Wang^{1, 5,*}

Affiliations:

¹ Institute for Ocean Engineering, Shenzhen International Graduate School, Tsinghua University, Shenzhen, China

² Institute of Deep Sea Science and Engineering, Chinese Academy of Sciences, Sanya, China

³ College of Environment, Zhejiang University of Technology, Hangzhou, China

⁴ Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China

⁵ Shenzhen Key Laboratory of Advanced Technology for Marine Ecology, Shenzhen International Graduate School, Tsinghua University, Shenzhen, China

***Corresponding author:**

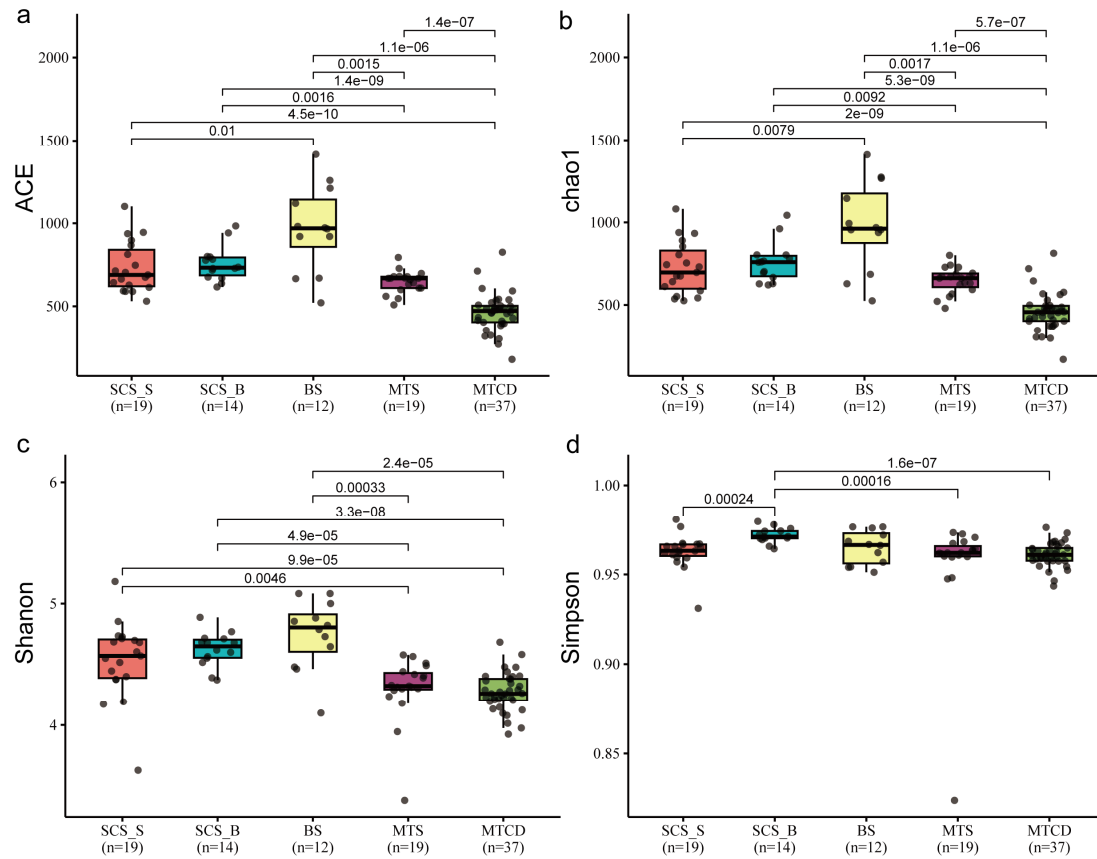
Yong Wang, PhD

Institute for Ocean Engineering, Shenzhen International Graduate School, Tsinghua University, P.R. of China

Phone: (86)-755-86709751

E-mail: wangyong@sz.tsinghua.edu.cn

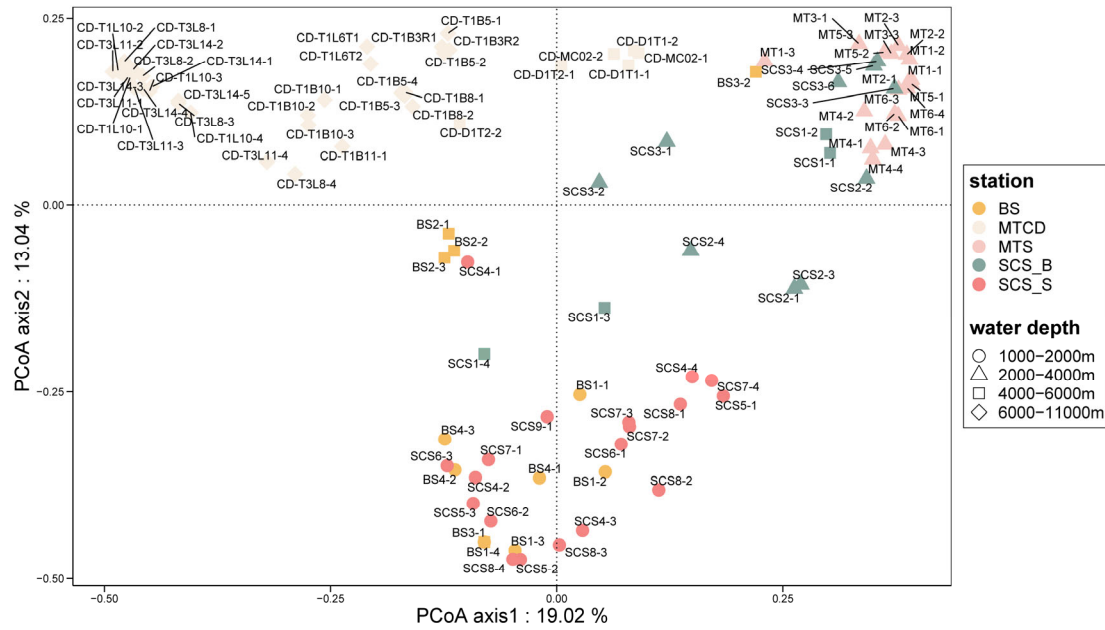
Supplementary Figures



Supplementary Figure 1: Alpha diversity indices of microbial communities.

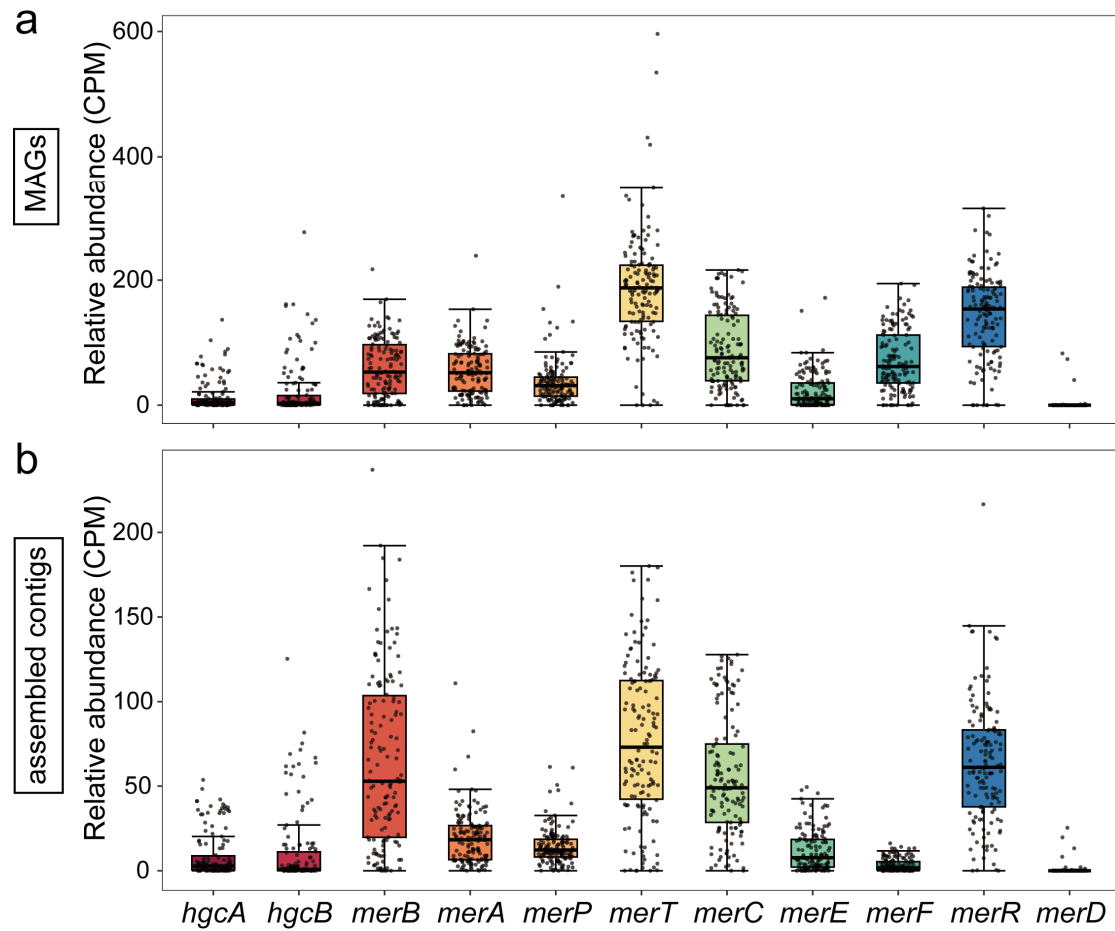
(a-d) ACE, Chao1, Shannon, and inverse Simpson indices of prokaryotic communities. *p* values were estimated using the two-sided Wilcoxon test for pairwise comparisons. In box plot, center lines indicate median values. The lower and upper bounds represent 25th and 75th percentiles, respectively. The lower/upper whiskers represent minima/maxima no further than 1.5 times interquartile range from the hinge.

SCS_S: South China Sea slope; SCS_B: South China Sea basin; BS: Bashi Channel; MTS: Mariana Trench slope; MTCD: Challenger Deep of Mariana Trench. Source data are provided as a Source Data file.



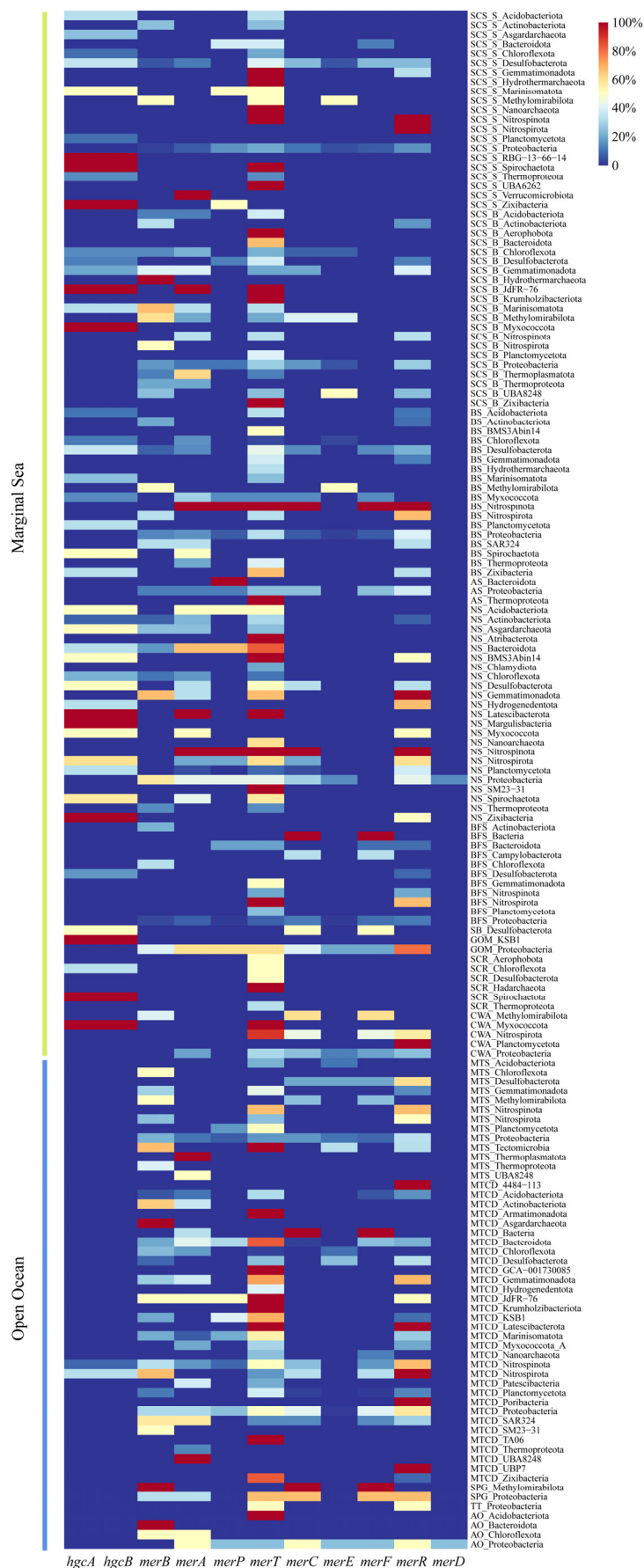
Supplementary Figure 2: PCoA analysis of prokaryotic community structures.

Principal coordinate analysis (PCoA) using Bray-Curtis dissimilarity matrix of the prokaryotic community structures in different stations. SCS_S: South China Sea slope; SCS_B: South China Sea basin; BS: Bashi Channel; MTS: Mariana Trench slope; MTCD: Challenger Deep of Mariana Trench. Source data are provided as a Source Data file.



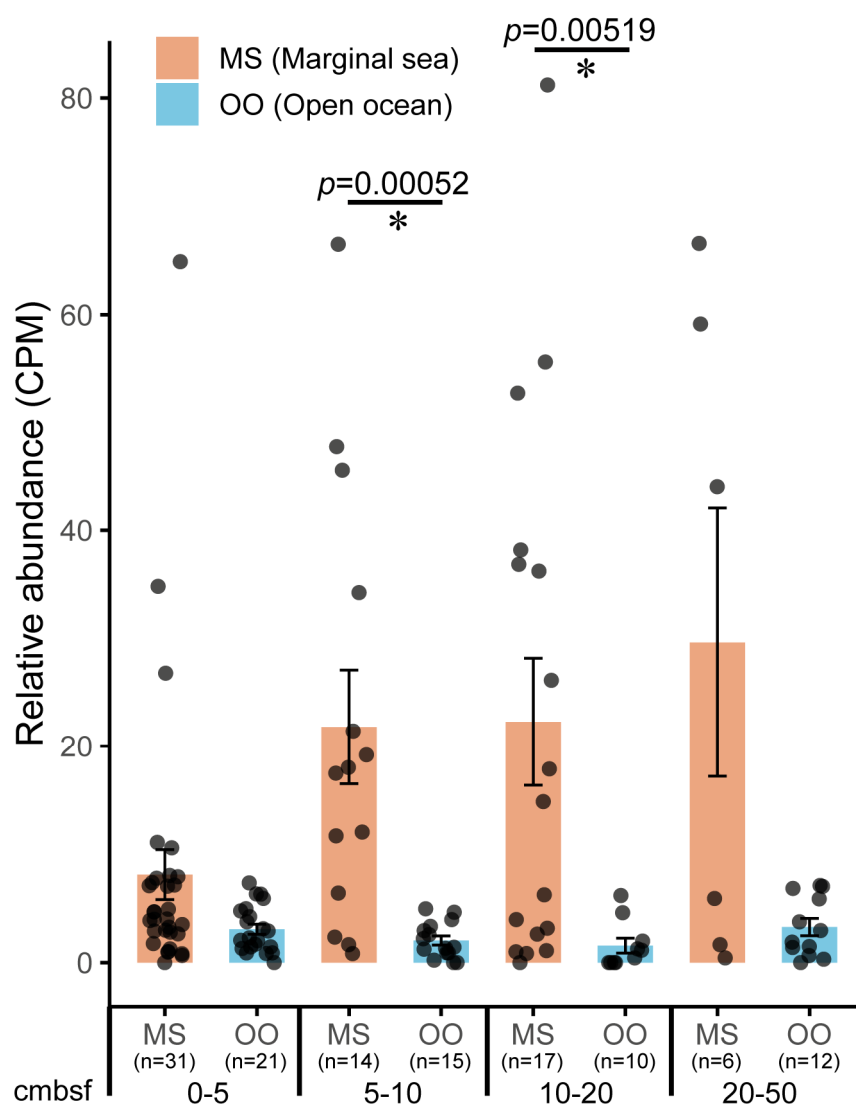
Supplementary Figure 3: Relative abundance of mercury transforming genes in 142 metagenomes.

(a, b) The relative abundance of the genes extracted from MAGs (a) and contigs (b) is denoted by CPM values calculated in individual metagenomes (n=142), respectively. In box plot, center lines indicate median values. The lower and upper bounds represent 25th and 75th percentiles, respectively. The lower/upper whiskers represent minima/maxima no further than 1.5 times interquartile range from the hinge, and the points falling outside of the whiskers represent the outliers. Source data are provided as a Source Data file.



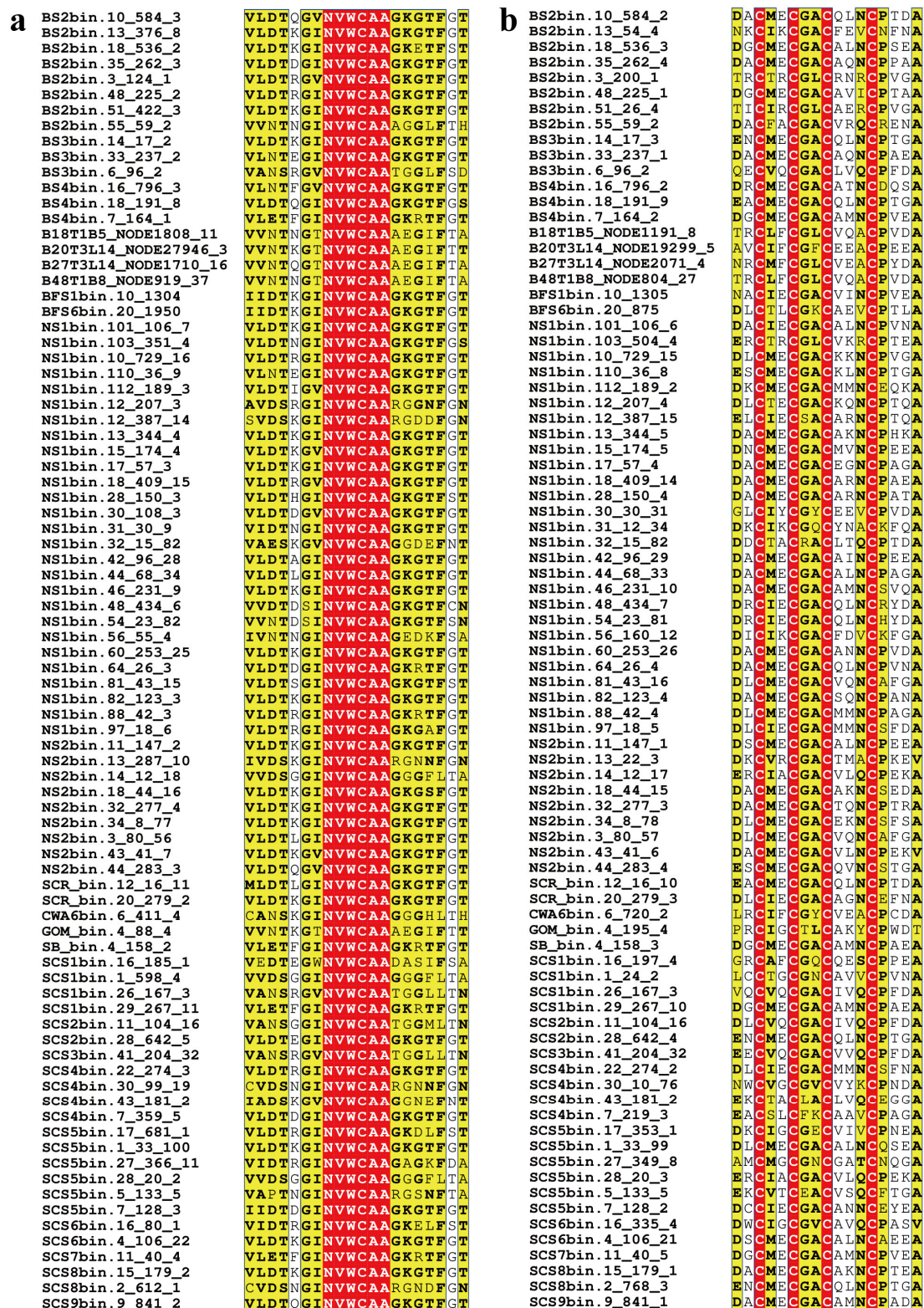
Supplementary Figure 4: Distribution of mercury transforming genes in MAGs.

The y-axis shows the percentage of MAGs from different stations that include mercury transforming genes. Color indicates the percentage of MAGs containing the genes. The phylum in the ID is presented with the sampling station information at the front. SCS_S: South China Sea slope; SCS_B: South China Sea basin; BS: Bashi Channel; AS: Arabian sea; NS: Norwegian Sea; BFS: Beaufort Sea; SB: Stellwagen Bank; GOM: Gulf of Mexico; SCR: Shore of Costa Rica; CWA: Coastal Western Antarctica; MTS: Mariana Trench slope; MTCD: Challenger Deep of Mariana Trench; SPG: South Pacific gyre U1369; TT: Tonga Trench; AO: Atlantic Ocean. Source data are provided as a Source Data file.

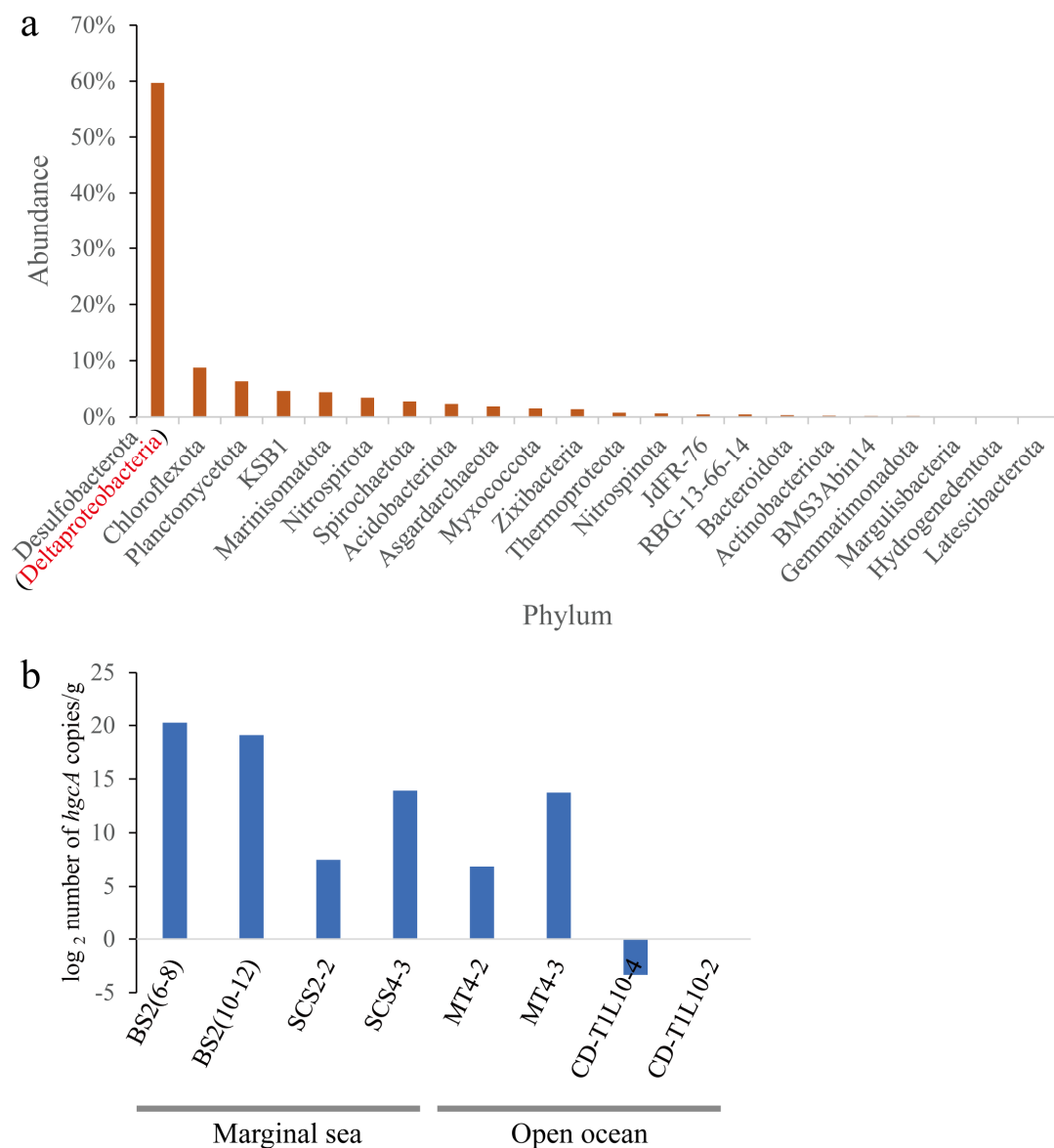


Supplementary Figure 5: Mean relative abundance of the *hgcA* gene in the metagenomes of marginal sea (MS) and open ocean (OO) stations.

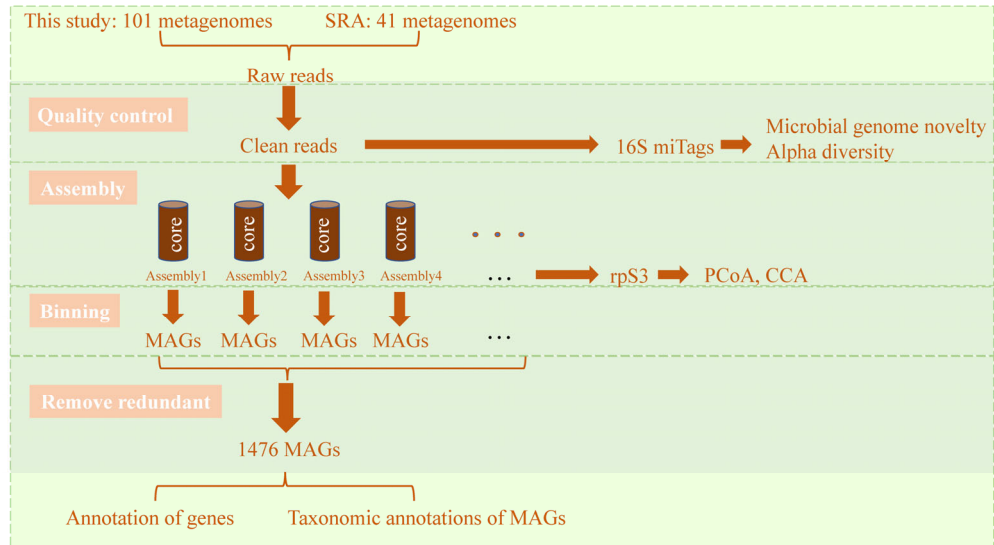
The x-axis represents the sampling depth of the samples. Error bars represent standard error and asterisks indicate significant difference between the two groups (two-sided Wilcoxon test, $p < 0.05$). Source data are provided as a Source Data file.



Supplementary Figure 6: Alignments of identified HgcAB protein sequences from dereplicated MAGs recovered from all sediments. (a) Alignments of identified HgcA protein sequences. (b) Alignments of identified HgcB protein sequences.



Supplementary Figure 7: Taxonomic distribution and quantitative abundance of the *hgcA* gene in sediment samples. (a) Phylum-level percentage distribution of *hgcA* sequences relative to total *hgcA* abundance (quantified as CPM) across 142 metagenomic datasets. (b) Abundance of the *hgcA* gene in eight sediment samples was quantified by qPCR. Source data are provided as a Source Data file.



Supplementary Figure 8: Workflow diagram of metagenomic analysis.