

Supplementary information

Salinization alters microbial methane cycling in freshwater sediments

Authors: Lorena Selak, Dimitri V. Meier, Maja Marinović, Andrea Čačković, Katarina Kajan, Petra Pjevac, Sandi Orlić

Supplementary protocol I

q-PCR protocol for quantification of functional and phylogenetic marker genes

Preparing standards

The qPCR standards were generated from genes amplified from environmental samples using the PCR method (Table 1) using a 96 – Well Thermal Cycler (Applied Biosystems, Foster City, CA, USA). The resulting PCR products were purified using the OneStep PCR Inhibitor Removal kit (Zymo Research, CA, USA), and purified amplicons were ligated into PCR4-TOPO plasmids (Invitrogen, MA, USA) and transformed into XL1-Blue Competent cells (Agilent, Santa Clara, CA, USA) following the manufacturer's instructions. Positive clones were identified through colony PCR with M13 primer pairs, and plasmid isolation was performed using the QIAprep Spin Miniprep kit (Qiagen, Valencia, CA, USA). The isolated plasmid DNA was then subjected to Sanger sequencing to ensure the accuracy and success of the cloning process. The concentrations of the plasmid DNA containing the successfully cloned sequences of interest were measured using a Qubit fluorometer (Thermo Fisher Scientific, MA, USA). The number of copies was adjusted within the range of 10^7 - 10^1 by diluting the initial samples.

Table 1. PCR conditions and details on primers used in reactions.

Gene	<i>pmoA</i>	<i>mcrA</i>	<i>dsrAB</i>	515-806 (V4, 16S)
Sequence length (bp)	491	469	221	291
Primers	A189F Mmb661r	mlas mcrA-rev	DSR-F+ DSR-R	515F 'Parada' 806F 'Appri1'
Sequence	GGNGACTGGGACTTCTGG CCCGMGCAACGTCYTTAC	GGTGGGTGTMGGDTTCACMCARTA CGTTCATBGCCTAGTTVGGRTAGT	ACSCACTGGAAGCACGGCGG GTGGMRCCGTGCAKRTTGG	GTG YCA GCM GCC GCG GTA A GGA CTA CNV GGG TWT CTA AT
Literature	Kolb et al., 2003	Steinberg & Regan, 2008	Leloup et al., 2007	Parada et al. 2015 / Appri et al., 2015
Initial denaturation	94 °C 30 s	94 °C 30 s	94 °C 30 s	98 °C 30 s
Denaturation	94 °C 30 s	94 °C 30 s	94 °C 30 s	98 °C 10 s
Annealing	56 °C 50 s	58 °C 45 s	56 °C 50 s	62 °C 30 s
Extension	68 °C 40 s	68 °C 30 s	68 °C 25 s	72 °C 30 s
Final extension	68 °C 5 min	68 °C 5 min	68 °C 5 min	72 °C 5 min
Number of cycles	30	30	33	29

Quantitative PCR assays

Quantitative PCR assays were carried out using the Bio-Rad CFX96 Touch system (Bio – Rad, CA, USA). The reactions consisted of 10 µL 1x SsoAdvanced Universal SYBR Green Supermix

(BioRad, CA, USA), 1 μ L of each primer (10 μ mol L⁻¹), 1 μ L of DNA, and 8 μ L of ultrapure DNase-free water. The extracted DNA was diluted 10 to 50 times to ensure results within the valid range and minimize the potential impact of reaction inhibitors. qPCR was performed in triplicates with positive controls and negative controls (DNase-free water). Details of the primers and qPCR conditions are provided in Table 2. The gene copy number per gram of sediment was calculated using the mean Cq value, normalized by the DNA concentration of the sample, and adjusted based on a calibration curve generated from standards with known copy numbers.

Table 2. qPCR conditions and details on primers used in reactions.

Gene	<i>pmoA</i>	<i>mcrA</i>	<i>dsrAB</i>	515-806 (V4, 16S rRNA)
Sequence length (bp)	491	469	221	291
Primers	A189F Mmb661r	mlas mcrA-rev	DSR-F+ DSR-R	515F 'Parada' 806F 'April'
Sequence	GGNGACTGGGACTTCTGG CCGGMGCAACGTCYTTACC	GGTGGTGTMGGDTTCACMCARTA CGTTCATBGCAGTTVGGRTAGT	ACSCACTGGAAGCACGGCGG GTGGMRCCTGCAKRTTGG	GTG YCA GCM GCC GCG GTA A GGA CTA CNV GGG TWT CTA AT
Literature	Kolb et al., 2003	Steinberg & Regan, 2008	Leloup et al., 2007	Parada et al., 2015 / Appril et al., 2015
Initial denaturation	95 °C 3 min	95 °C 3 min	94 °C 10 min	95 °C 3 min
Denaturation	95 °C 10 s	95 °C 10 s	94 °C 15 s	95 °C 30 s
Annealing & Extension	62 °C 1 min	57 °C / 72 °C 20 s / 30 s	59 °C 1 min	50 °C 1 min
Fluorescence data acquisition	87 °C 8 s	83 °C 10 s	72 °C 45 s	72 °C 45 s
Number of cycles	40	40	41	40
Melting curve generation	60-95 °C 5 s	60-95 °C 5 s	45-95 °C 5 s	65-95 °C 5 s

Supplementary protocol II

Amplicon data processing

The FASTQ workflow (default parameters; BaseSpace; Illumina) was employed to extract amplicon pools from raw sequencing data. PhiX contamination was removed using BBduk (BBtools, [1]). Data was demultiplexed using the Python package demultiplex (Laros JFJ, github.com/jfjaros/demultiplex), allowing for 1, 2, and 2 mismatches to the barcode, linker and primer sequences, respectively. FASTQ reads were trimmed at 220 nt for the forward and 150 nt for the reverse read. Amplicon sequence variants (ASVs) were inferred using the DADA2 R package version 1.20.0 [2] with default settings. Taxonomic assignments were performed with the SINA version 1.6.1 classifier [3] against the SILVA reference database (SSU RefNR 99, v. 138.1).

The 16S rRNA gene amplicon data was preprocessed, removing ASVs classified as eukaryotes, mitochondria, or chloroplasts, along with unclassified and singleton ASVs. The phyloseq (v. 1.32.0, [4]), vegan (v. 2.5-6, [5]), microbiome (v. 1.10.0, [6]), and ggplot2 packages (v. 3.3.2, [7]) were used for analysis and visualization. Differences in qPCR data and functional group abundances were first assessed using one-way ANOVA, followed by Games-Howell post-hoc tests to evaluate pairwise group differences while accounting for unequal variances and sample sizes. Spatial variation in functional group distribution within surface sediments was assessed using a distance-decay analysis based on 16S rRNA gene amplicons. Multiple regression with a dissimilarity matrix derived from Bray-Curtis coefficients was employed focusing on observing changes in community structure among different geographic locations.

Supplementary protocol III

Metagenome bioinformatics analysis

Metagenomic DNA libraries were prepared using the NEBNext Ultra FS II DNA Library Prep kit for Illumina and NEBNext Multiplex Oligos for Illumina. Barcoded libraries were pooled and sequenced on the NovaSeq 6000 platform (2 x 150 bp). Metagenomic read sets were trimmed and quality filtered (Q30), and Illumina adapters and Phix sequences were removed using bbdduk (bbmap v. 39.01., [1]). Sequencing depth in each read set was normalized using bbnorm (bbmap v. 39.01., [1]), de novo assembled with MEGAHIT (min. 1000nt, v. 1.2.9, [8]), and binned into metagenome assembled genomes (MAGs) with MetaBAT2 [9] using specific settings for high - quality binning (percentage of 'good' contigs 90-97, maximum number of edges per node 200-500, minimum score of an edge for binning 60-70). DAS-Tool (v. 1.1.6, [10]) was used to identify and remove redundant MAGs, followed by de-replication using dRep (v. 3.4.2, [11]). Representative MAGs (>99% ANI) were assessed for completeness and contamination with checkM2 (v. 1.0.1, [12]), while taxonomy was assigned with GTDB-Tk (v. 2.1.1, [13]). Manual refinement of MAGs was conducted based on checkM2 results, single copy gene (SCG) abundance, and contig classification (Kaiju with maximum number of 5 mismatches, cutoff E-value 0.00001, v. 1.9.2, [14]) using the Anvio platform (v. 7.1, [15]). Taxonomy of the refined MAGs was again assigned with GTDB-tk (v. 2.1.1), and MAG quality was re-assessed with checkM2. MAGs with a completeness > 50% and contamination < 10% were used to calculate the relative abundance of MAGs across samples based on contig mean coverage. These thresholds were chosen to enable

broader representation of community members relevant to methane-cycling. While more stringent cutoffs are commonly used for metabolic reconstructions, our aim was to capture taxonomic and functional patterns across sites rather than detailed physiological predictions. Functional annotation of MAGs was performed using Interproscan (v. 5.62.-92.0, [16]) and the Anvio platform. MAGs were queried for genes involved in osmoregulation processes, including aquaporins, stress response proteins, and synthesis of osmoprotectants. Additionally, MAGs were searched for the presence of multi-heme cytochromes and pili. A manual search for sequences containing heme-binding motifs ($CX_{(n-2-5)}CH$) was performed using the 3of5 application [17]. Amino acid sequences containing more than 3 heme-binding motifs were used in protein localization within the cell using PSORTb (v. 3.0.3, [18]) and the results were compared to previously assigned functional annotations. Phylogenetic trees of hmm search detected functional genes of interest, namely *mcrA*, *dsrAB*, and *pmoA*, were constructed by clustering amino acid sequences of each functional gene separately using Clustal Omega (v.1.2.4, [19]). The resulting alignments were trimmed and refined using trimAl (v. 1.2., [20]). Fasttree, a tool from the GToTree software suite [21], was used to calculate phylogenetic trees based on the refined alignments. The trees were visualized and edited using iTOL (v.6, [22]). Furthermore, the 16S rRNA gene sequences obtained from the contigs were mapped against the SILVA database (v. 138.1) using bbmap. The mapped sequences were then clustered based on a 98% similarity threshold using phyloFLASH (v. 3.4, [23]).

Network analysis

For network analysis, we generated estimated microbial community abundances by multiplying Q2Q3 mean coverage values of MAGs with 16S rRNA gene qPCR data. All MAGs with completeness >50 and contamination <10 were utilized in CoNet (v.1.1.1, [24]) for network analysis. Pearson and Spearman correlations were calculated ($R^2_{\min} = 0.8$) along with distances based on Bray Curtis dissimilarity ($R^2_{\min} = 0.8$). The data was randomized with 1000 iterations and resampled using the bootstrap method. Significance was determined using the Simes method for *p*-value combination, and a Benjamini-Hochberg multiple test correction was applied with a threshold at $p < 0.05$. In addition, the MAG co-occurrence was estimated using FastSpar with default parameters (v.1.0.0, [25]) and the two results were overlapped. The resulting network was visualized in Gephi (v0.10.1, [26]), where node size corresponded to the number of connections. Nodes were color-coded based on the functional group, and their ID-s were assigned based on the

lowest taxonomic rank assigned via GTDBk. The Louvain method was employed to assess the strength of the network clustering, and edges were distinguished by mutual exclusion or copresence and co-occurrence percentage [27]. Network clusters were identified using the Modularity algorithm in Gephi to detect communities based on modularity optimization and assigned nodes to distinct modularity classes. To determine the spatial and geochemical affiliation of each cluster, the presence or absence of nodes corresponding to MAGs was examined across sediment zones.

Supplementary results

Phosphate concentrations averaged 0.21 mmol L^{-1} in the salinization profile and 0.24 mmol L^{-1} in the freshwater profile, and similarly to nitrate concentrations, no clear trend with sediment depth was evident in either profile (Fig. 1Sa,b.). Nitrate concentrations increased from inlets to cove, and peaked at station 12, with a tenfold difference between sediment profiles at the surface (Fig. S1b.).

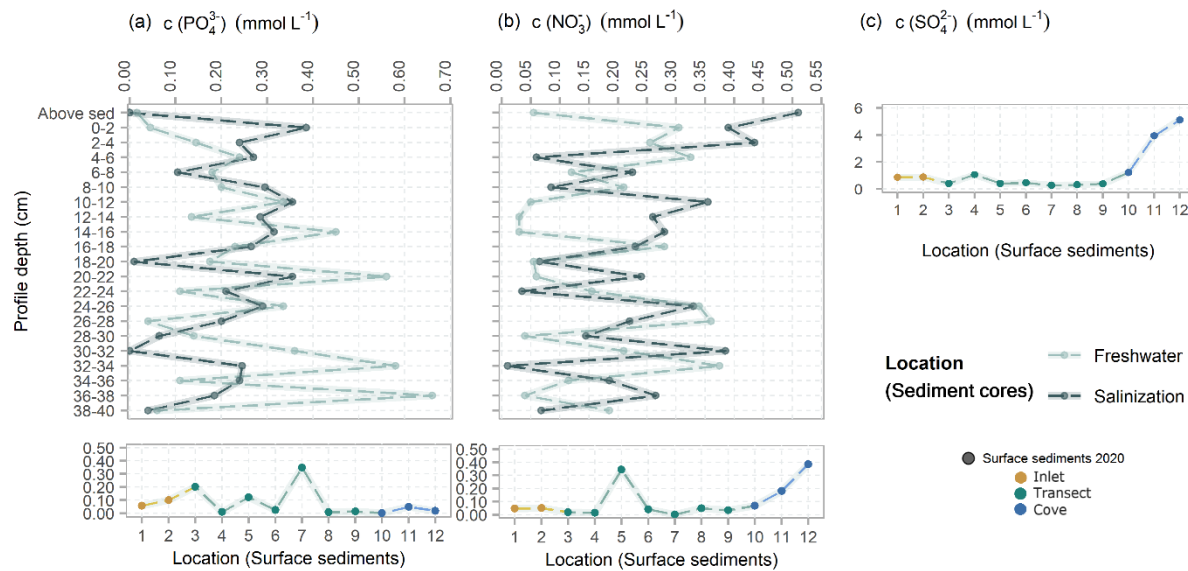


Figure S1 Physicochemical parameters of pore water retrieved from surface sediments namely inlets (1-2), transect (3-9), and cove (10-12), and deeper sediment profiles, namely Salinization profile, and Freshwater profile.

Functional gene relative abundances and gene copy numbers derived from qPCR were used to calculate Spearman correlation between the genes and physicochemical parameters in surface sediments (Table S1a.) and deeper sediment profiles (Table S1b.).

Table S1 a. Spearman correlation of functional gene relative abundances in surface sediments.

		<i>dsrAB</i>	<i>pmoA</i>
<i>pmoA</i>	ρ	-0.13	/
	<i>p</i>	>0.05	/
<i>mcrA</i>	ρ	0.05	0.64
	<i>p</i>	>0.05	<0.05

Table S1 b. Spearman correlation of functional gene copy numbers and physicochemical parameters of sediment profiles.

		<i>dsrAB</i>		<i>pmoA</i>		<i>mcrA</i>		Methane		Salinity		SO ₄ ²⁻	
		Jugovir	Lateral	Jugovir	Lateral	Jugovir	Lateral	Jugovir	Lateral	Jugovir	Lateral	Jugovir	Lateral
<i>pmoA</i>	ρ	-0.67	-0.06	/	/								
	<i>p</i>	<0.05	>0.05	/	/								
<i>mcrA</i>	ρ	0.81	0.43	-0.75	-0.79	/	/						
	<i>p</i>	<0.05	>0.05	<0.05	<0.05	/	/						
Methane	ρ	-0.62	-0.02	0.41	-0.28	-0.48	0.24	/	/				
	<i>p</i>	<0.05	>0.05	>0.05	>0.05	<0.05	>0.05	/	/				
Salinity	ρ	-0.58	0.17	0.41	0.43	-0.64	-0.44	0.68	0.01	/	/		
	<i>p</i>	<0.05	>0.05	>0.05	>0.05	<0.05	>0.05	<0.05	>0.05	/	/		
SO ₄ ²⁻	ρ	-0.3	-0.14	0.40	0.50	-0.49	-0.38	-0.22	-0.23	0.11	0.20	/	/
	<i>p</i>	>0.05	>0.05	>0.05	<0.05	<0.05	>0.05	>0.05	>0.05	>0.05	>0.05	/	/

Functional gene relative abundances and gene ratios in surface sediment samples were obtained through qPCR analysis, providing insights into the spatial distribution of key metabolisms associated with the microbial methane cycle (Fig. S2.). We acknowledge that *dsrAB* gene abundances obtained with the here selected primer pair likely are an underestimation, due to biased primer coverage. However, the trends and ratios obtained are likely correct, as they align well with relative abundance data obtained from 16S rRNA gene amplicon sequencing.

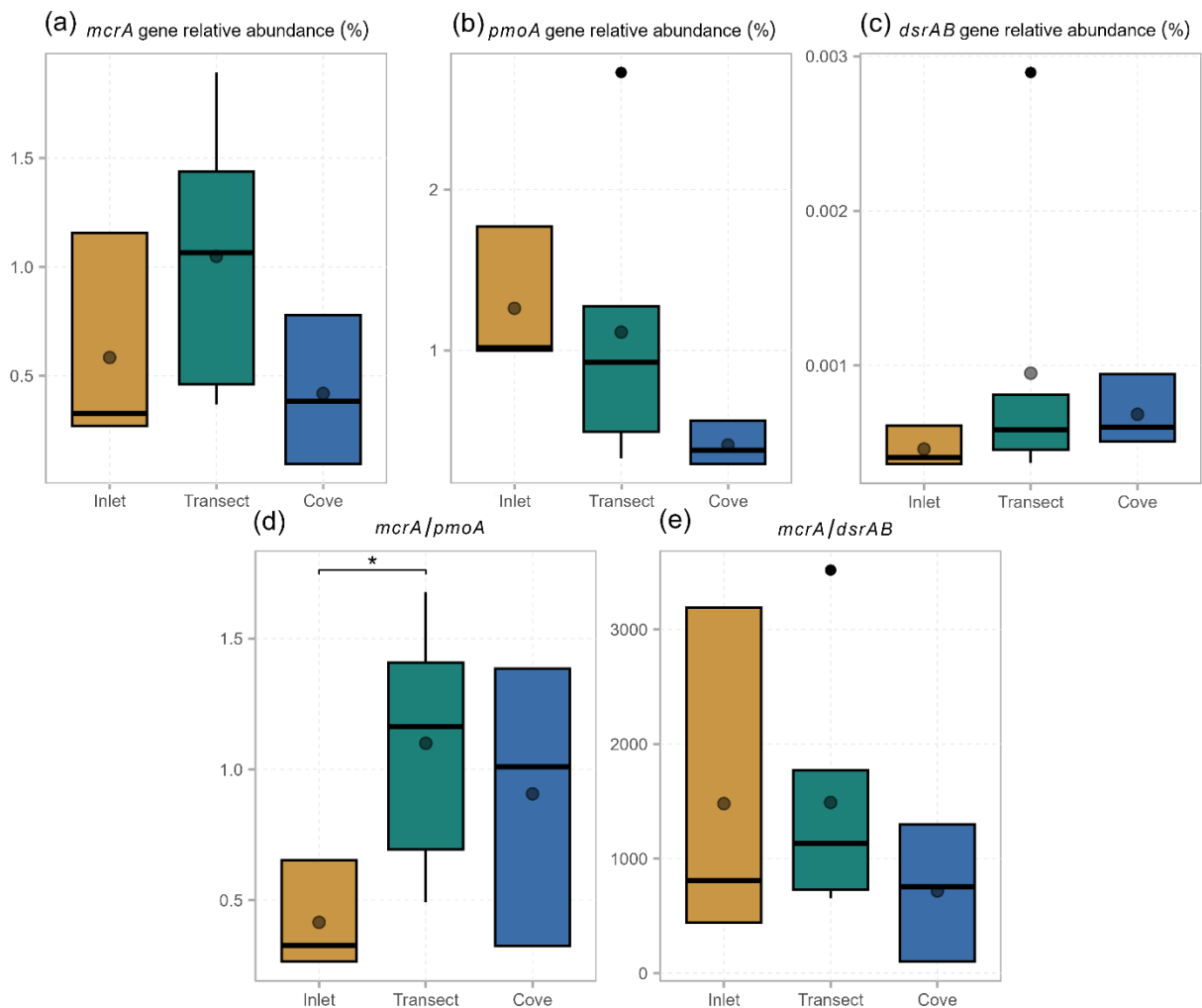


Figure S2. Gene relative abundance and gene ratios of select functional genes relevant to the methane cycle. Gene copy numbers were determined by qPCR in surface sediment samples and normalized by 16S rRNA gene copy numbers. The large black circle represents the mean value, the line indicates the median, while filled circles represent statistical outliers. Asterisks indicate statistically significant differences between group means where the adjusted $p < 0.05$.

Alpha diversity of key functional groups in the microbial methane cycle of surface sediments was assessed by categorizing them into geographical locations and calculating species richness and the Shannon index from rarefied 16S rRNA gene amplicon sequencing data (Fig. S3.). Relative abundance of key functional groups involved in the microbial methane cycle of surface sediments was categorized by geographical locations determined from 16S rRNA gene amplicon sequencing data (Fig. S4.). In the 16S rRNA gene sequencing datasets generated from surface sediment samples, methanotrophs, methanogens, and SRB were classified into 36, 26, and 52 ASVs,

respectively, constituting approximately 4.0%, 2.2%, and 5.4% of the microbial community, with average read count of 300. Their alpha diversity displayed similar values across the lake transect (Fig. S3.). Methanotroph diversity was highest in the inlet samples (stations 1 and 2, Fig. S3b., e.), while methanogen and SRB diversity remained relatively constant across all surface samples (Fig. S3a., d.), with no statistically significant difference between the locations. Amongst the methanogens, 7 orders with a mainly uniform spatial distribution across the surface sediment samples were detected (Fig. S4.). The exceptions were the *Methanobacterium* genus, predominantly present near the inlet locations (1-4). *Methanolinea* was the most prevalent methanogenic genus overall. Methanotroph surface sediment communities were dominated by ANME-1b in nearly all samples. Alongside ANME-1, the dominant methanotrophs included the groups of aerobic methanotrophs including Candidatus *Methylospira*, *Crenothrix*, *Methylocaldum*, and the clade Sh765B-Tzt-3. Dominant SRB groups included *Desulfatiglans*, *Desulfobacca*, and representatives of the Desulfobacterales order. PERMANOVA results indicated that salinity influences the spatial distribution of this functional group (Table S2.). A Mantel test confirmed

variations in geographic distribution of salinity, *pmoA* and *mcrA* gene copy numbers, as well as the distribution of methanotrophs and SRB groups across the lake transect (Table S3.).

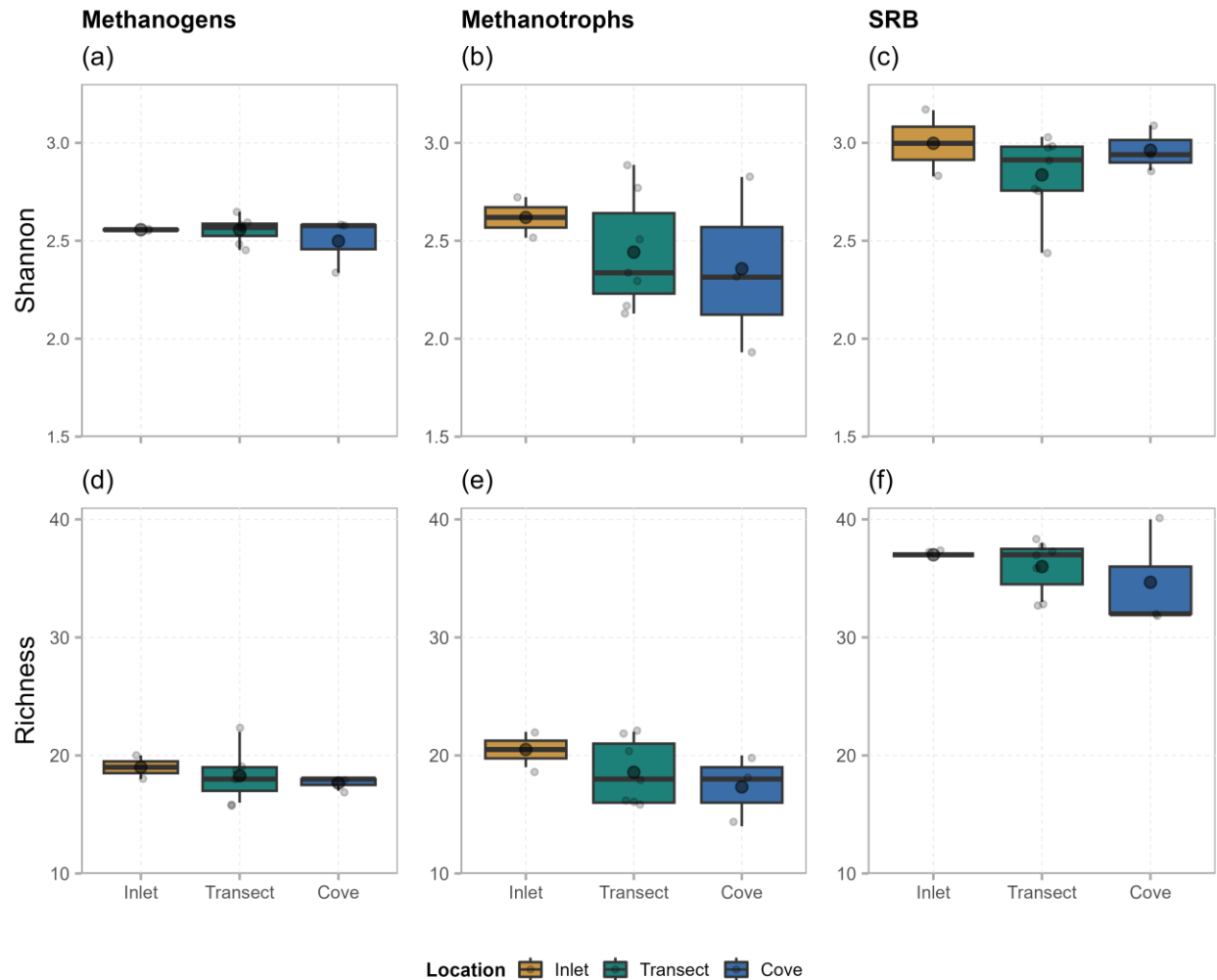


Figure S3. Alpha diversity of selected functional groups in surface sediments determined based on the Shannon diversity index and species richness divided by geographical location (inlet, transect, cove).

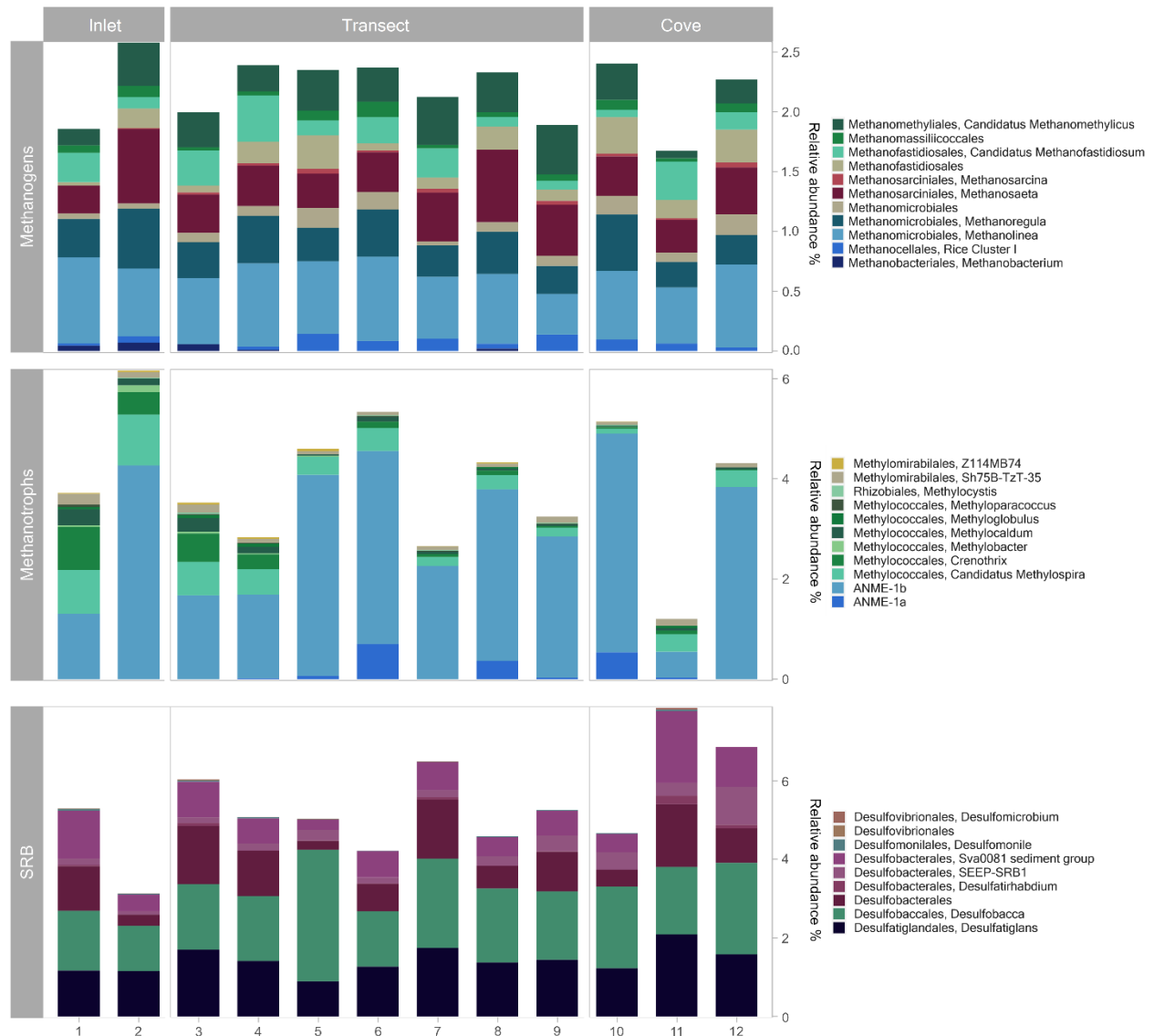


Figure S4. The relative abundance of methanogens, methanotrophs, and SRB communities in surface sediments at the lowest assigned taxonomy level. The numbers represent different sampling stations (1-12, see Fig. 1a.) belonging to the assigned geographical location.

Distance – decay analysis suggested decreasing surface sediment community similarity with increasing geographic distance for methanotrophs and SRB (Fig. S5).

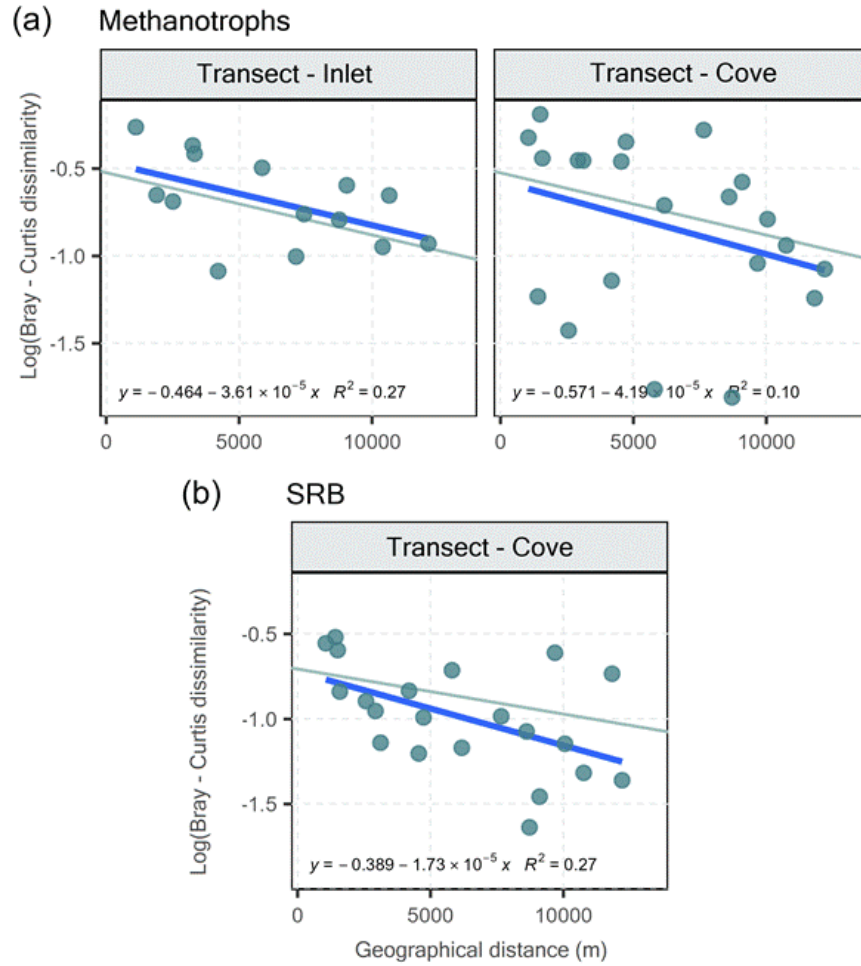


Figure S5. The results of distance-decay analysis between communities of a) methanotrophs and b) SRB from different geographical locations. The blue line represents the linear model of the subset of samples, while the gray-blue line depicts the linear regression that includes all samples ($p < 0.05$).

The PERMANOVA test was employed to examine the associations between microbial functional group 16S rRNA gene amplicon sequence dissimilarity and physicochemical parameters in surface sediments (Table S2.).

Table S2. Results of PERMANOVA test on Bray - Curtis dissimilarity matrices of 16S rRNA gene amplicon sequence data from each functional group (ASVs) and physicochemical parameters of surface sediments

		Methanotrophs	Methanogens	SRB
Salinity	R^2	0,168	0,136	0,253
	p	0,091	0,153	0,016
NH₄⁺	R^2	0,309	0,226	0,178
	p	0,005	0,011	0,067
NO₃⁻	R^2	0,098	0,178	0,228
	p	0,354	0,037	0,013
SO₄²⁻	R^2	0,125	0,149	0,175
	p	0,208	0,106	0,05
PO₄³⁻	R^2	0,135	0,078	0,108
	p	0,168	0,539	0,293

Table S3. Results of Mantel test based on Bray - Curtis dissimilarity matrices of 16S rRNA gene amplicon sequence data from each functional group (ASVs), gene copy numbers, and physicochemical parameters across geographical locations (inlet, transect, cove) of surface sediments.

	R^2	p
Salinity	0,226	0,041
NH ₄ ⁺	0,071	0,244
NO ₃ ⁻	-0,020	0,498
SO ₄ ²⁻	0,162	0,081
PO ₄ ³⁻	-0,032	0,558
<i>pmoA</i>	0,418	<0,001
<i>mcrA</i>	0,390	0,005
<i>dsrAB</i>	0,149	0,101
Methanotrophs	0,554	<0,001
Methanogens	0,155	0,108
SRB	0,239	0,034

Alpha diversity of select functional groups relevant to the methane cycle of sediment profiles was assessed by calculating species richness and the Shannon index from 16S rRNA gene amplicon sequence data (Fig. S6a.), and samples from deeper freshwater and salinization zones clustered separately in all three functional groups (Fig. S6b.).

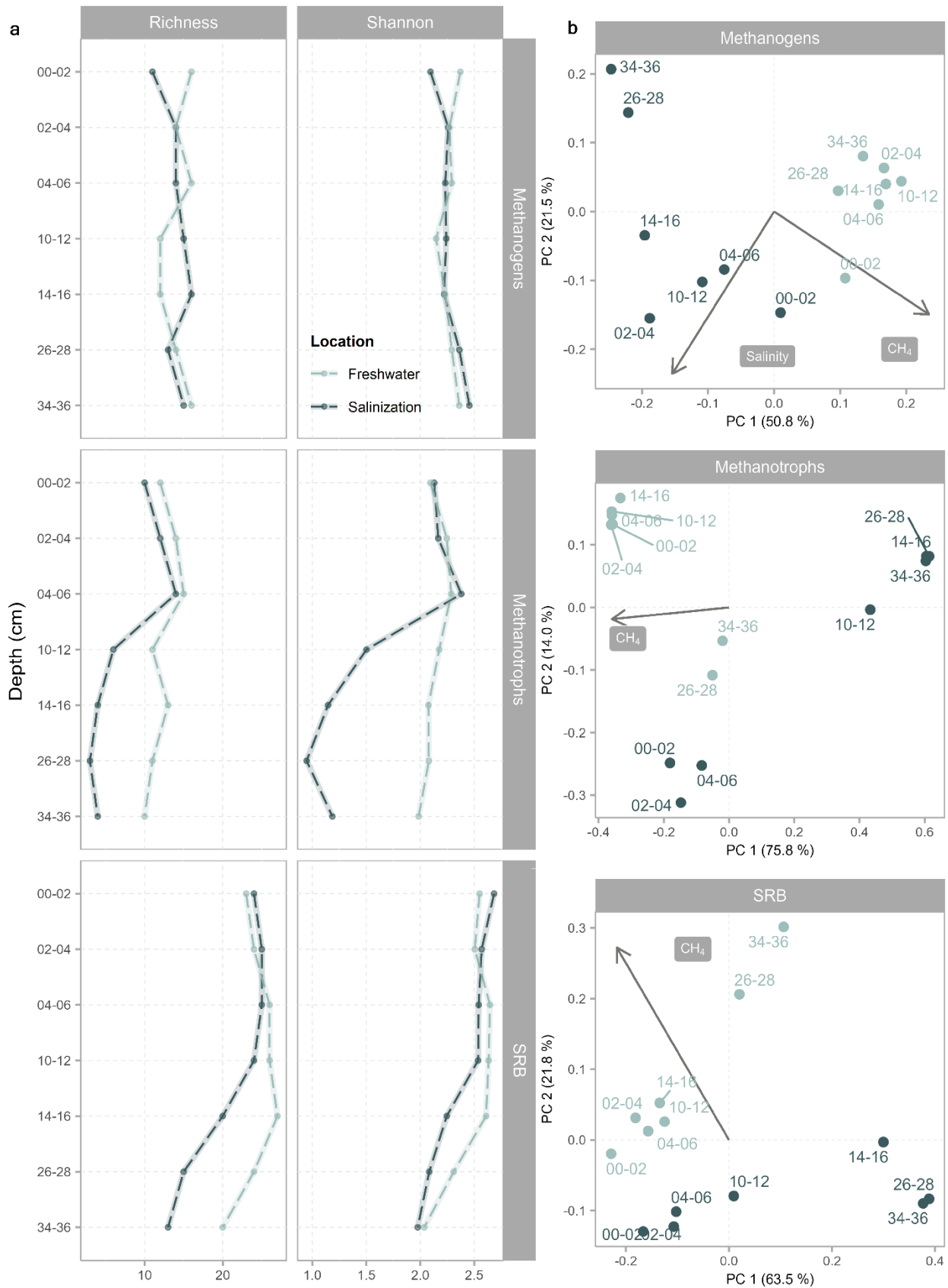
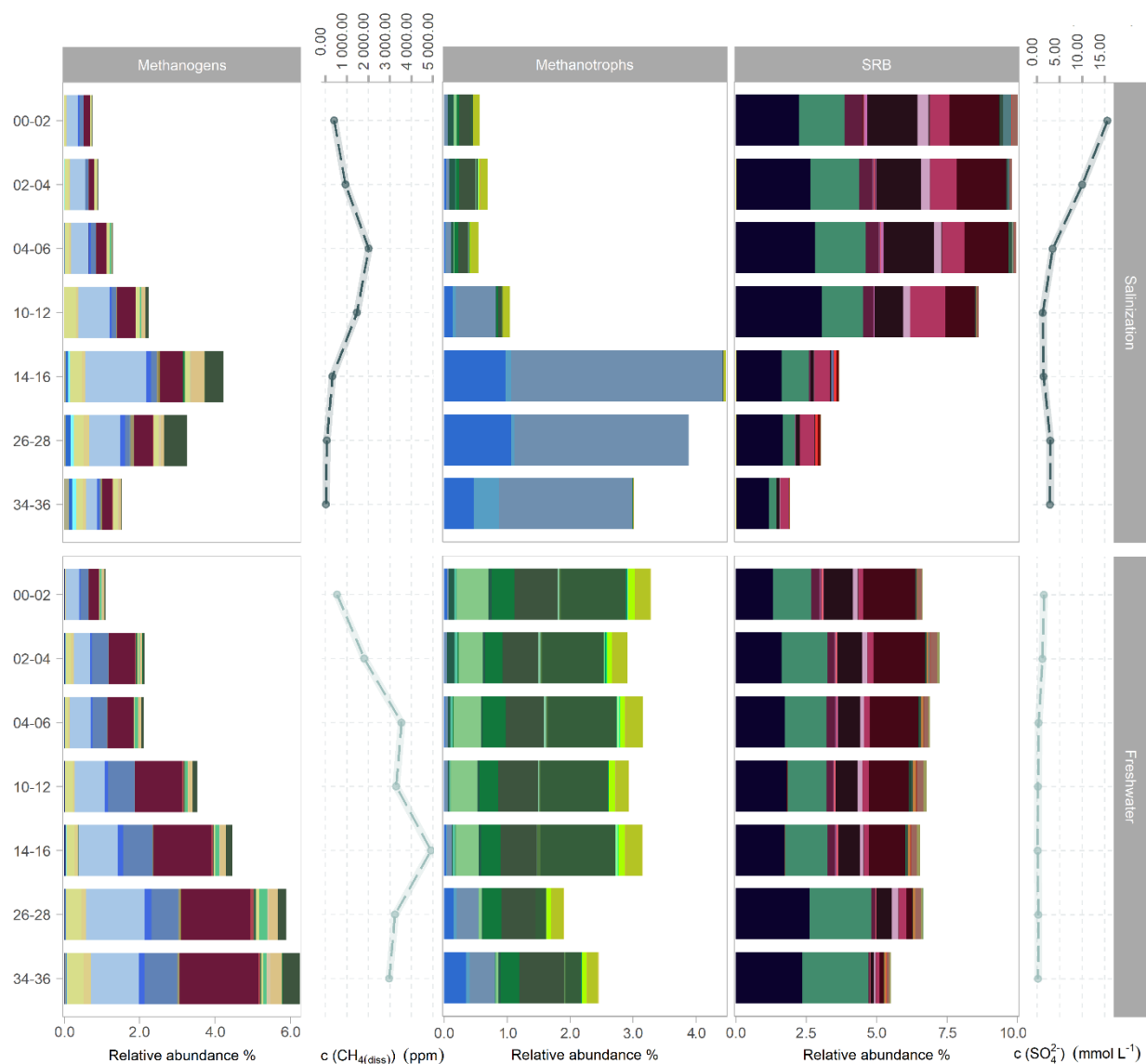


Figure S6. a) Alpha diversity of select functional groups in sediment profiles determined based on the Shannon diversity index and species richness of 16S rRNA gene amplicon sequence data in the salinization profile (Jugovir) and the freshwater profile (Lateral canal inlet). b) PCoA clustering of microbial communities based on Bray-Curtis dissimilarity matrix of 16S rRNA gene sequence data derived with PhyloFLASH from metagenomes that are taxonomically affiliated with known methanogens, methanotrophs, and SRB. The communities are arranged according to the depth interval and sediment profile: salinization profile (Jugovir; dark green) and freshwater profile (Lateral canal inlet; light green). Arrows represent physicochemical parameters significantly associated with the respective microbial communities.

Relative abundances of microbial communities in sediment profiles from selected functional groups, derived from 16S rRNA gene data in metagenomes (Fig. S7.). Relevant geochemical profiles of dissolved methane and sulfate concentrations were added to the corresponding relative abundance profiles: methanogens, methanotrophs, and SRB.



Methanogens

- Methanobacteriales, Methanobacterium
- Methanocellales
- Methanocellales, Methanocellaceae
- Methanocellales, Methanocella
- Methanomicrobiales, Methanoculleus
- Methanocellales, Rice Cluster I
- Methanomicrobiales
- Methanomicrobiales, Methanomicrobiaceae
- Methanomicrobiales, Methanogenium
- Methanomicrobiales, Methanolinea
- Methanomicrobiales, Methanoplanus
- Methanomicrobiales, Methanoregulaceae
- Methanomicrobiales, Methanoregula
- Methanomicrobiales, Methanospirillum
- Methanosarcinales
- Methanosarcinales, Methanosarcinaceae
- Methanosarcinales, Methanosarcina
- Methanosarcinales, Methanolobus
- Methanosarcinales, Methanomethylovorans
- Methanofastidiosales
- Methanofastidiosales, Candidatus Methanofastidiosum
- Methanomassilicoccales
- Methanomassilicoccales, Methanomassilicoccaceae
- Methanomassilicoccales, Methanomassilicoccus
- Methanomethyliales, Candidatus Methanomethylicus

Methanotrophs

- ANME-1
- ANME-1, ANME-1a
- ANME-1, ANME-1b
- Methylacidiphilales, Methylacidiphilaceae
- Methylococcales
- Methylococcales, Candidatus Methylospira
- Methylococcales, Crenothrix
- Methylococcales, Methylobacter
- Methylococcales, Methylocaldum
- Methylococcales, Methylococcaceae
- Methylococcales, Methyloglobulus
- Methylococcales, Methylochromium
- Methylococcales, Methylomonadaceae
- Methylococcales, Methylomonas
- Methylococcales, Methyloparacoccus
- Methylococcales, Methyloprofundus
- Methylococcales, Methylothericola
- Rhizobiales, Methylocystis
- Methylomirabilales
- Methylomirabilales, Rokubacteriales
- Methylomirabilales, Sh765B-TzT-35
- Methylomirabilales, Z114MB74

SRB

- Desulfarculariaceae bacterium JdFR-97
- Desulfatiglandales, Desulfatiglans
- Desulfotobacteriales, TC1
- Desulfobacterales, Desulfobacca
- Desulfobacterales
- Desulfobacterales, Desulfatiferula
- Desulfobacterales, Desulfatirhabdium
- Desulfobacterales, Desulfatitalea
- Desulfobacterales, Desulfobacteraceae
- Desulfobacterales, Desulfococcaceae
- Desulfobacterales, Desulfococcus
- Desulfobacterales, Desulfonema
- Desulfobacterales, Desulfosarcina
- Desulfobacterales, Desulfosarcinaceae
- Desulfobacterales, Incertae Sedis
- Desulfobacterales, LCP-80
- Desulfobacterales, SEEP-SRB1
- Desulfobacterales, Sva0081 sediment group
- Desulfobulbales
- Desulfobulbales, Desulfobulbus
- Desulfobulbales, Desulfocapsaceae
- Desulfobulbales, Desulforhopalus
- Desulfobulbales, MSBL7
- Desulfomoniales, Desulfomonie
- Desulfotomaculales, Cryptanaerobacter
- Desulfotomaculales, Pelotomaculaceae
- Desulfovibrionales, Desulfomicrobium
- Desulfovibrionales, Desulfovibrio
- Desulfuromonadales
- Desulfuromonadales, Desulfuromonadaceae
- Desulfuromonadales, Geothermobacter
- Desulfuromonadales, Syntrophotalea
- Syntrophobacteriales, Syntrophobacter
- Syntrophorhabdiales, Syntrophorhabdus
- Syntrophomonadales, Syntrophomonas

Figure S7. The relative abundance of methanogens, methanotrophs, and SRB communities in sediment profiles at the lowest assigned taxonomy level along with dissolved methane and sulfate concentrations. The numbers represent different sampling depth intervals (cm) belonging to the designated sediment profile: salinization profile and freshwater profile.

Basic genome statistics of all MAGs containing the *mrcA* gene. Genomes were examined for the presence of genes related to methanogenesis and methanotrophy to evaluate pathway completeness (Fig. S8.).

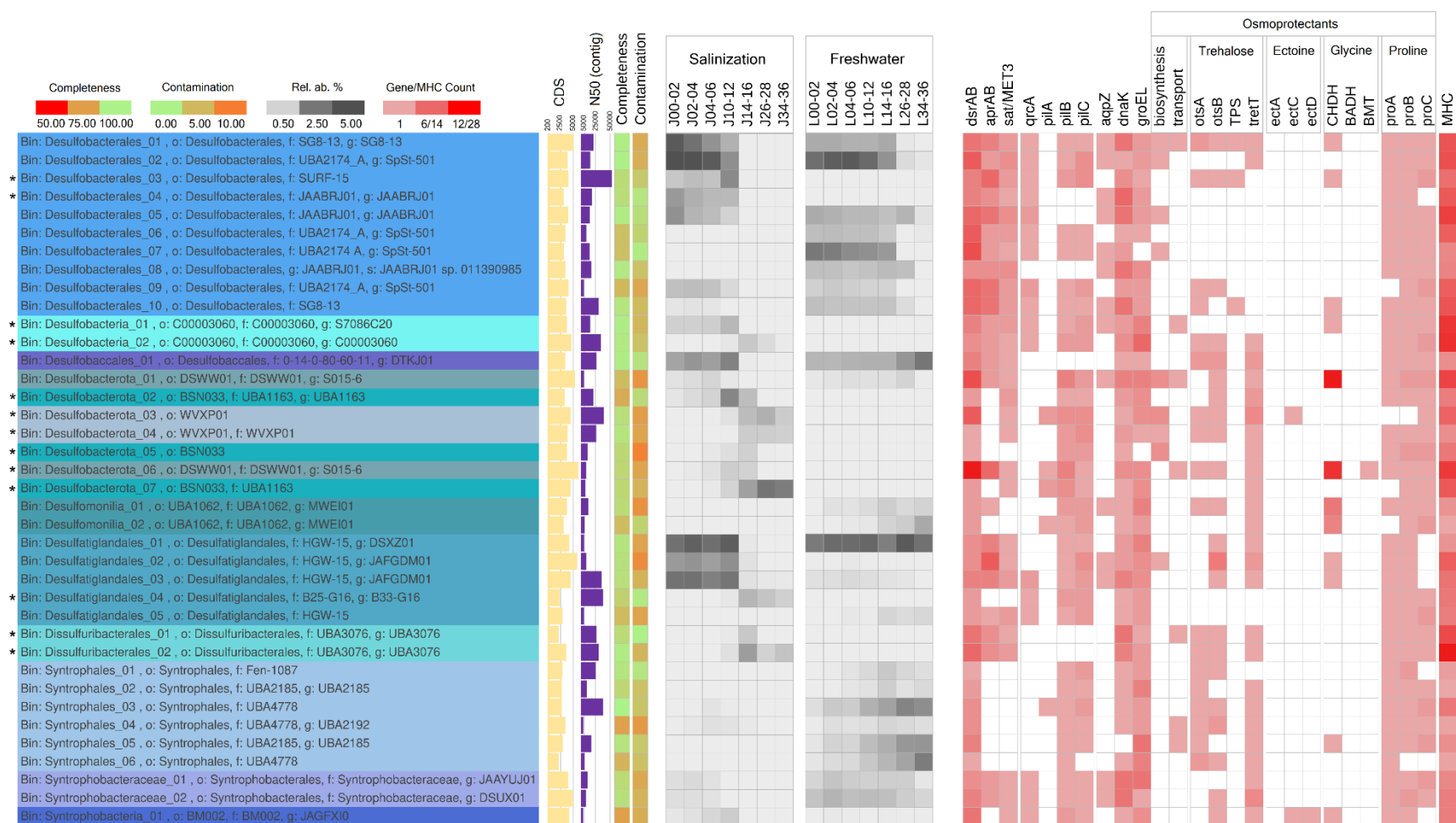


Figure S8. MAGs (Metagenome-Assembled Genomes) (completeness > 50.00%, contamination < 10.00%) containing the *dsrAB* gene with assigned GTDB taxonomy, number of coding sequences (CDS), N50, completeness and contamination levels, relative abundance in the salinization and freshwater sediment profiles, and functional gene annotations of key genes involved in dissimilatory sulfate reduction, interspecies electron transfer, and osmoregulation processes, respectively. The asterisk (*) denotes MAGs exclusive to salinization profile.

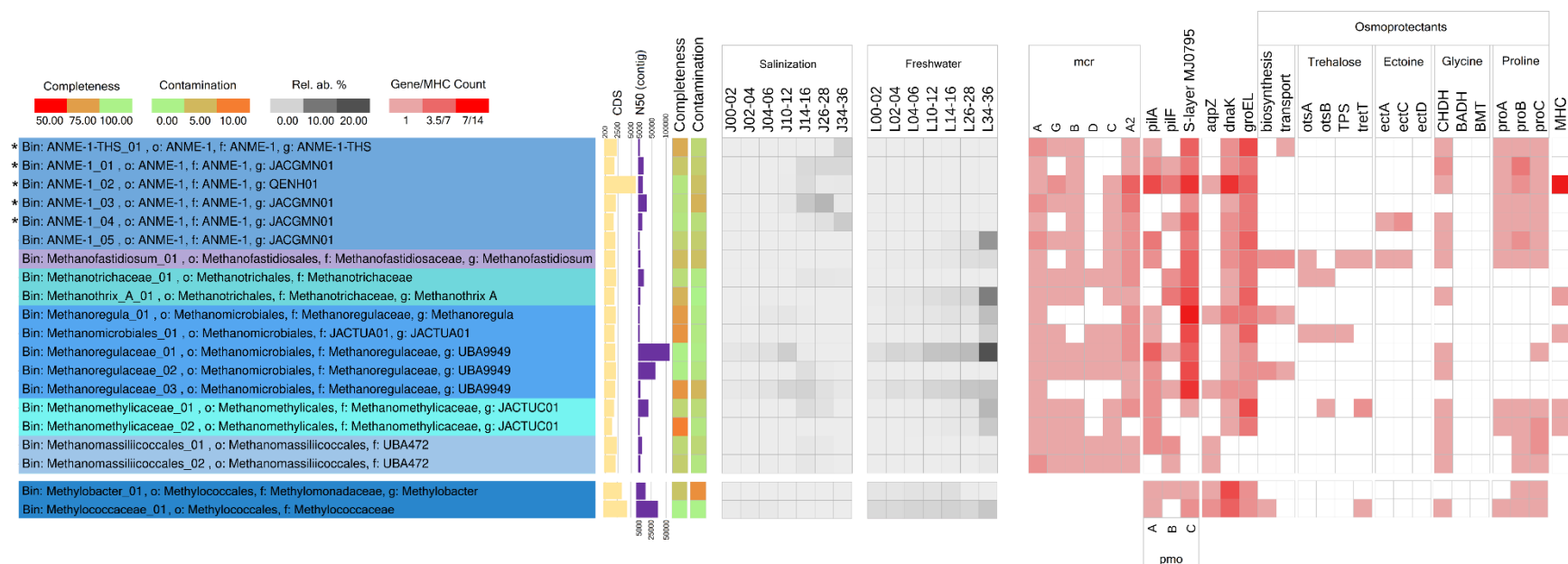


Figure S9. MAGs (Metagenome-Assembled Genomes) (completeness > 50.00%, contamination < 10.00%) based on *mcrA* gene (for methanogens and ANME affiliated MAGs) and *pmoA* gene (for aerobic methanotrophs) with assigned GTDB taxonomy, number of coding sequences (CDS), N50, completeness and contamination levels, relative abundance in the salinization and freshwater sediment profiles, and functional gene annotations of key genes involved in methanogenesis/methanotrophy, interspecies electron transfer, and osmoregulation processes, respectively. The asterisk (*) denotes MAGs exclusive to salinization profile.

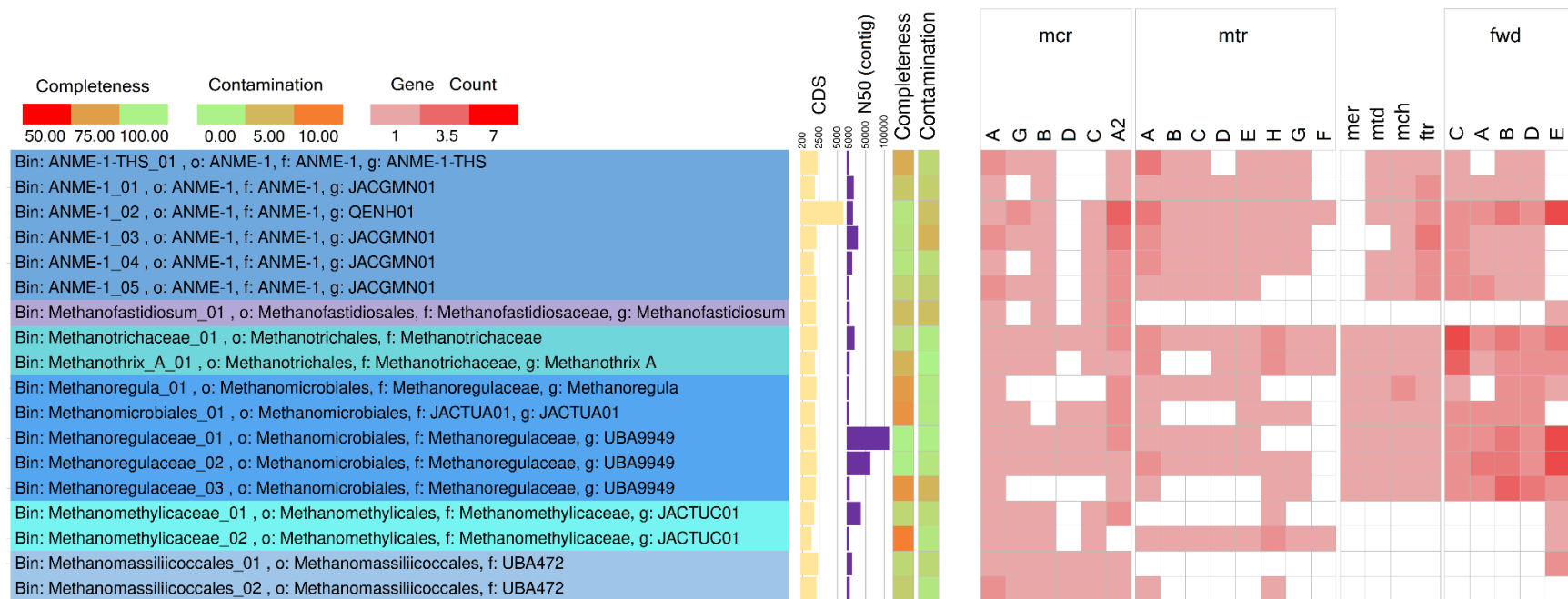


Figure S10. MAGs (completeness > 50.00%, contamination < 10.00%) present in deeper sediments containing the *mcrA* gene (methanogens and ANME affiliated MAGs) recovered from sediment profile samples, including assigned GTDB taxonomy, number of coding sequences (CDS), N50, completeness and contamination levels, and presence/absence of annotated functional gene involved in methanogenesis/reverse methanogenesis.

References

- [1] Bushnell, B. BBMap sourceforge.net/projects/bbmap/, 2015.
- [2] Callahan BJ, Sankaran K, Fukuyama JA, McMurdie PJ, Holmes SP. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000Res*. 2016;5:1492.
- [3] Pruesse E, Peplies J, Glöckner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *J Bioinform*. 2012;28(14):1823–1829.
- [4] McMurdie PJ, Holmes S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 2013;8:e61217.
- [5] Oksanen J, Blanchet FG, Friendly M, et al. vegan: Community Ecology Package. R package version 2.5-6. 2019. URL: <https://cran.r-project.org/package=vegan>.
- [6] Lahti L, Shetty S, Blake T, Salojärvi J. microbiome R package. *Tools Microbiome Anal R*. 2017. URL: <https://microbiome.github.io/>.
- [7] Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.
- [8] Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *J Bioinform*. 2015;31(10):1674–1676.
- [9] Kang DD, Li F, Kirton E, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7:e7359.
- [10] Sieber CMK, Probst AJ, Sharrar A, et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol*. 2018;3:836–843.
- [11] Olm MR, Brown CT, Brooks B, Banfield JF. dRep: A tool for fast and accurate genome de-replication that enables tracking of microbial genotypes and improved genome recovery from metagenomes. *ISME J*. 2017;11:2864–2868.
- [12] Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *bioRxiv*. 2022;11.499243.
- [13] Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *J Bioinform*. 2020;36(6):1925–1927.
- [14] Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun*. 2016;7:11257.
- [15] Eren AM, Kiefl E, Shaiber A, et al. Community-led, integrated, reproducible multi-omics

with anvi'o. Nat Microbiol. 2021;6:3–6.

[16] Jones P, Binns D, Chang HY, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236–1240.

[17] Seiler M, Mehrle A, Poustka A, Wiemann S. The 3of5 web application for complex and comprehensive pattern matching in protein sequences. BMC Bioinformatics. 2006;7:144.

[18] Yu NY, Wagner JR, Laird MR, et al. PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics. 2010;26(13):1608–1615.

[19] Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.

[20] Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25(15):1972–1973.

[21] Lee MD. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics. 2019;35(20):4162–4164.

[22] Letunić I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics. 2006;23(1):127-8.

[23] Gruber-Vodicka HR, Seah BKB, Pruesse E. phyloFlash: Rapid small-subunit rRNA profiling and targeted assembly from metagenomes. mSystems. 2020;5(5):00920-20.

[24] Faust K, Raes J. CoNet app: inference of biological association networks using Cytoscape. F1000 Research. 2016; 5:1519.

[25] Watts, S. C., Ritchie, S. C., Inouye, M., & Holt, K. E. FastSpar: rapid and scalable correlation estimation for compositional data. Bioinformatics. 2018; 35(6):1064-1066.

[26] Bastian M, Heymann S, Jacomy M. Gephi: an open-source software for exploring and manipulating networks. In: Third international AAAI conference on weblogs and social media. 2009.

[27] Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech. 2008; 10:P1008.