Unravelling Biosynthesis and Biodegradation Potentials of Microbial Dark Matters in Hypersaline Lakes

Zhiguang Qiu[1,2], Yuanyuan Zhu[1], Qing Zhang[1], Xuejiao Qiao[1], Rong Mu[1], Zheng Xu[3,4], Yan Yan[5], Fan Wang[6], Tong Zhang[7], Ke Yu*[1,2]

[1] *School of Environment and Energy, Peking University Shenzhen Graduate School, Shenzhen, 518055, China*

[2] *AI for Science (AI4S)-Preferred Program, Peking University, Shenzhen, 518055, China*

[3] *Southern University of Sciences and Technology Yantian Hospital, Shenzhen, 518081, China*

[4] *Institute of Biomedicine and Biotechnology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, China*

[5] *State Key Laboratory of Isotope Geochemistry, CAS Center for Excellence in Deep Earth Science, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, China*

[6] *School of Atmospheric Sciences, Sun Yat-sen University, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519082, China*

[7] *Department of Civil Engineering, University of Hong Kong, Hong Kong, 999077, China*

*Corresponding author: Ke Yu; E-mail address: yuke.sz@pku.edu.cn

Highlights

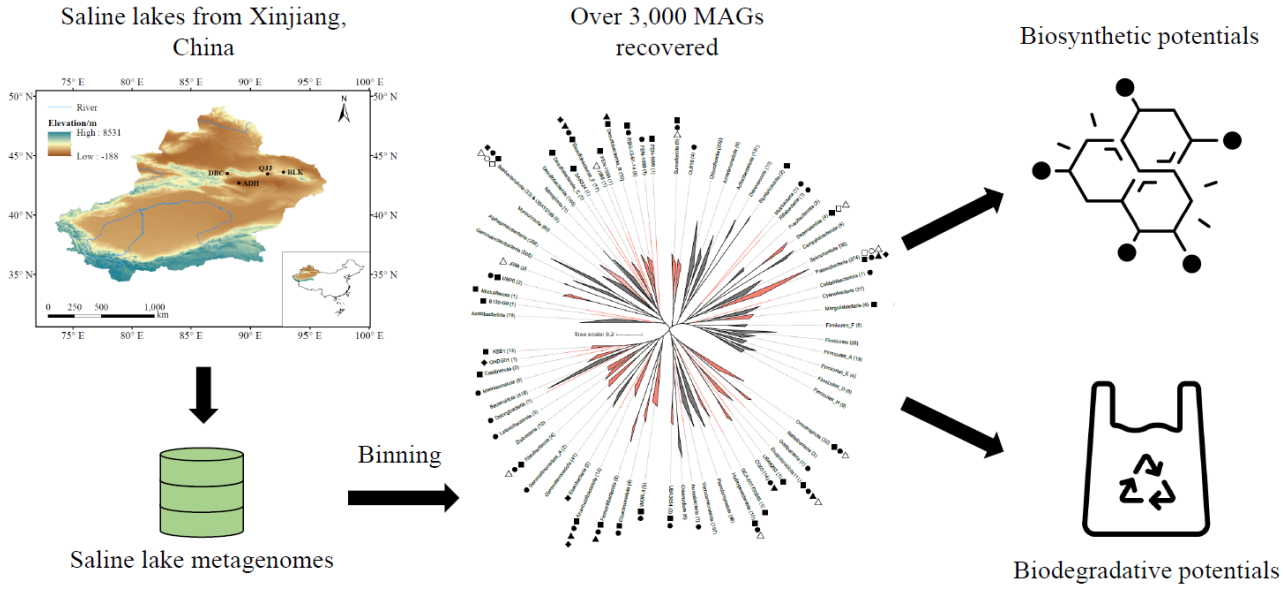
- Over 3,000 MAGs obtained from hypersaline lakes that enriching genomic resources.
- Identified 8,000+ potential biosynthetic gene clusters in uncultured microbes.
- Uncovered biodegradation potential in several microbial dark matter lineages.
- Key insights for biotechnology, environmental management applications.
- Versatile and resilient microorganisms in extreme environments.

Abstract

Biosynthesis and biodegradation of microorganisms critically underpin the development of biotechnology, new drugs and therapies, and environmental remediation. However, the vast majority of uncultured microbial species along with their metabolic capacities in extreme environments remain obscured. To unravel the metabolic potential of these microbial dark matters (MDMs), we investigated four deep-inland hypersaline lakes with largely diversified environmental parameters in Xinjiang,

China. Metagenomic binning obtained 3,030 metagenome-assembled genomes (MAGs) spanning 82 phyla, of which 2,363 MAGs could not be assigned to a known genus. MDM MAGs were abundantly observed with distinct taxa among lakes, possibly linked to the diversification of physiochemical conditions. Analysis of biosynthetic potentials identified 9,635 biosynthesis gene clusters (BGCs), in which 8,966 BGCs considered novel. We found that some MDM phyla consistently comprised enriched BGCs, which were considered to play essential roles in biotechnological applications. In addition, biodegradation potentials such as dehalogenation, anaerobic ammonium oxidation (Anammox), polycyclic aromatic hydrocarbon (PAH), and plastics degradation were found in new microbial clades from hypersaline lakes. These findings substantially expanded the genetic repository of biosynthesis and biodegradation potentials, which can further assist the development of new and innovative applications in biotechnology.

Graphical abstract



Keywords

Biosynthesis; Biodegradation; Hypersaline Lakes; Metagenome-assembled Genomes; Microbial Dark Matters

1. Introduction

Microorganisms are vital in the global ecosystems that underpin life on Earth, as they play essential roles contributing to biogeochemical cycling, supporting food webs, and maintaining the fitness of plants and animals (Bahramet al. 2018; Falkowskiet al. 2008; Loucaet al. 2016; Senderet al. 2016). Particularly, the microbial metabolic capabilities such as decomposing pollutants and producing secondary metabolites like natural products have fostered the development of biosynthesis and biodegradation, thereby propelling biotechnological processes and addressing environmental challenges (Barabadiet al. 2019; Chemler and Koffas 2008; Wanget al. 2014; Zhanget al. 2022). Nonetheless, a vast majority of the microbial species are unculturable, leading to the inefficient exploration of metabolic capacities across diverse environments.

To address this issue, a series of state-of-art sequencing technologies along with bioinformatic tools, such as metagenomic sequencing and binning, have revolutionised our ability to investigate and decode the metabolic pathways of uncultured microorganisms. Employing such approaches, we can unveil previously uncharacterised microorganisms within underexplored environment, considerably expanding our recognition of the specific microbial roles and functions in biosynthesis and biodegradation. This knowledge could be instrumental in harnessing more efficient and sustainable methods for producing natural products or devising more effective pathways for contaminant degradation (Zallotet al. 2021). However, current progress in the application of microbial resources for biosynthesis is still hindered by the efficiency of the enzymes for large-scale production, sustainability of methodologies, and the financial demands of culturing (Saravananet al. 2021; Wuet al. 2014). On the other hand, environmental variability, complexity of microbial communities, and the high costs of pre-treatment are main issues that obstructed the efficiency of biodegradation (Azubuikeet al. 2016; Taghaviet al. 2021). Furthermore, our understanding of the underlying microbial mechanisms of biosynthesis and biodegradation, including novel enzymes and metabolic pathways involved in these processes, are still largely unknown. Therefore, to enhance our understanding of these mechanisms and facilitate more efficient and effective biosynthesis and biodegradation processes, it is imperative to investigate uncultured microbial species and their metabolic capacities in extreme environments. Such exploration can deepen the knowledge of novel enzymes and metabolic pathways, leading to more sustainable and environmentally friendly solutions across a wide range of applications.

Hypersaline lakes are typical examples of extreme environment, which usually characterised by high salinity (≥ 35 g/L), elevated aridity and evaporation, poor in nutrients such as organic carbon, and some with high alkalinity (Sorokinet al. 2014; Vavourakiset al. 2018; Ventosaet al. 2015). These habitats accommodate distinct extremophilic microbial communities that have evolved to adapt the conditions inhospitable to other living organisms (Rothschild and Mancinelli 2001; Shu and Huang 2022). Deep

inland hypersaline lakes usually exhibit distinct environmental parameters from lake to lake due to the influences of climate, geology, and human activities. Consequently, a wide range of microbial communities residing in these hypersaline habitats remain unexplored, presenting potential resources for biotechnological applications (Krienitz and Kotut 2010; Mesbah and Wiegel 2012; Rojaset al. 2018; Schultz and Rosado 2020). Genome-resolved analyses indicates that these deep inland hypersaline lakes harbour a large number of uncultured microorganisms, namely microbial dark matter (MDM) such as DPANN archaea and CPR bacteria (Vavourakiset al. 2018; Xieet al. 2022; Zhaoet al. 2020). These findings substantially broadened our recognition of MDM from an evolutionary perspective. However, these pioneering studies merely scratched the surface of hypersaline lake microbiomes, restricted by the scope of study areas and limitations in bioinformatic analysis techniques. Therefore, to comprehensively uncover the microbial niches within hypersaline lake microbiomes and further reveal metabolic capacities in these extreme ecosystems, more in-depth and expansive investigations on hypersaline lake microbial genomes should be envisaged.

In this study, we examined four saline lakes in Xinjiang Uygur Autonomous Region, China, each with distinct salinity (1.5 g/L to 237 g/L) and altitude (-153 m to 1,585 m). Previous study observed that salinity significantly influences the microbial assemblages of saline lake microbiota (Jiet al. 2019), but the roles that lake planktonic and sediment microbial communities devote to biosynthesis and biodegradation resolved from metagenomes remain uncertain. Given such a diverse range of geographic and environmental parameters, we assembled over 3,000 Metagenome-Assembled Genomes (MAGs) from these four lakes, aiming to unearth new microbial lineages and decode the metabolic characteristics of hypersaline lake prokaryotes.

2. Methods

2.1. Sampling and measurement of physiochemical parameters

Hypersaline lake samples were collected in July 2018 from Aiding Lake (ADH), Barkol Lake (BLK), Dabancheng Lake (DBC) and Qijiaojing Lake (QJJ) (Fig. S1), Xinjiang Uygur Autonomous Region, China. The detailed sampling and geographic information were recorded in Table S1. For water samples from each lake (n = 4), a total amount of 5 L of lake water from the upper 50 cm surface was randomly collected into a 5 L sterile container. Lake water was firstly filtered through a filter paper (Whatman, GE Healthcare, NY, USA) to remove large particles before filtering through 0.22 μm PES (Polyethersulfone) membranes (Millipore, Billerica, MA, USA) to enrich planktonic microorganisms. For sediment samples from each lake, about 50 grams of sediment samples (n = 4) were randomly collected at 0-10 cm depth at the bottom of lake into sterile 50 ml falcon tubes. Membranes and

119 sediment samples were immediately placed on dry ice before brought to laboratory transferred to a -
120 80 °C freezer until further DNA extraction was performed.

121 The salinity and pH of lake water were measured *in situ* using a Hydrolab DS5 multiparameter water
122 quality meter (Hach Company, Loveland, CO, USA). Total organic carbon (TOC) was analyzed with
123 a TOC/TN-VCPH analyzer (Shimadzu, Tokyo, Japan). The concentrations of lithium (Li^+), sodium
124 (Na^+), ammonium (NH_4^+), potassium (K^+), magnesium (Mg^{2+}), calcium (Ca^{2+}), chloridion (Cl^-) and
125 sulfate (SO_4^{2-}) were determined using an inductively coupled plasma mass spectrometry (Agilent
126 Technologies Inc., Bellevue, WA, USA).

127 **2.2. DNA extraction and metagenomic sequencing**

128 Total DNA was extracted from frozen filtered membrane or sediment samples using the DNeasy
129 PowerSoil Pro Kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. Extracted
130 DNA was quality checked by NanoDrop 2000 (Thermo Fisher Scientific, Waltham, Massachusetts,
131 US) and quantity checked by Qubit Fluorometer (Thermo Fisher Scientific, USA). Next generation
132 metagenomic sequencing was performed at Novogene (Tianjin, China) on an Illumina NovaSeq 6000
133 platform using the PE150 strategy.

134 **2.3. Assembly, binning and phylogenetic analyses**

135 Raw paired-end reads were initially filtered using fastp with default parameters (Chenet al. 2018).
136 Clean reads were individually and co-assembled using SPAdes (v3.15.2) (Bankevichet al. 2012) under
137 "--meta" mode into contigs specifying k-mer sizes of 21, 33, 55, 77 and finally reserved contigs >
138 1,000 bps. Metagenomic binning was applied to both single sample assembly and co-assembly by
139 using BASALT (v1.0) under default mode (Yuet al. 2021). The completeness and contamination of
140 bins were assessed using CheckM v1.0.2 with "lineage_wf" workflow and default parameters (Parkset
141 al. 2015), with only medium and high qualitied bins that meet the MIMAG standard (completeness ≥
142 50, contamination ≤ 10) (Bowerset al. 2017) kept as metagenomic-assembled genomes (MAGs) for
143 further analysis.

144 Phylogenetic analysis of bacteria and archaea MAGs were performed based on a multiple sequence
145 alignment of 120 bacterial- and 122 archaeal-specific single-copy marker proteins (Chaumeilet al.
146 2020), respectively. A concatenated alignment of these marker proteins was created using HMMER
147 v3.1.b2 with default parameters (Eddy 2011). Bacterial and archaeal hylogenetic trees were inferred
148 using IQ-TREE under the best fitted models with 1,000 bootstrap replications (Minhet al. 2020), before
149 being visualized and annotated in iTOL v6 (Letunic and Bork 2021).

150 **2.4. Identification of non-coding RNA genes and functional annotation**

151 Non-coding RNA genes including transfer RNA genes and ribosomal RNA genes were identified to
152 evaluate the integrity of MAGs. Briefly, tRNA genes were predicted using tRNAscan-SE (v2.0.9)
153 (Chanet al. 2021), while rRNA genes including 5S rRNA, 16S rRNA, and 23S rRNA were predicted
154 using Barrnap (v0.9, https://github.com/tseemann/barrnap). For genetic analysis, Opening Reading
155 Frames (ORFs) of assembled contigs were predicted using Prodigal (v2.6.3) with default parameters
156 (Hyattet al. 2010). Predicted ORFs were dereplicated with CD-HIT (v4.6) at 90% identity (Li and
157 Godzik 2006). Dereplicated ORFs were annotated against Kyoto Encyclopedia of Genes and Genomes
158 (KEGG) database (version 58) and NCBI-NR database (version 20220806), respectively with
159 Diamond (v2.0.11.149) (Buchfinket al. 2015) using "more-sensitive" mode and e-value at 1e-10. In
160 addition, dereplicated ORFs were further annotated against Pfam (v34.0) and TIGRFAM (v15.0)
161 databases using InterProScan (v5.54-87.0) (Joneset al. 2014). To estimate the relative abundance of
162 each genes, dereplicated fasta sequences of ORFs were mapped against raw sequence reads by Bowtie2
163 (v2.4.2) (Langdon 2015). Then, the depth of aligned reads were calculated using SAMtools (v1.7)
164 (Liet al. 2009) and summarised with a perl script 'calc.coverage.in.bam.depth.pl' from a previous study
165 (Albertsenet al. 2013). Finally, coverage-normalised KOs were clustered into modules and pathways
166 using a python script "pathway_pipeline.py" in PICRUSt2 (Douglaset al. 2020). For genome-centric
167 analysis, each MAG was predicted using Prodigal with default parameters and being annotated against
168 KEGG database using the same parameter described above.

169 **2.5. Annotation of Biosynthetic Gene Clusters (BGCs)**

170 Genome sequences were used as input using antiSMASH (v.6.1.1) (Blinet al. 2021) to identify BGCs
171 from saline lake MAGs. The gbk-formatted output of each MAG was further processed using BiG-
172 SCAPE (v1.1.2) (Navarro-Muñozet al. 2020) against MIBiG database (Kautsaret al. 2020) to cluster
173 Gene Cluster Families (GCFs) at 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 distance threshold, respectively.
174 To determine the novelty of BGCs, we followed the criteria described in Paoli *et al.* (Paoliet al. 2022),
175 that any BGCs with cosine distance larger than 0.2 to a known BGC in the MIBiG database were
176 considered novel.

177 **2.6. Detection of Biodegradation potentials**

178 To identify MAGs with potential biodegradation functions including dechlorinating, nitrogen removal,
179 and degradation of polycyclic aromatic hydrocarbons (PAHs), dereplicated MAG ORFs were further
180 annotated against two homebrew databases: EDB-DHG and EDB-Ncyc for dechlorinating and
181 nitrogen removal, and the Darhd Database for PAH degradation (Liet al. 2022), as well as the
182 PlasticDB for plastics degradation (Gambariniet al. 2022), respectively, using Diamond with the same
183 parameter described above, finally reserve matched sequences at identity ≥ 50%. As the Darhd

database only contained nucleotide sequences, we pre-treated all sequences by translating into amino acids with Prodigal to keep the consistency with other databases. Annotated sequences were further checked by comparing results with KEGG, NCBI-NR, and InterProScan, respectively, to avoid false annotation. EDB-DHG and EDB-Ncyc Databases are available for download in FigShare (https://figshare.com/ndownloader/articles/23504874/versions/1).

2.7. Statistical analysis and data visualization

Data organization and formatting were conducted using the R packages "tibble", "dplyr", "tidyr", and "stringr". Statistical analysis was conducted using the R packages "vegan", "stats", and "FSA". Data visualization was performed using the R packages "ggplot2" and "cowplot".

3. Results

3.1. Large proportion of metagenome-assembled genomes from hypersaline lakes were considered novel

Metagenomic sequencing generated ~1.2 Tb DNA sequence data from 30 samples taken from four hypersaline lakes. The metagenomic assembly generated 26 Gb of contig data with sequence length > 1,000 bp. After metagenomic binning, we recovered 3,030 non-redundant bins that meet the MIMAG stand (Completeness ≥ 50%, Purity ≥ 90%, mean Completeness = 79.8% ± 14.8%, mean Purity = 97.8% ± 2.0%) (Bowerset al. 2017). These bins were considered as metagenomic-assembled genomes (MAGs) for the following analyses (Fig. 1A). There were 76.9%, 51.5%, and 44.6% of the MAGs contained 5S, 16S, and 23S rRNA genes, respectively, while 23.7% possessed all three types of rRNA genes (Fig. 1B). In terms of tRNA genes, 75% of the MAGs contained more than 13 types of tRNA genes (Fig. 1C). There were 97.1%, 78.0% and 56.7% of MAGs affiliated with unclassified or close to a tentatively assigned species, genera, and families, respectively, suggesting that a high proportion of the genomes discovered in the hypersaline lakes were unknown or candidates for future identification (Fig. 1D). SDetails of the sequencing data size, contig summary, number of genomes, and specifics of the MAGs were supplied in Table S1.

3.2. Distinct genetic composition and functional potential among lakes

Distinct genetic compositions were identified among the lakes and sample types. Overall, the planktonic genetic richness in Chao1 index of ADH and DBC were significantly higher than that of QJJ, while the genetic richness and evenness in Shannon index of DBC was significantly higher than that of QJJ ($P < 0.05$). However, no significant difference was found between lakes in sedimentary samples (Fig. 2A). Genetic compositions of lake planktonic and sedimentary communities were significantly different ($P < 0.05$, Fig. S2), which were mainly driven by salinity including Li^+, Na^+,

217 SO_4^{2-}, Ca^{2+}, Mg^{2+}, and K^+ in planktonic communities, and TOC, SO_4^{2-}, and Na^+ in sedimentary
218 communities (Fig. 2B). Functional capacity profiles suggested that microbial genetic potentials varied
219 distinctly among sample types and lakes, particularly in aspects such as nitrogen and sulfur
220 metabolisms, carbon fixation, methanogenesis, and multidrug resistance, among others (Fig. 3C).
221 Specific analysis within pathways showed that in carbon fixation, Dicarboxylate−hydroxybutyrate
222 cycle was mainly enriched in QJJ planktons, while rTCA cycle, Hydroxypropionate−hydroxybutylate
223 cycle, and 3−Hydroxypropionate bi−cycle were enriched in QJJ sediments. On the other hand, BLK
224 sediment communities mainly performed Reductive pentose phosphate cycle (CBB cycle) and
225 Wood−Ljungdahl pathway for carbon fixation. For methane metabolism, BLK was found to be
226 enriched in Methanogenisis in sediments, while DBC and QJJ were found to be enriched in methane
227 oxidation in sediments. F420 biosynthesis were found to be abundant in QJJ plankton, possibly due to
228 the high relative abundance of archaea communities. Nitrogen cycle related modules including
229 nitrogen fixation, nitrification, and denitrification were mainly found enriched in QJJ sediments, while
230 dissimilatory nitrate reduction and assimilatory nitrate reduction module were found enriched in DBC
231 and QJJ plankton, respectively. For sulfur metabolism, thiosulfate oxidation was found to be abundant
232 in DBC plankton, while dissimilatory sulfate reduction was found enriched in BLK sediments (Fig.
233 2D). These findings suggest that distinct metabolic pathways were found among these four lakes.

234 **3.3. Microbial dark matters were highly diverse in the hypersaline lakes**

235 Of the 3,030 non-redundant MAGs, 2,685 were classified as bacteria spanning 70 phyla, with 18 of
236 them (25.7%) considered as MDM phyla. The remaining 345 MAGs were identified as archaea,
237 spanning 12 phyla, including 8 MDM phyla. While a vast majority of the MAGs were assigned to non-
238 MDM phyla (96.6% in plankton and 92.5% in sediment in relative abundance), those minorities which
239 classified as MDM demonstrated higher diversity that comprised 16.7% and 30.1% of total MAGs in
240 numbers in the plankton and sediments, respectively (Fig. 2A). This observation suggested that MDMs
241 could exhibit considerable diversity in hypersaline environments. Lineages of common phyla were
242 ubiquitously found across most samples, but distinct distributions of MDM phyla (except
243 *Patescibacteria*) were only identified in specific lakes or sample types (Fig. 2B). These MDM phyla
244 were predominantly located in samples with high microbial diversity (*i.e.,* ADH and BLK sediment,
245 Fig. 2B & C), but low relative abundance (< 0.5%) among recovered MAGs, indicating that a
246 considerable number of unknown microorganisms inhabitant in these environments were yet to be
247 discovered.

248 **3.4. Biosynthetic Potentials of saline lake MDMs**

249 In the 3,030 hypersaline lake MAGs, we identified 9,635 biosynthetic gene clusters (BGCs), including
250 1,170 BGCs clustered into known gene cluster families (GCFs) as per the MiBIG database. A total of

251 8,966 BGCs (93% of overall BGCs) that including 501 clustered BGCs with cosine distance ≥ 0.2, as
252 well as 8,465 unclustered BGCs, were considered novel (Fig. 4A). Notably, a number of bacterial
253 MDM phyla were enriched in BGCs (BGCs per MAG > 5), including *OLB16*, *Moduliflexota*, *KSB1*,
254 *Fibrobacterota*, and *B130-G9*. This enrichment was only observed in *Planctomycetota*, *Myxococcota*,
255 and *Acidobacteriota* among the non-MDM phyla (Fig. 4B), suggesting that bacterial MDMs in these
256 saline lakes may possess abundant biosynthetic potentials. Fewer BGCs were found in archaeal MAGs
257 (BGCs per MAG < 2), possibly due to the insufficient data collected in the database.

258 To further explore biosynthetic potentials of MDM MAGs, bacterial and archaeal phylogenetic trees
259 were constructed, selecting MAGs with Completeness ≥ 70% or BGCs ≥ 5. There were 16 bacterial
260 and 2 archaeal phyla that contained MAG(s) with BGCs ≥ 5, suggesting potential biosynthetic
261 enrichment in these phyla. Bacterial and archaeal phyla with small genome sizes (*i.e., Patescibacteria*
262 and *Nanoarchaeota*) generally lacked BGCs, except one MAG from *Nanoarchaeota* which was
263 considered an outlier. Some phyla with only a single MAG obtained from saline lake metagenomes,
264 such as *B130-G9*, *Moduliflexota*, *UBP6*, etc. (labelled individually on the tree, Figure 4B), were not
265 further discussed in this study as their biosynthetic pattern and consistency cannot be validated due to
266 insufficient data. However, other MDM phyla consistently demonstrated enrichment with specific
267 BGC Classes. For example, phylum *Armatimonadota* primarily consisted of RiPPs, Terpene and Other
268 BGC classes, while phylum OLB16 was enriched in NRPS, RiPPs, and Terpene. Most of the MAGs
269 in phylum *Bdellovibrionota* had BGCs in Terpene, RiPPs, and Other BGC classes, whereas KSB1
270 MAGs featured NRPS, PKS-NRPS hybrid, and Terpene classes (Figure 4C). These observations
271 suggest that MAGs within these phyla may have similar biosynthetic potentials. Two *Asgardarchaeota*
272 MAGs were found to be enriched with BGCs, although this pattern was not consistent across the entire
273 phyla, necessitating more genomes analysis for further exploration.

274 **3.5. Biodegradation potentials of saline lake MDMs**

275 To dissect the biodegradation potentials of hypersaline lake microorganisms, we analysed genes
276 involved in the nitrogen cycle, dehalogenation, plastics degradation, and PAHs degradation within
277 MDM MAGs. A number of bacterial MDM phyla were found to be rich in nitrogen metabolism genes,
278 such as *GCA-001730085*, *Hydrogenedentota*, *OLB16*, *Fibrobacterota*, *Krumholzibacteriota*, *FEN-*
279 *1099*, *RBG-13-61-14*, and *Zixibacteria* (Figure 5A), suggesting that these MDMs may play a crucial
280 role in the nitrogen cycle in hypersaline lakes. Interestingly, although MAGs belong to typical
281 anaerobic ammonium oxidizing bacteria (AnAOB), *i.e.,* families *Scalinduaceae* and *Brocadiaceae* in
282 the order *Brocadiales*, class *Brocadiae*, were not found in this study, anaerobic ammonium oxidation
283 (Anammox)-related genes (*i.e.,* gen *hzs* and *hdh*) were found in several MDM phyla in addition to the
284 known AnAOB groups, such as GCA-001730085, *Hydrogenedentota*, *Calditrichota*, *KSB1*, *FEN-*

285 *1099*, and *SAR324* (Figure 5A). While fewer nitrogen cycle related genes were found in archaea than
286 bacteria, we found that Asgard archaea universally contain *nirBD* and *norBC* genes, suggesting that
287 they might be functionally involved in the nitrogen cycle in hypersaline lakes (Figure 5B). In terms of
288 the dehalogenation genes *rdh*, we found a number of MDM lineages with *rdh* genes in the hypersaline
289 lake MAGs, such as bacterial phyla *Eisenbacteria*, *Krumholzibacteriota*, *Zixibacteria*, *KSB1*, and
290 *Marinisomatota* (Figure 5A). In archaeal taxa, *rdh* gene were found in phyla *Thorarchaeota*,
291 *Lokiarchaeota*, and *Hadarchaeota* (Figure 5B). These findings suggested that these MDMs may have
292 dehalogenation capacity in hypersaline lakes.

293 To explore plastic degradation and PAHs potentials of saline lake MAGs, dereplicated ORFs were
294 annotated against the PlasticDB and Darhd databases. Results showed that MDM MAGs with plastic
295 degradation capabilities distributed in most of the phyla, including for previously known as non-
296 biodegradable plastics, such as polyethylene (PE), polyethylene terephthalate (PET), polyamide
297 (Nylon), polystyrene (PS), and polyurethane (PU), were found in phyla *GCA-001730085*,
298 *Hydrogenedentota*, *Krumholzibacteriota*, *Fibrobacterota*, *Marinisomatota*, *KSB1*, *Moduliflexota*,
299 *FEN-1099*, *SAR324*, *RBG-13-61-14*, and *Lokiarchaeota* (Figure 5A & B). For PAHs degradation
300 potentials, genes such as aromatic ring-hydroxylating dioxygenase, phenylpropionate dioxygenase,
301 and benzoate dioxygenase were found across MDM MAGs, spanning phyla *Sumerlaeota*,
302 *Krumholzibacteriota*, *Calditrichota*, and *Lokiarchaeota* (Figure 5A & B), suggesting that many
303 MDMs may also have the potential to degrade microbially degradable plastics and PAHs.

304

305 # 4. Discussion

306 Uncharacterised microorganisms possess a vast repository of undiscovered enzymes and metabolic
307 pathways that hold the potential for applications in various fields including biotechnology, medicine,
308 agriculture, and environmental remediation (Atanasovet al. 2021; Bhushanet al. 2019; Chenet al. 2021;
309 Rinkeet al. 2013; Waschulinet al. 2022; Xuet al. 2022). By identifying and characterising both known
310 or novel enzymes and pathways in MDMs, we can uncover new technologies for the production of
311 compounds and natural products, as well as innovative strategies for biodegradation of pollutants and
312 toxic wastes. In this study, we investigated four hypersaline lakes with distinct biochemical
313 characteristics and successfully uncovered over 3,000 MAGs with biosynthesis and biodegradation
314 potentials. Although a metatranscriptomic analysis was not included in this study to confirm
315 biosynthetic activity, our metagenomic analyses revealed nearly 10,000 potential BGCs, including ~80%
316 of which were considered novel. Notably, our findings concerning MDM phyla that are enriched in
317 BGCs aligned with a previous studies that included metatranscriptomic analysis from marine
318 environments (Geller-McGrathet al. 2023; Paoliet al. 2022), such as *Omnitrophota*, *KSB1*, *OLB16*, and

319 *FEN-1099*, etc., further validated our findings in biosynthetic analysis. On the other hand,
320 biodegradation capabilities such as nitrogen removals (especially Anammox), dehalogenation, plastics
321 degradation, and PAHs degradation, are areas of intense interest for environmental remediation efforts
322 (Heet al. 2021; Liuet al. 2023a; Liuet al. 2020; Qiaoet al. 2023; Zhanget al. 2019; Zrimecet al. 2021).
323 The presence of biodegradation genes in MDMs observed in this study offered new potentials for the
324 degradation of pollutants, plastics, and toxic compounds. Overall, these findings improved our
325 understanding of the metabolic potentials of novel microbial species in developing biosynthetic and
326 biodegradation capabilities, which can contribute to a broad range of applications including biofuels,
327 medicine, agricultural and industrial applications, as well as environmental remediation efforts.
328 Moreover, the recovery of MAGs provided valuable genomic resources that enables insights into
329 microbial evolution, potentially expanding our understanding of the Tree of Life (Huget al. 2016).

330 Extreme environments are typically inhospitable to a large proportion of known microorganisms (Lee
331 and Park 2018; Rothschild and Mancinelli 2001). However, such environments offer a group of
332 specialised extremophiles that have evolved unique strategies to maintain their stability and
333 functionality under such extreme conditions (Malavasiet al. 2020; Wanget al. 2022; Yinet al. 2019).
334 Recent studies have uncovered dozens of thousands of MAGs from various extreme environments,
335 including deserts, oil fields, hydrothermal vents, cold seeps, hypersaline lakes, and glaciers, etc.
336 (Andersonet al. 2017; Donget al. 2022; Ezeet al. 2020; Finstadet al. 2017; Liuet al. 2019; Liuet al.
337 2022; Vavourakiset al. 2018). These findings vastly improved our recognition of uncultured
338 microorganisms, including their genetic and functional potentials in these extreme environments.
339 Microorganisms in hypersaline lakes are known as halotolerant or halophilic live forms, with distinct
340 physiological and biochemical features that enable them to adapt such extreme environments (Liuet al.
341 2023b; Muet al. 2020). Interestingly, while hypersaline lakes were previously known to be nutrient-
342 poor, recent studies revealed that some of them are nutrient-rich, particularly in the lake sediment
343 (Konget al. 2020; Yueet al. 2019; Zhaoet al. 2022), which aligns with the results of this study. The
344 remarkable adaptations and compatibilities of these organisms to high salinity environment, along with
345 specialised functions including osmoregulation, "salt-in" strategy, membrane formation, and
346 halophilic enzymes, etc. (Gunde-Cimermanet al. 2018; Oren 2002; 2020; Rodriguez-Valera 2020;
347 Salma and andMehnaz Samina 2020), offer exceptional opportunities for biotechnological applications.
348 These applications range from non-sterile bioproduction and bioremediation of high salinity wastes,
349 to salt production and desalination. Collectively, mining and utilising microbial resources from
350 hypersaline lakes can significantly advance our understanding of the metabolic potentials of
351 extremophiles. This knowledge can be harnessed for applications in biotechnological industries and
352 environmental management practices, broadening the spectrum of practical applications and
353 contributing to sustainable development and environmental preservation.

Despite the promising potential of exploiting uncultured hypersaline microorganisms (namely MDMs) for various applications, limitations of exploiting these MDMs cannot be overlooked. One key challenge is that MDMs often contain many genes or enzymes undocumented in the current databases. This can affect the accuracy of predictions, annotations, or classification in the genomic analyses (Dröge and McHardy 2012; Lobbet al. 2020; Rinkeet al. 2013). While advancements in state-of-art sequencing technologies and bioinformatic tools, such as long-read sequencing (Jainet al. 2016; Wanget al. 2015), single-cell sequencing (Nawy 2014), deep refinement (Yuet al. 2021), and machine learning (Hoarfrostet al. 2022) approaches, have assisted to mitigate these discrepancies, each of these technologies has its own limitations. These still hampers the robust *de novo* assembly of low abundant MDM MAGs from the microbial communities. Moreover, since successfully culturing of a microorganism is still considered as the "gold standard" approach of the microbial characterisation, it is still challenging to replicate the extreme conditions such as hypersaline environments under laboratory conditions, making it difficult to culture MDMs. Even if the culturing approaches can be managed, validating their functional and metabolic properties to confirm biosynthesis and biodegradation potentials remains a major challenge. Cutting-edge technologies and in-depth analysis tools, such as multi-omics approaches (Gutlebenet al. 2018), Raman spectroscopy methods (Cuiet al. 2022), and high-throughput single-cell sequencing technologies (Zhenget al. 2022) may assist the isolation and culturing of uncharacterised microorganisms. However, most of these techniques have been sophisticatedly applied to human gut microbiome studies, thus such application to extreme environmental samples remains challenging, both in terms of laboratory experiments and *in silico* analyses.

Indeed, further exploration of MDMs from hypersaline lakes and other extreme environments is complex, with numerous technical challenges to be addressed. One of the key conundrums lies in the increase of resolution for low coverage of microbial genomes, which is difficult to recover via binning. The successful enrichment of these MDMs is also hindered by intricate nutrient requirements, complex microbial interactions, and slow growth rates associated with these extremophiles (Dedysh 2011; Imachiet al. 2020; Lewiset al. 2021). Possible solutions might be conceived such as 1) utilisation of advanced sequencing technologies such as Hi-C, Pore-C, and HiFi sequencing (Bickhartet al. 2022; Deshpandeet al. 2022; Wengeret al. 2019), which could potentially yield higher-quality and more complete genomic assemblies from metagenomic datasets, offering a more detailed and accurate view of the landscape of MDM genomes; 2) development of new bioinformatic tools to enhance the accuracy and resolution of recovered genomes from metagenomic datasets. This could be achieved through the application of machine learning algorithms or the improvement of existing bioinformatic tools; 3) upgraded laboratory techniques for culture enrichment (*e.g.,* machine learning based high-throughput microbial culturomics) (Huanget al. 2023) or "counterselection" strategy on extreme

389 environmental metagenomes. Nevertheless, successful development and application of these
390 approaches could significantly advance our discovery of MDMs and their associated biotechnological
391 potentials. It could also unfold possibilities for exploration of life in most extreme environments on
392 Earth, including abyssal, hadal, deep subsurface, and even beyond, into extraterrestrial environments.

393

394 ## 5. Conclusions

395 Our study implemented genome-centric analyses on over 3,000 MAGs recovered from four
396 hypersaline lakes with distinct different environmental parameters in Xinjiang, China, offered valuable
397 resources of genomic data and contributions to our understanding of microbial life in extreme
398 environments, particularly hypersaline lakes. We found over 8,000 potential BGCs and uncovered
399 several MDM phyla that may be enriched in biosynthetic capacities, which could substantially support
400 future research in biotechnological applications. In addition, the discovery of biodegradation genes
401 within certain MDM lineages also suggests promising prospects for environmental remediation
402 strategies. In summary, this study expands our knowledge of microbial diversity and function in
403 extreme environments, paving the path to the future discovery of uncultured microorganisms. This
404 research not only deepens our understanding of the Tree of Life but also opens doors to a multitude of
405 potential applications in biotechnology and environmental remediation.

406

407 ## CRediT authorship contribution statement

408 ZQ and KY designed the experiment; XQ and YK collected the samples; RM and ZQ processed the
409 samples; ZQ, YZ, and QZ analysed the data; ZQ wrote the manuscript; ZQ, YK and TZ revised the
410 manuscript; and all co-authors reviewed and agreed the manuscript for publication.

411 ## Declaration of competing interests

412 The authors have no competing interests.

413 ## Acknowledgements

414 The authors would like to thank Dr. Luhua Xie, Dr. Bing Li, Ms. Liyu Zhang, and Ms. Xiaoling Gao
415 for technical assistance and valuable discussion.

416 ## Funding

417 This research was supported by the National Key Research and Development Program of China
418 (2021YFA1301300), Nature Science Foundation of China (62202014 and 61972217), Science and

Technology Planning Project of Shenzhen Municipality (JCYJ20200109120416654), and Shenzhen Basic Research Program (JCYJ20220813151736001).

Availability of data and materials

Metagenomic sequence data of saline lakes are available in the CNCB-NGDC database under project PRJCA014712.

Consent for publication

The authors have all reviewed and agreed the manuscript for publication.

References

Albertsen, M.; Hugenholtz, P.; Skarshewski, A.; Nielsen, K.L.; Tyson, G.W.; Nielsen, P.H. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat Biotechnol 2013;31:533-538

Anderson, R.E.; Reveillaud, J.; Reddington, E.; Delmont, T.O.; Eren, A.M.; McDermott, J.M.; Seewald, J.S.; Huber, J.A. Genomic variation in microbial populations inhabiting the marine subseafloor at deep-sea hydrothermal vents. Nature communications 2017;8:1-11

Atanasov, A.G.; Zotchev, S.B.; Dirsch, V.M.; Supuran, C.T. Natural products in drug discovery: advances and opportunities. Nature reviews Drug discovery 2021;20:200-216

Azubuike, C.C.; Chikere, C.B.; Okpokwasili, G.C. Bioremediation techniques–classification based on site of application: principles, advantages, limitations and prospects. World Journal of Microbiology and Biotechnology 2016;32:1-18

Bahram, M.; Hildebrand, F.; Forslund, S.K.; Anderson, J.L.; Soudzilovskaia, N.A.; Bodegom, P.M.; Bengtsson-Palme, J.; Anslan, S.; Coelho, L.P.; Harend, H. Structure and function of the global topsoil microbiome. Nature 2018;560:233-237

Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012;19:455-477

Barabadi, H.; Tajani, B.; Moradi, M.; Damavandi Kamali, K.; Meena, R.; Honary, S.; Mahjoub, M.A.; Saravanan, M. Penicillium family as emerging nanofactory for biosynthesis of green nanomaterials: a journey into the world of microorganisms. Journal of Cluster Science 2019;30:843-856

Bhushan, A.; Egli, P.J.; Peters, E.E.; Freeman, M.F.; Piel, J. Genome mining-and synthetic biology-enabled production of hypermodified peptides. Nature chemistry 2019;11:931-939

Bickhart, D.M.; Kolmogorov, M.; Tseng, E.; Portik, D.M.; Korobeynikov, A.; Tolstoganov, I.; Uritskiy, G.; Liachko, I.; Sullivan, S.T.; Shin, S.B.; Zorea, A.; Andreu, V.P.; Panke-Buisse, K.; Medema, M.H.; Mizrahi, I.; Pevzner, P.A.; Smith, T.P.L. Generating lineage-resolved, complete metagenome-assembled genomes from complex microbial communities. Nat Biotechnol 2022;

Blin, K.; Shaw, S.; Kloosterman, A.M.; Charlop-Powers, Z.; Van Wezel, G.P.; Medema, M.H.; Weber, T. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 2021;49:W29-W35

Bowers, R.M.; Kyrpides, N.C.; Stepanauskas, R.; Harmon-Smith, M.; Doud, D.; Reddy, T.; Schulz, F.; Jarett, J.; Rivers, A.R.; Eloe-Fadrosh, E.A. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol 2017;35:725-731

Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. Nature Methods 2015;12:59

Chan, P.P.; Lin, B.Y.; Mak, A.J.; Lowe, T.M. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res 2021;49:9077-9096

Chaumeil, P.-A.; Mussig, A.J.; Hugenholtz, P.; Parks, D.H. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 2020;

Chemler, J.A.; Koffas, M.A. Metabolic engineering for plant natural product biosynthesis in microbes. Curr Opin Biotechnol 2008;19:597-605

Chen, S.-C.; Budhraja, R.; Adrian, L.; Calabrese, F.; Stryhanyuk, H.; Musat, N.; Richnow, H.-H.; Duan, G.-L.; Zhu, Y.-G.; Musat, F. Novel clades of soil biphenyl degraders revealed by integrating isotope probing, multi-omics, and single-cell analyses. The ISME Journal 2021;15:3508-3521

Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018;34:i884-i890

Cui, D.; Kong, L.; Wang, Y.; Zhu, Y.; Zhang, C. In situ identification of environmental microorganisms with Raman spectroscopy. Environmental Science and Ecotechnology 2022:100187

Dedysh, S.N. Cultivating uncultured bacteria from northern wetlands: knowledge gained and remaining gaps. Frontiers in microbiology 2011;2:184

Deshpande, A.S.; Ulahannan, N.; Pendleton, M.; Dai, X.; Ly, L.; Behr, J.M.; Schwenk, S.; Liao, W.; Augello, M.A.; Tyer, C. Identifying synergistic high-order 3D chromatin conformations from genome-scale nanopore concatemer sequencing. Nat Biotechnol 2022;40:1488-1499

Dong, X.; Zhang, C.; Peng, Y.; Zhang, H.-X.; Shi, L.-D.; Wei, G.; Hubert, C.R.; Wang, Y.; Greening, C. Phylogenetically and catabolically diverse diazotrophs reside in deep-sea cold seep sediments. Nature Communications 2022;13:4885

Douglas, G.M.; Maffei, V.J.; Zaneveld, J.R.; Yurgel, S.N.; Brown, J.R.; Taylor, C.M.; Huttenhower, C.; Langille, M.G.I. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol 2020;38:685-688

Dröge, J.; McHardy, A.C. Taxonomic binning of metagenome samples generated by next-generation sequencing technologies. Briefings in bioinformatics 2012;13:646-655

Eddy, S.R. Accelerated profile HMM searches. PLoS Comp Biol 2011;7:e1002195

Eze, M.O.; Lütgert, S.A.; Neubauer, H.; Balouri, A.; Kraft, A.A.; Sieven, A.; Daniel, R.; Wemheuer, B. Metagenome assembly and metagenome-assembled genome sequences from a historical oil field located in Wietze, Germany. Microbiology Resource Announcements 2020;9

Falkowski, P.G.; Fenchel, T.; Delong, E.F. The microbial engines that drive Earth's biogeochemical cycles. science 2008;320:1034-1039

Finstad, K.M.; Probst, A.J.; Thomas, B.C.; Andersen, G.L.; Demergasso, C.; Echeverría, A.; Amundson, R.G.; Banfield, J.F. Microbial community structure and the persistence of cyanobacterial populations in salt crusts of the hyperarid Atacama Desert from genome-resolved metagenomics. Frontiers in microbiology 2017;8:1435

Gambarini, V.; Pantos, O.; Kingsbury, J.M.; Weaver, L.; Handley, K.M.; Lear, G. PlasticDB: a database of microorganisms and proteins linked to plastic biodegradation. Database 2022;2022

Geller-McGrath, D.; Mara, P.; Taylor, G.T.; Suter, E.; Edgcomb, V.; Pachiadaki, M. Diverse secondary metabolites are expressed in particle-associated and free-living microorganisms of the permanently anoxic Cariaco Basin. Nature Communications 2023;14:656

Gunde-Cimerman, N.; Plemenitaš, A.; Oren, A. Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. FEMS Microbiol Rev 2018;42:353-375

Gutleben, J.; Chaib De Mares, M.; Van Elsas, J.D.; Smidt, H.; Overmann, J.; Sipkema, D. The multi-omics promise in context: from sequence to microbial isolate. Crit Rev Microbiol 2018;44:212-229

He, H.; Li, Y.; Shen, R.; Shim, H.; Zeng, Y.; Zhao, S.; Lu, Q.; Mai, B.; Wang, S. Environmental occurrence and remediation of emerging organohalides: A review. Environ Pollut 2021;290:118060

Hoarfrost, A.; Aptekmann, A.; Farfañuk, G.; Bromberg, Y. Deep learning of a bacterial and archaeal universal language of life enables transfer learning and illuminates microbial dark matter. Nature communications 2022;13:2606

Huang, Y.; Sheth, R.U.; Zhao, S.; Cohen, L.A.; Dabaghi, K.; Moody, T.; Sun, Y.; Ricaurte, D.; Richardson, M.; Velez-Cortes, F. High-throughput microbial culturomics using automation and machine learning. Nat Biotechnol 2023:1-10

Hug, L.A.; Baker, B.J.; Anantharaman, K.; Brown, C.T.; Probst, A.J.; Castelle, C.J.; Butterfield, C.N.; Hernsdorf, A.W.; Amano, Y.; Ise, K. A new view of the tree of life. Nature microbiology 2016;1:1-6

Hyatt, D.; Chen, G.-L.; LoCascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: prokaryotic gene recognition and translation initiation site identification. Bioinformatics 2010;11:119

Imachi, H.; Nobu, M.K.; Nakahara, N.; Morono, Y.; Ogawara, M.; Takaki, Y.; Takano, Y.; Uematsu, K.; Ikuta, T.; Ito, M. Isolation of an archaeon at the prokaryote–eukaryote interface. Nature 2020;577:519-525

Jain, M.; Olsen, H.E.; Paten, B.; Akeson, M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome biology 2016;17:1-11

523 Ji, M.; Kong, W.; Yue, L.; Wang, J.; Deng, Y.; Zhu, L. Salinity reduces bacterial diversity, but increases network
524 complexity in Tibetan Plateau lakes. FEMS Microbiol Ecol 2019;95:fiz190
525 Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.;
526 Nuka, G. InterProScan 5: genome-scale protein function classification. Bioinformatics 2014;30:1236-
527 1240
528 Kautsar, S.A.; Blin, K.; Shaw, S.; Navarro-Muñoz, J.C.; Terlouw, B.R.; Van Der Hooft, J.J.; Van Santen, J.A.;
529 Tracanna, V.; Suarez Duran, H.G.; Pascal Andreu, V. MIBiG 2.0: a repository for biosynthetic gene
530 clusters of known function. Nucleic Acids Res 2020;48:D454-D458
531 Kong, X.; Jiang, Z.; Zheng, Y.; Xiao, M.; Chen, C.; Yuan, H.; Chen, F.; Wu, S.; Zhang, J.; Han, C. Organic
532 geochemical characteristics and organic matter enrichment of mudstones in an Eocene saline lake,
533 Qianjiang Depression, Hubei Province, China. Mar Pet Geol 2020;114:104194
534 Krienitz, L.; Kotut, K. Fluctuating algal food populations and the occurrence of lesser flamingos (Phoeniconaias
535 minor) in three Kenyan rift valley lakes 1. J Phycol 2010;46:1088-1096
536 Langdon, W.B. Performance of genetic programming optimised Bowtie2 on genome comparison and analytic
537 testing (GCAT) benchmarks. BioData mining 2015;8:1-7
538 Lee, Y.S.; Park, W. Current challenges and future directions for bacterial self-healing concrete. Appl Microbiol
539 Biotechnol 2018;102:3059-3070
540 Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and
541 annotation. Nucleic Acids Res 2021;49:W293-W296
542 Lewis, W.H.; Tahon, G.; Geesink, P.; Sousa, D.Z.; Ettema, T.J. Innovations to culturing the uncultured
543 microbial majority. Nature Reviews Microbiology 2021;19:225-240
544 Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The
545 sequence alignment/map format and SAMtools. Bioinformatics 2009;25
546 Li, S.; Shen, W.; Lian, S.; Wu, Y.; Qu, Y.; Deng, Y. Darhd: A sequence database for aromatic ring-
547 hydroxylating dioxygenase analysis and primer evaluation. J Hazard Mater 2022;436:129230
548 Li, W.; Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide
549 sequences. Bioinformatics 2006;22
550 Liu, J.; Xu, G.; Zhao, S.; Chen, C.; Rogers, M.J.; He, J. Mechanistic and microbial ecological insights into the
551 impacts of micro-and nano-plastics on microbial reductive dehalogenation of organohalide pollutants.
552 J Hazard Mater 2023a;448:130895
553 Liu, J.; Zheng, Y.; Lin, H.; Wang, X.; Li, M.; Liu, Y.; Yu, M.; Zhao, M.; Pedentchouk, N.; Lea-Smith, D.J.
554 Proliferation of hydrocarbon-degrading microbes at the bottom of the Mariana Trench. Microbiome
555 2019;7:1-13
556 Liu, L.; Wang, Y.; Che, Y.; Chen, Y.; Xia, Y.; Luo, R.; Cheng, S.H.; Zheng, C.; Zhang, T. High-quality bacterial
557 genomes of a partial-nitritation/anammox system by an iterative hybrid assembly method. Microbiome
558 2020;8:1-17
559 Liu, Y.-H.; Mohamad, O.A.A.; Gao, L.; Xie, Y.-G.; Abdugheni, R.; Huang, Y.; Li, L.; Fang, B.-Z.; Li, W.-J.
560 Sediment prokaryotic microbial community and potential biogeochemical cycle from saline lakes
561 shaped by habitat. Microbiol Res 2023b;270:127342
562 Liu, Y.; Ji, M.; Yu, T.; Zaugg, J.; Anesio, A.M.; Zhang, Z.; Hu, S.; Hugenholtz, P.; Liu, K.; Liu, P. A genome
563 and gene catalog of glacier microbiomes. Nat Biotechnol 2022;40:1341-1348
564 Lobb, B.; Tremblay, B.J.-M.; Moreno-Hagelsieb, G.; Doxey, A.C. An assessment of genome annotation
565 coverage across the bacterial tree of life. Microbial Genomics 2020;6
566 Louca, S.; Parfrey, L.W.; Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome.
567 Science 2016;353:1272-1277
568 Malavasi, V.; Soru, S.; Cao, G. Extremophile microalgae: the potential for biotechnological application. J
569 Phycol 2020;56:559-573
570 Mesbah, N.M.; Wiegel, J. Life under multiple extreme conditions: diversity and physiology of the halophilic
571 alkalithermophiles. Appl Environ Microbiol 2012;78:4074-4082
572 Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; Von Haeseler, A.; Lanfear, R.
573 IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol
574 Evol 2020;37:1530-1534
575 Mu, D.-S.; Wang, S.; Liang, Q.-Y.; Du, Z.-Z.; Tian, R.; Ouyang, Y.; Wang, X.-P.; Zhou, A.; Gong, Y.; Chen,
576 G.-J. Bradymonabacteria, a novel bacterial predator group with versatile survival strategies in saline
577 environments. Microbiome 2020;8:1-15
578 Navarro-Muñoz, J.C.; Selem-Mojica, N.; Mullowney, M.W.; Kautsar, S.A.; Tryon, J.H.; Parkinson, E.I.; De
579 Los Santos, E.L.; Yeong, M.; Cruz-Morales, P.; Abubucker, S. A computational framework to explore
580 large-scale biosynthetic diversity. Nat Chem Biol 2020;16:60-68

581 Nawy, T. Single-cell sequencing. Nat Methods 2014;11:18-18
582 Oren, A. Molecular ecology of extremely halophilic Archaea and Bacteria. FEMS Microbiol Ecol 2002;39:1-7
583 Oren, A. Ecology of extremely halophilic microorganisms. The biology of halophilic bacteria: CRC Press; 2020
584 Paoli, L.; Ruscheweyh, H.-J.; Forneris, C.C.; Hubrich, F.; Kautsar, S.; Bhushan, A.; Lotti, A.; Clayssen, Q.;
585 Salazar, G.; Milanese, A. Biosynthetic potential of the global ocean microbiome. Nature 2022;607:111-
586 118
587 Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: assessing the quality of
588 microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 2015;25:1043-
589 1055
590 Qiao, X.; Fu, C.; Chen, Y.; Fang, F.; Zhang, Y.; Ding, L.; Yang, K.; Pan, B.; Xu, N.; Yu, K. Molecular insights
591 into enhanced nitrogen removal induced by trace fluoroquinolone antibiotics in an anammox system.
592 Bioresour Technol 2023;374:128784
593 Rinke, C.; Schwientek, P.; Sczyrba, A.; Ivanova, N.N.; Anderson, I.J.; Cheng, J.-F.; Darling, A.; Malfatti, S.;
594 Swan, B.K.; Gies, E.A. Insights into the phylogeny and coding potential of microbial dark matter.
595 Nature 2013;499:431-437
596 Rodriguez-Valera, F. Introduction to saline environments. The biology of halophilic bacteria: CRC Press; 2020
597 Rojas, P.; Rodríguez, N.; de la Fuente, V.; Sánchez-Mata, D.; Amils, R.; Sanz, J.L. Microbial diversity
598 associated with the anaerobic sediments of a soda lake (Mono Lake, California, USA). Canadian journal
599 of microbiology 2018;64:385-392
600 Rothschild, L.J.; Mancinelli, R.L. Life in extreme environments. Nature 2001;409:1092-1101
601 Salma, M.; andMehnaz Samina, M.K.A. Osmoadaptation in halophilic bacteria and archaea. Research Journal
602 of Biotechnology Vol 2020;15:5
603 Saravanan, A.; Kumar, P.S.; Karishma, S.; Vo, D.-V.N.; Jeevanantham, S.; Yaashikaa, P.; George, C.S. A
604 review on biosynthesis of metal nanoparticles and its environmental applications. Chemosphere
605 2021;264:128580
606 Schultz, J.; Rosado, A.S. Extreme environments: a source of biosurfactants for biotechnological applications.
607 Extremophiles 2020;24:189-206
608 Sender, R.; Fuchs, S.; Milo, R. Revised estimates for the number of human and bacteria cells in the body. PLoS
609 Biol 2016;14:e1002533
610 Shu, W.-S.; Huang, L.-N. Microbial diversity in extreme environments. Nature Reviews Microbiology
611 2022;20:219-235
612 Sorokin, D.Y.; Berben, T.; Melton, E.D.; Overmars, L.; Vavourakis, C.D.; Muyzer, G. Microbial diversity and
613 biogeochemical cycling in soda lakes. Extremophiles 2014;18:791-809
614 Taghavi, N.; Udugama, I.A.; Zhuang, W.-Q.; Baroutian, S. Challenges in biodegradation of non-degradable
615 thermoplastic waste: From environmental impact to operational readiness. Biotechnol Adv
616 2021;49:107731
617 Vavourakis, C.D.; Andrei, A.-S.; Mehrshad, M.; Ghai, R.; Sorokin, D.Y.; Muyzer, G. A metagenomics roadmap
618 to the uncultured genome diversity in hypersaline soda lake sediments. Microbiome 2018;6:1-18
619 Ventosa, A.; de la Haba, R.R.; Sanchez-Porro, C.; Papke, R.T. Microbial diversity of hypersaline environments:
620 a metagenomic approach. Curr Opin Microbiol 2015;25:80-87
621 Wang, M.; Beck, C.R.; English, A.C.; Meng, Q.; Buhay, C.; Han, Y.; Doddapaneni, H.V.; Yu, F.; Boerwinkle,
622 E.; Lupski, J.R. PacBio-LITS: a large-insert targeted sequencing method for characterization of human
623 disease-associated chromosomal structural variations. BMC Genomics 2015;16:1-12
624 Wang, S.; Chng, K.R.; Wilm, A.; Zhao, S.; Yang, K.-L.; Nagarajan, N.; He, J. Genomic characterization of
625 three unique Dehalococcoides that respire on persistent polychlorinated biphenyls. Proceedings of the
626 National Academy of Sciences 2014;111:12103-12108
627 Wang, S.; Wang, J.; Liu, Z.; Zhang, B. Unraveling diverse survival strategies of microorganisms to vanadium
628 stress in aquatic environments. Water Res 2022;221:118813
629 Waschulin, V.; Borsetto, C.; James, R.; Newsham, K.K.; Donadio, S.; Corre, C.; Wellington, E. Biosynthetic
630 potential of uncultured Antarctic soil bacteria revealed through long-read metagenomic sequencing.
631 The ISME journal 2022;16:101-111
632 Wenger, A.M.; Peluso, P.; Rowell, W.J.; Chang, P.-C.; Hall, R.J.; Concepcion, G.T.; Ebler, J.; Fungtammasan,
633 A.; Kolesnikov, A.; Olson, N.D.; Töpfer, A.; Alonge, M.; Mahmoud, M.; Qian, Y.; Chin, C.-S.;
634 Phillippy, A.M.; Schatz, M.C.; Myers, G.; DePristo, M.A.; Ruan, J.; Marschall, T.; Sedlazeck, F.J.;
635 Zook, J.M.; Li, H.; Koren, S.; Carroll, A.; Rank, D.R.; Hunkapiller, M.W. Accurate circular consensus
636 long-read sequencing improves variant detection and assembly of a human genome. Nat Biotechnol
637 2019;37:1155-1162

638 Wu, J.; Du, G.; Zhou, J.; Chen, J. Systems metabolic engineering of microorganisms to achieve large-scale
639 production of flavonoid scaffolds. J Biotechnol 2014;188:72-80
640 Xie, Y.-G.; Luo, Z.-H.; Fang, B.-Z.; Jiao, J.-Y.; Xie, Q.-J.; Cao, X.-R.; Qu, Y.-N.; Qi, Y.-L.; Rao, Y.-Z.; Li, Y.-
641 X. Functional differentiation determines the molecular basis of the symbiotic lifestyle of Ca.
642 Nanohaloarchaeota. Microbiome 2022;10:1-13
643 Xu, Z.; Park, T.-J.; Cao, H. Advances in mining and expressing microbial biosynthetic gene clusters. Crit Rev
644 Microbiol 2022:1-20
645 Yin, W.; Wang, Y.; Liu, L.; He, J. Biofilms: the microbial "protective clothing" in extreme environments.
646 International journal of molecular sciences 2019;20:3423
647 Yu, K.; Qiu, Z.; Mu, R.; Qiao, X.; Zhang, L.; Lian, C.-A.; Deng, C.; Wu, Y.; Xu, Z.; Li, B.; Pan, B.; Zhang, Y.;
648 Fan, L.; Liu, Y.-x.; Cao, H.; Jin, T.; Chen, B.; Wang, F.; Yan, Y.; Xie, L.; Zhou, L.; Yi, S.; Chi, S.;
649 Zhang, C.; Zhang, T.; Zhuang, W. Recovery of high-qualitied Genomes from a deep-inland Salt Lake
650 Using BASALT. bioRxiv 2021:2021.2003.2005.434042
651 Yue, L.; Kong, W.; Ji, M.; Liu, J.; Morgan-Kiss, R.M. Community response of microbial primary producers to
652 salinity is primarily driven by nutrients in lakes. Sci Total Environ 2019;696:134001
653 Zallot, R.; Oberg, N.; Gerlt, J.A. Discovery of new enzymatic functions and metabolic pathways using genomic
654 enzymology web tools. Curr Opin Biotechnol 2021;69:77-90
655 Zhang, J.; Peng, Y.; Li, X.; Du, R. Feasibility of partial-denitrification/anammox for pharmaceutical wastewater
656 treatment in a hybrid biofilm reactor. Water Res 2022;208:117856
657 Zhang, S.; Hu, Z.; Wang, H. Metagenomic analysis exhibited the co-metabolism of polycyclic aromatic
658 hydrocarbons by bacterial community from estuarine sediment. Environ Int 2019;129:308-319
659 Zhao, D.; Zhang, S.; Xue, Q.; Chen, J.; Zhou, J.; Cheng, F.; Li, M.; Zhu, Y.; Yu, H.; Hu, S. Abundant taxa and
660 favorable pathways in the microbiome of soda-saline lakes in Inner Mongolia. Frontiers in microbiology
661 2020;11:1740
662 Zhao, Z.; Grohmann, S.; Zieger, L.; Dai, W.; Littke, R. Evolution of organic matter quantity and quality in a
663 warm, hypersaline, alkaline lake: The example of the Miocene Nördlinger Ries impact crater, Germany.
664 Frontiers in Earth Science 2022;10
665 Zheng, W.; Zhao, S.; Yin, Y.; Zhang, H.; Needham, D.M.; Evans, E.D.; Dai, C.L.; Lu, P.J.; Alm, E.J.; Weitz,
666 D.A. High-throughput, single-microbe genomics with strain resolution, applied to a human gut
667 microbiome. Science 2022;376:eabm1483
668 Zrimec, J.; Kokina, M.; Jonasson, S.; Zorrilla, F.; Zelezniak, A. Plastic-degrading potential across the global
669 microbiome correlates with recent pollution trends. MBio 2021;12:e02155-02121

670

671 **Fig. 1.** Summary of MAGs obtained from hypersaline lakes. A) A breakdown of completeness and
672 purity of recovered MAGs. Red: medium-quality MAGs, Green: high-quality MAGs, and Blue: Near-
673 complete (NC) MAGs. B) The proportion of 5S, 16S, 23S, and all three types of rRNA genes predicted
674 in MAGs. C) Summary of genes encoding standard tRNAs predicted in MAGs. D) Proportion of
675 MAGs that affiliated to a classified, candidate, or unknown taxon at different levels.

676

677 **Fig. 2.** Microbial composition and genetic characteristics of four saline lakes. A) Alpha indices of
678 predicted genes from planktonic and sedimentary communities in different lakes. B) RDA analysis of
679 predicted genes from planktonic and sedimentary communities in different lakes. C) Relative
680 abundance of genes involved in key pathways including carbon fixation, fatty acid metabolism,
681 methane metabolism, nitrogen metabolism, photosynthesis, and sulfur metabolism. D) Z-score
682 normalized relative abundance of modules in carbon fixation, methane metabolism, nitrogen
683 metabolism, and sulfur metabolism.

684

685 **Fig. 3.** Phylogenetic analysis of MDM MAGs. Unrooted maximum likelihood trees of archaeal and
686 bacterial MAGs were constructed with 122 and 120 concatenated marker genes in IQ-TREE using
687 Q.pfam+R10 and LG+F+R10 models, respectively. A) MDM proportion of planktonic and
688 sedimentary communities in terms of relative abundance and number of MAGs. Red: MDMs; Grey:
689 non-MDMs. B) Microbial structure of planktonic and sedimentary communities in different lakes in
690 terms of relative abundance and number of MAGs classified at phylum level. C) Phylogenetic tree of
691 bacterial and archaeal MAGs. MDM lineages (bacterial) or phyla (archaeal) highlighted in pink, and
692 MDM MAG(s) present in the corresponding lineages/phyla were indicated with shapes (ADH: square,
693 BLK: circle, DBC: triangle, QJJ: diamond, solid symbol: sediment, hallowed symbol: plankton).

694

695 **Fig. 4.** Analysis of biosynthetic potentials of hypersaline lake MAGs. A) A summary of predicted
696 BGCs. In the pie chart on top, blue indicates that predicted BGCs were clustered to known BGCs in
697 the MiBIG database with cosine distance ≤ 0.2; red indicates that predicted BGCs clustered to known
698 BGCs but with cosine distance > 0.2, which was considered novel; grey indicates that BGCs were not
699 clustered to a known BGC in the MiBIG database. The pie chart at bottom summarised the number of
700 BGCs annotated to a BGC class. B) Normalised number of BGCs per MAG in each phylum. C)
701 Phylogenetic trees of archaeal and bacterial MAGs with completeness ≥ 70, or number of BGCs ≥ 5.
702 Midpoint-rooted maximum likelihood trees of archaeal and bacterial MAGs were constructed with 122
703 and 120 concatenated marker genes in IQ-TREE using Q.pfam+R10 and LG+F+R10 models,
704 respectively. Black branches indicate bacterial MAGs, while purple branches indicate archaeal MAGs.

705

706 **Fig. 5.** Phylogenetic analysis of biodegradative potentials in A) bacterial and B) archaeal MDM MAGs.
707 Maximum likelihood trees of bacterial and archaeal MAGs were constructed with 120 and 122
708 concatenated marker genes in IQ-TREE using LG+F+R10 and Q.pfam+R10 models, respectively. The
709 bacterial and archaeal trees were rooted with phylum *GCA-001730085* and *Thorarchaeota*,
710 respectively. Key biodegradative genes were annotated in orange (nitrogen metabolism), teal
711 (dehalogenation), light blue (plastics degradation), and purple (PAHs degradation).

712

713

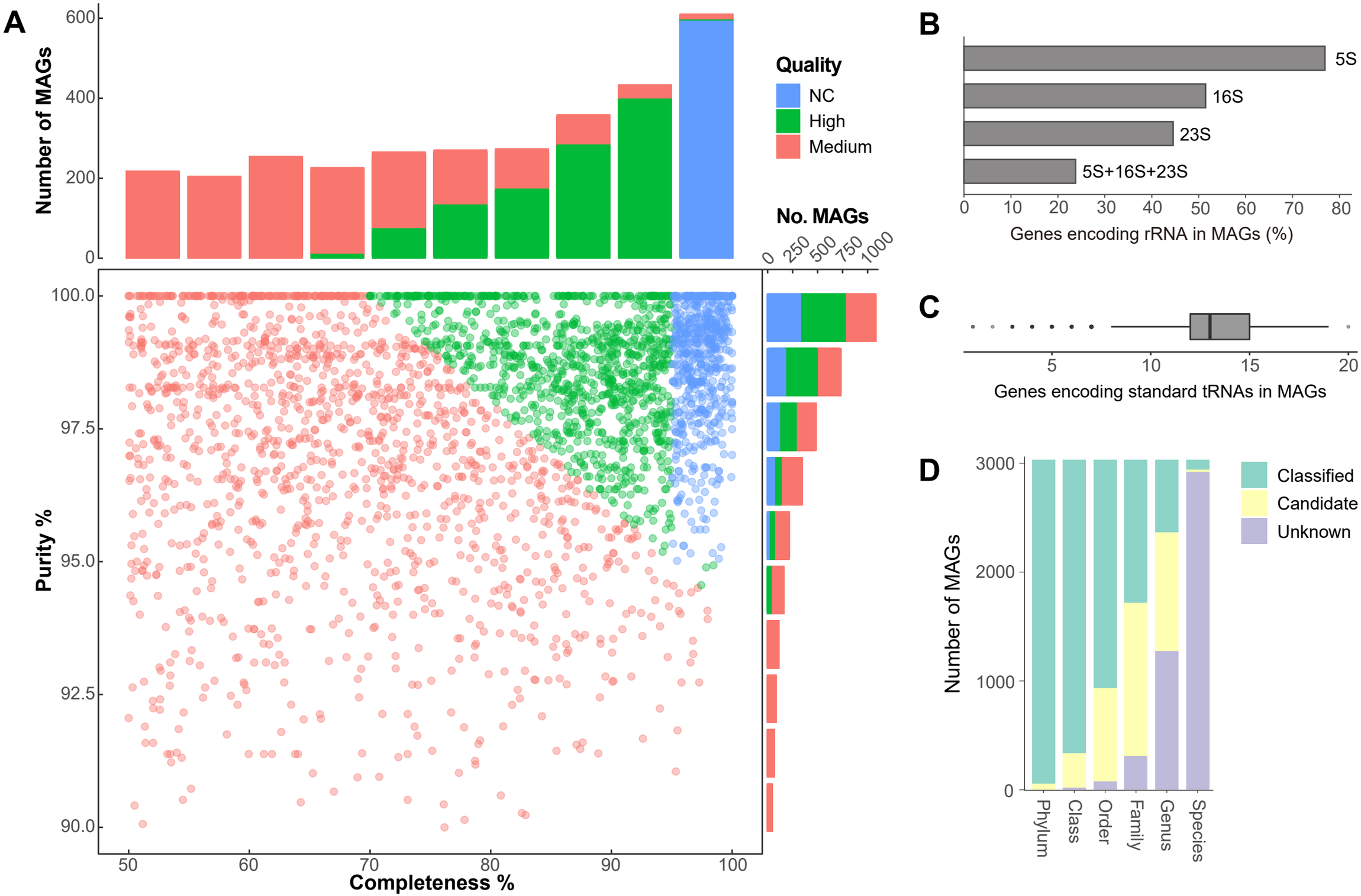
714

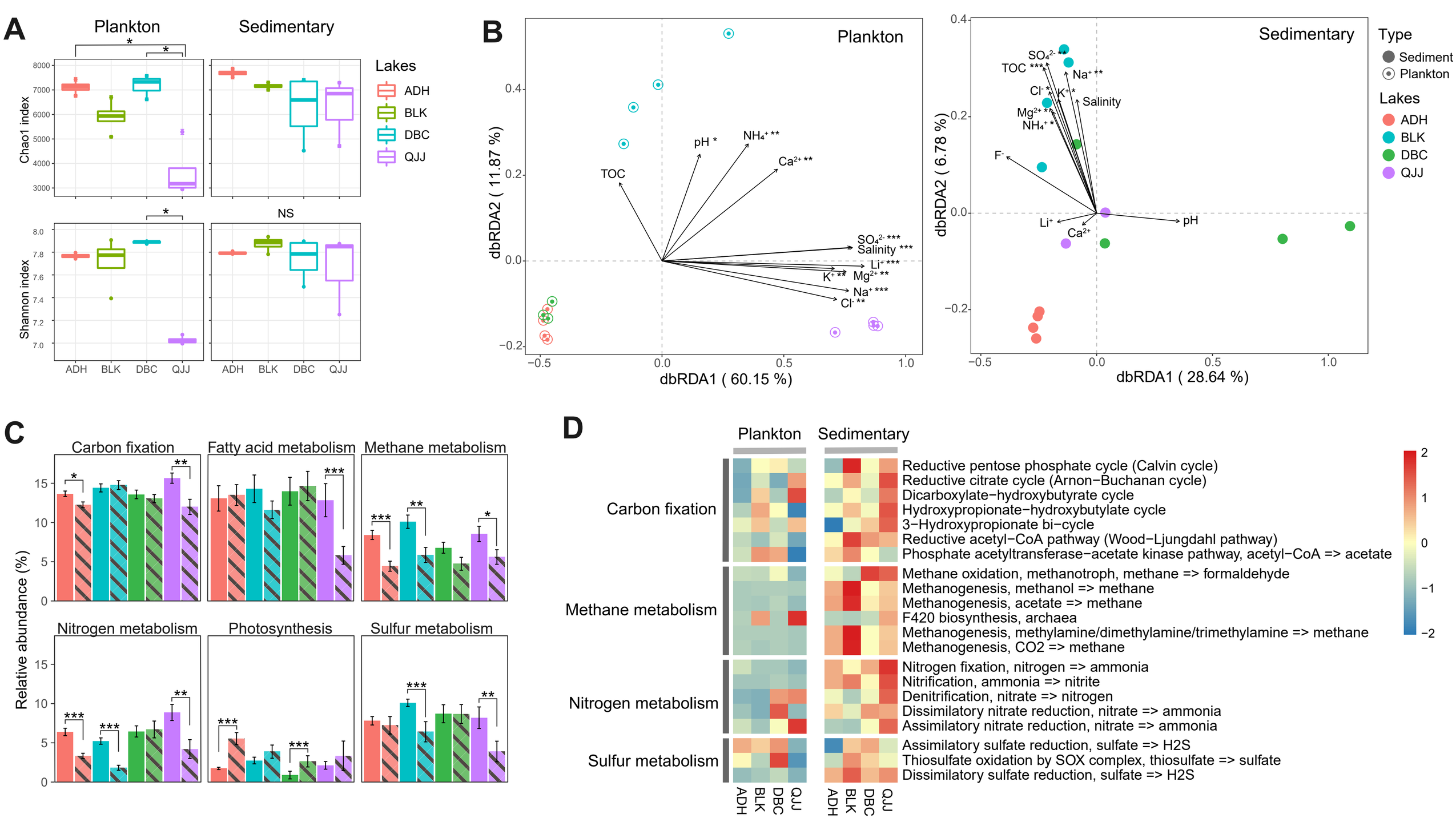
715

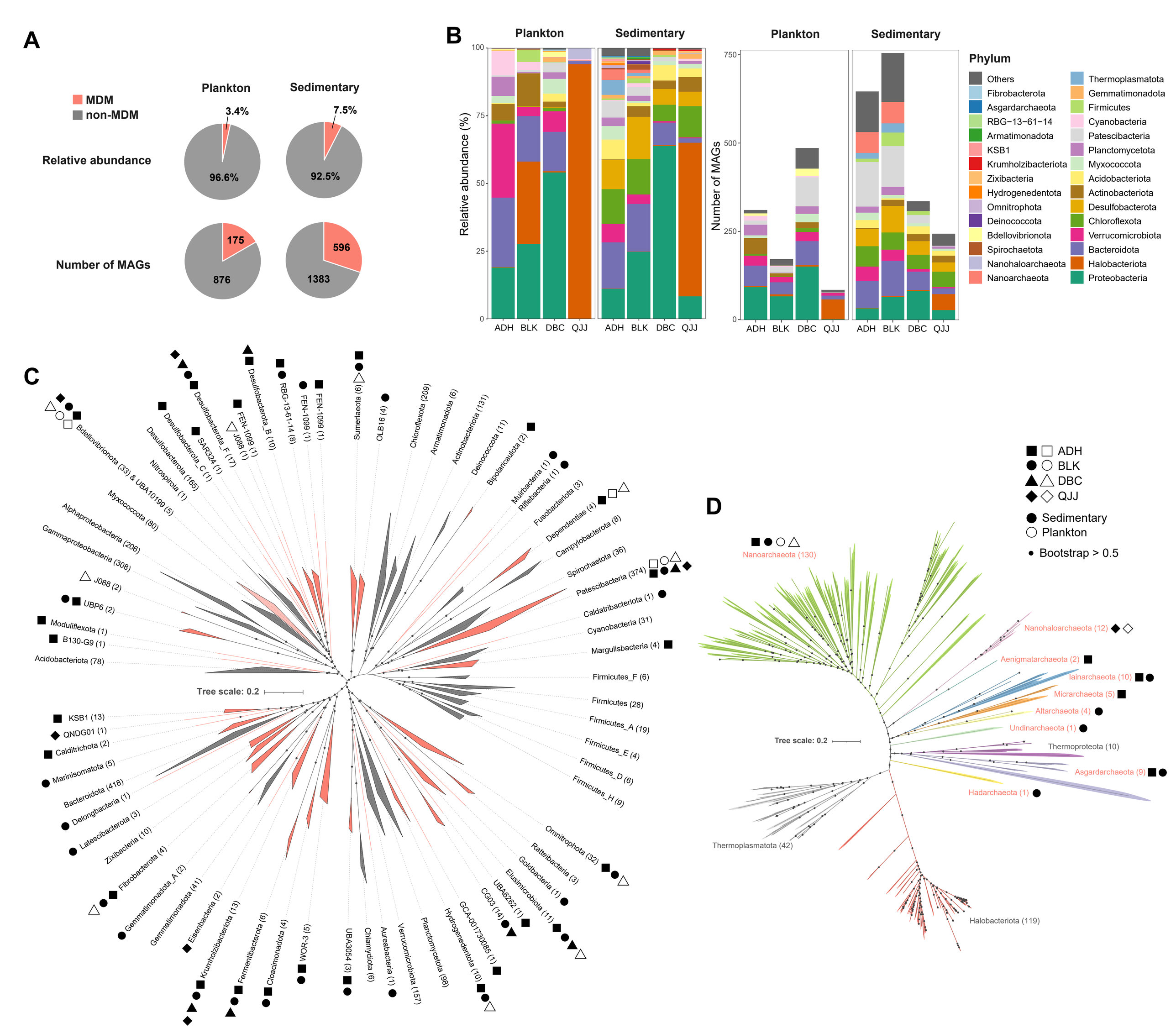
716

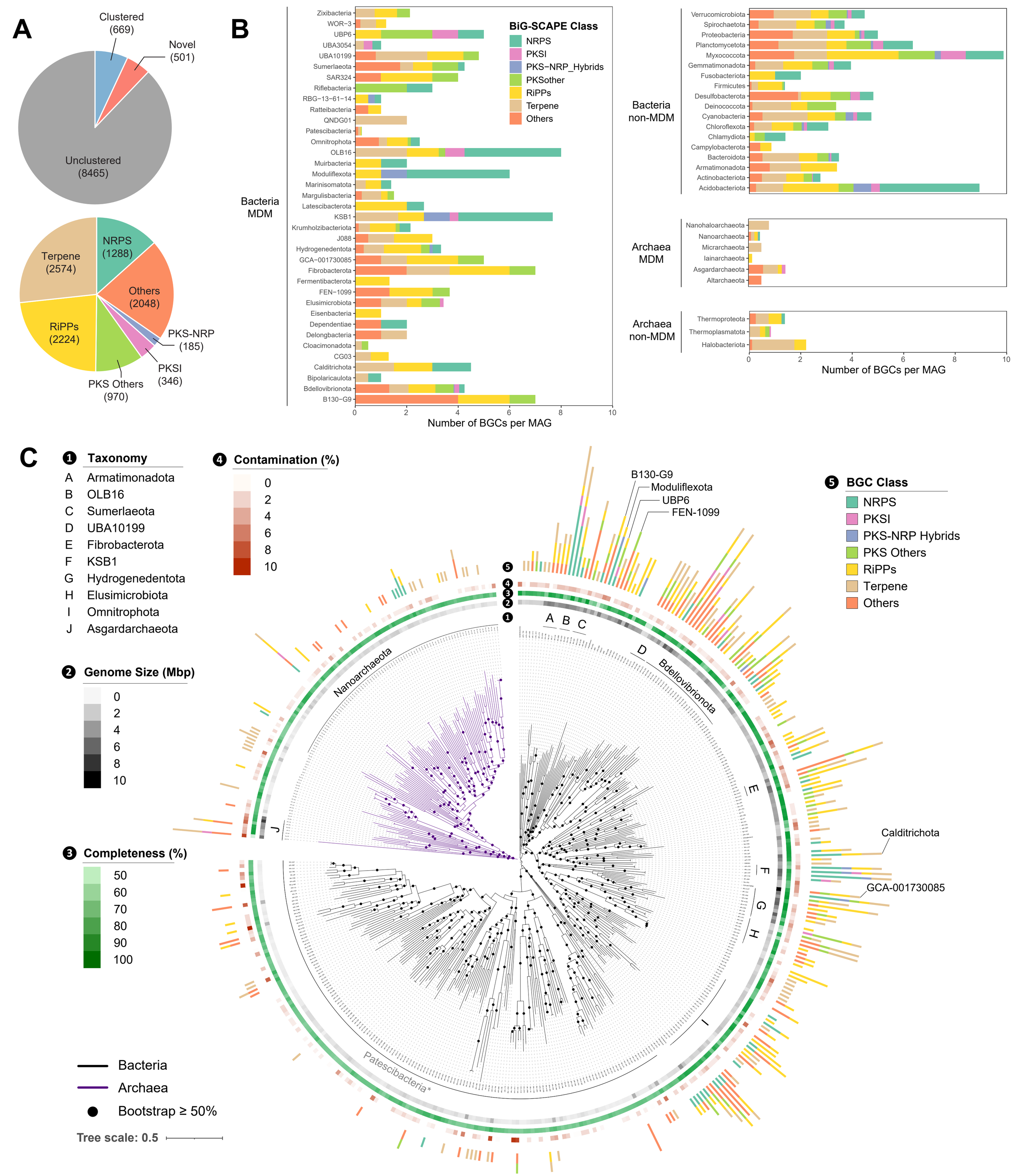
Fig. S1. Location of Aiding Lake (ADH), Barkol Lake (BLK), Dabancheng Lake (DBC), and Qijiaojing Lake (QJJ).

Fig. S2. NMDS plot of genetic diversities analysed in Bray-Curtis dissimilarities. Different colours indicate different lakes, including red (ADH), cyan (BLK), green (DBC), and purple (QJJ), while different shapes indicate different lake compartments, including sediment (solid circle) and plankton (hollowed circle).

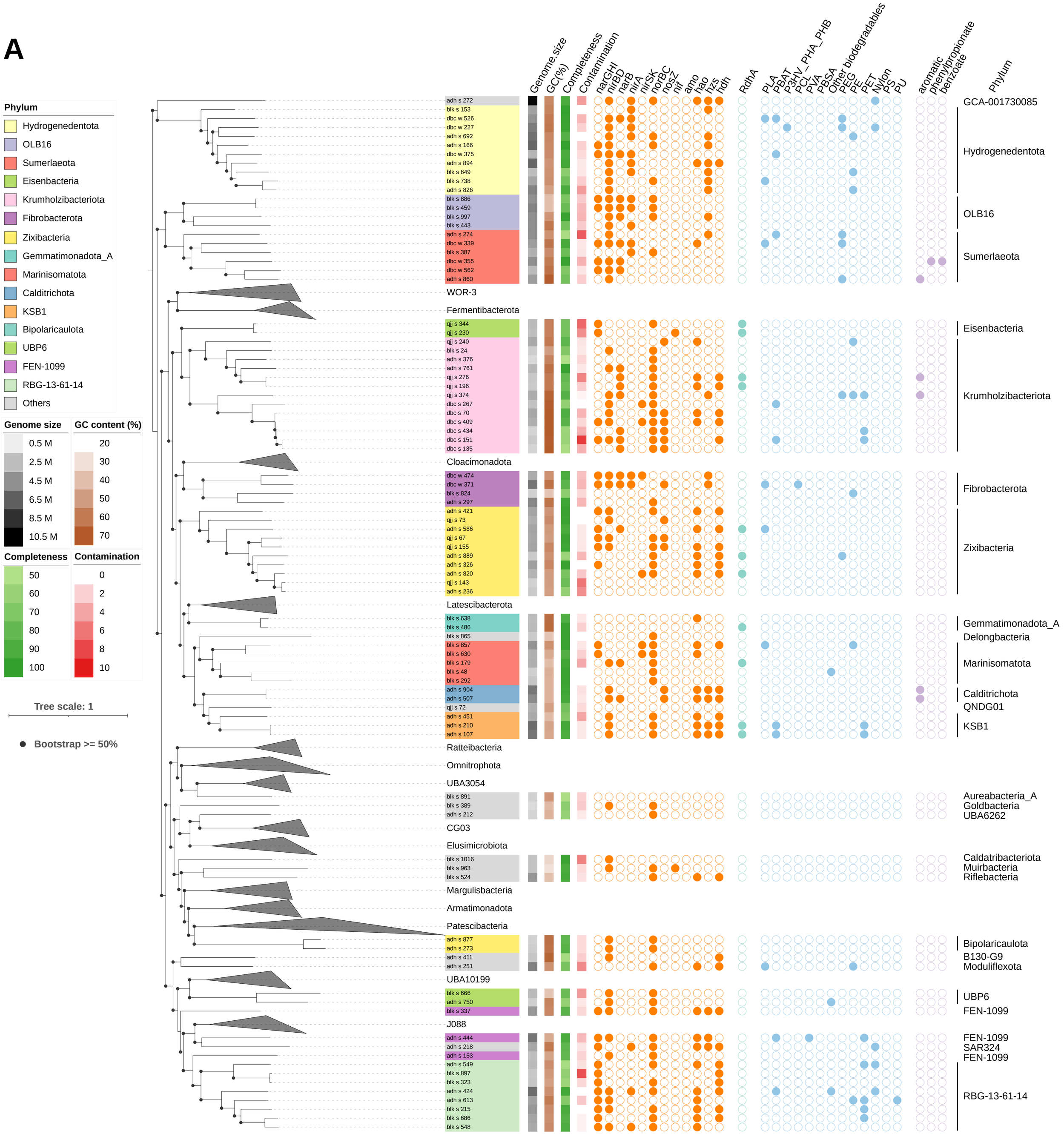

A
Number of MAGs
600
400
200
0
Quality
NC
High
Medium
No. MAGs
0
250
500
750
1000
Purity %
100.0
97.5
95.0
92.5
90.0
50
60
70
80
90
100
Completeness %
B
5S
16S
23S
5S+16S+23S
0
10
20
30
40
50
60
70
80
Genes encoding rRNA in MAGs (%)
C
5
10
15
20
Genes encoding standard tRNAs in MAGs
D
Number of MAGs
3000
2000
1000
0
Classified
Candidate
Unknown
Phylum
Class
Order
Family
Genus
Species

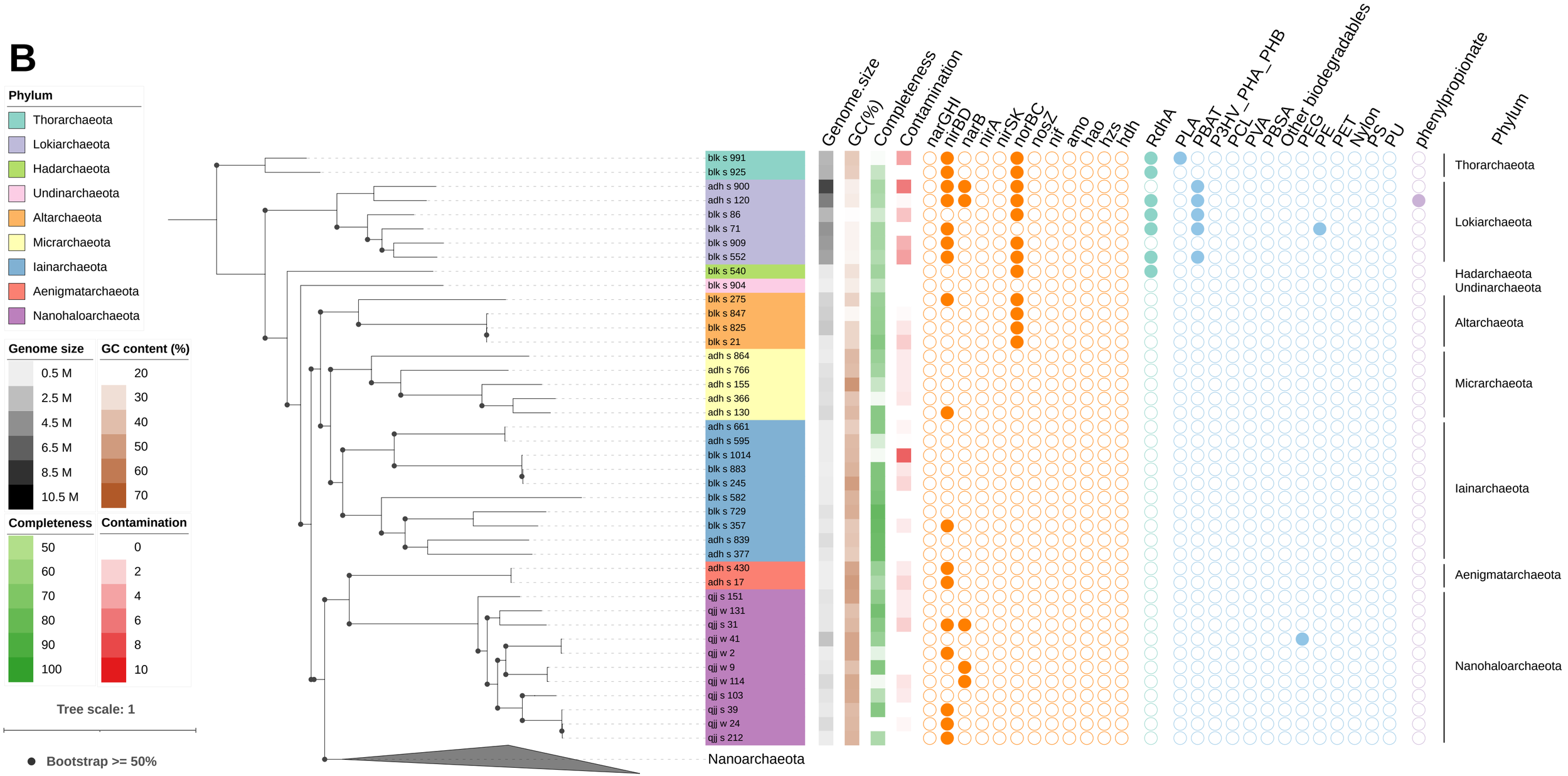

A
Plankton
Sedimentary
Chao1 index
Shannon index
NS
ADH
BLK
DBC
QJJ
Lakes
B
Plankton
Sedimentary
dbRDA1 (60.15 %)
dbRDA2 (11.87 %)
dbRDA1 (28.64 %)
dbRDA2 (6.78 %)
TOC
pH *
NH4+ **
Ca2+ **
SO42- ***
Salinity ***
Li+ ***
K+ **
Mg2+ **
Na+ ***
Cl- **
F-
pH
Type
Sediment
Plankton
Lakes
C
Carbon fixation
Fatty acid metabolism
Methane metabolism
Nitrogen metabolism
Photosynthesis
Sulfur metabolism
Relative abundance (%)
D
Plankton
Sedimentary
Carbon fixation
Methane metabolism
Nitrogen metabolism
Sulfur metabolism
Reductive pentose phosphate cycle (Calvin cycle)
Reductive citrate cycle (Arnon−Buchanan cycle)
Dicarboxylate−hydroxybutyrate cycle
Hydroxypropionate−hydroxybutylate cycle
3−Hydroxypropionate bi−cycle
Reductive acetyl−CoA pathway (Wood−Ljungdahl pathway)
Phosphate acetyltransferase−acetate kinase pathway, acetyl−CoA => acetate
Methane oxidation, methanotroph, methane => formaldehyde
Methanogenesis, methanol => methane
Methanogenesis, acetate => methane
F420 biosynthesis, archaea
Methanogenesis, methylamine/dimethylamine/trimethylamine => methane
Methanogenesis, CO2 => methane
Nitrogen fixation, nitrogen => ammonia
Nitrification, ammonia => nitrite
Denitrification, nitrate => nitrogen
Dissimilatory nitrate reduction, nitrate => ammonia
Assimilatory nitrate reduction, nitrate => ammonia
Assimilatory sulfate reduction, sulfate => H2S
Thiosulfate oxidation by SOX complex, thiosulfate => sulfate
Dissimilatory sulfate reduction, sulfate => H2S


A
MDM
non-MDM
Plankton
Sedimentary
Relative abundance
3.4%
96.6%
7.5%
92.5%
Number of MAGs
175
876
596
1383
B
Plankton
Sedimentary
Relative abundance (%)
Number of MAGs
ADH
BLK
DBC
QJJ
Phylum
Others
Fibrobacterota
Asgardarchaeota
RBG−13−61−14
Armatimonadota
KSB1
Krumholzibacteriota
Zixibacteria
Hydrogenedentota
Omnitrophota
Deinococcota
Bdellovibrionota
Spirochaetota
Nanohaloarchaeota
Nanoarchaeota
Thermoplasmatota
Gemmatimonadota
Firmicutes
Cyanobacteria
Patescibacteria
Planctomycetota
Myxococcota
Acidobacteriota
Actinobacteriota
Desulfobacterota
Chloroflexota
Verrucomicrobiota
Bacteroidota
Halobacteriota
Proteobacteria
C
Tree scale: 0.2
Bdellovibrionota (33) & UBA10199 (5)
Nitrospirota (1)
Desulfobacterota (165)
Desulfobacterota_C (1)
SAR324 (1)
Desulfobacterota_F (17)
J088 (1)
FEN-1099 (1)
Desulfobacterota_B (10)
RBG-13-61-14 (8)
FEN-1099 (1)
FEN-1099 (1)
Sumerlaeota (6)
OLB16 (4)
Chloroflexota (209)
Armatimonadota (6)
Actinobacteriota (131)
Deinococcota (11)
Bipolaricaulota (2)
Muirbacteria (1)
Riflebacteria (1)
Fusobacteriota (3)
Dependentiae (4)
Campylobacterota (8)
Spirochaetota (36)
Patescibacteria (374)
Caldatribacteriota (1)
Cyanobacteria (31)
Margulisbacteria (4)
Firmicutes_F (6)
Firmicutes (28)
Firmicutes_A (19)
Firmicutes_E (4)
Firmicutes_D (6)
Firmicutes_H (9)
Omnitrophota (32)
Rattelbacteria (3)
Goldbacteria (1)
Elusimicrobiota (11)
UBA6262 (1)
CG03 (14)
GCA-001730085 (1)
Hydrogenedentota (10)
Planctomycetota (98)
Verrucomicrobiota (157)
Aureabacteria (1)
Chlamydiota (6)
UBA3054 (3)
WOR-3 (5)
Cloacimonadota (4)
Fermentibacterota (6)
Krumholzibacteriota (13)
Eisenbacteria (2)
Gemmatimonadota (41)
Gemmatimonadota_A (2)
Fibrobacterota (4)
Zixibacteria (10)
Latescibacterota (3)
Delongbacteria (1)
Bacteroidota (418)
Marinisomatota (5)
Calditrichota (2)
QNDG01 (1)
KSB1 (13)
Acidobacteriota (78)
B130-G9 (1)
Moduliflexota (1)
UBP6 (2)
J088 (2)
Gammaproteobacteria (308)
Alphaproteobacteria (206)
Myxococcota (80)
D
Tree scale: 0.2
Nanoarchaeota (130)
Nanohaloarchaeota (12)
Aenigmatarchaeota (2)
Iainarchaeota (10)
Micrarchaeota (5)
Altarchaeota (4)
Undinarchaeota (1)
Thermoproteota (10)
Asgardarchaeota (9)
Hadarchaeota (1)
Thermoplasmatota (42)
Halobacteriota (119)
ADH
BLK
DBC
QJJ
Sedimentary
Plankton
Bootstrap > 0.5


A
Clustered (669)
Novel (501)
Unclustered (8465)
NRPS (1288)
Terpene (2574)
Others (2048)
RiPPs (2224)
PKS-NRP (185)
PKSI (346)
PKS Others (970)
B
BiG-SCAPE Class
NRPS
PKSI
PKS−NRP_Hybrids
PKSother
RiPPs
Terpene
Others
Bacteria MDM
Zixibacteria
WOR−3
UBP6
UBA3054
UBA10199
Sumerlaeota
SAR324
Riflebacteria
RBG−13−61−14
Ratteibacteria
QNDG01
Patescibacteria
Omnitrophota
OLB16
Muirbacteria
Moduliflexota
Marinisomatota
Margulisbacteria
Latescibacterota
KSB1
Krumholzibacteriota
J088
Hydrogenedentota
GCA−001730085
Fibrobacterota
Fermentibacterota
FEN−1099
Elusimicrobiota
Eisenbacteria
Dependentiae
Delongbacteria
Cloacimonadota
CG03
Calditrichota
Bipolaricaulota
Bdellovibrionota
B130−G9
0 2 4 6 8 10
Number of BGCs per MAG
Bacteria non-MDM
Verrucomicrobiota
Spirochaetota
Proteobacteria
Planctomycetota
Myxococcota
Gemmatimonadota
Fusobacteriota
Firmicutes
Desulfobacterota
Deinococcota
Cyanobacteria
Chloroflexota
Chlamydiota
Campylobacterota
Bacteroidota
Armatimonadota
Actinobacteriota
Acidobacteriota
Archaea MDM
Nanohaloarchaeota
Nanoarchaeota
Micrarchaeota
Iainarchaeota
Asgardarchaeota
Altarchaeota
Archaea non-MDM
Thermoproteota
Thermoplasmatota
Halobacteriota
0 2 4 6 8 10
Number of BGCs per MAG
C
1 Taxonomy
A Armatimonadota
B OLB16
C Sumerlaeota
D UBA10199
E Fibrobacterota
F KSB1
G Hydrogenedentota
H Elusimicrobiota
I Omnitrophota
J Asgardarchaeota
2 Genome Size (Mbp)
0 2 4 6 8 10
3 Completeness (%)
50 60 70 80 90 100
4 Contamination (%)
0 2 4 6 8 10
5 BGC Class
NRPS
PKSI
PKS-NRP Hybrids
PKS Others
RiPPs
Terpene
Others
B130-G9
Moduliflexota
UBP6
FEN-1099
Calditrichota
GCA-001730085
Nanoarchaeota
Bdellovibrionota
Patescibacteria*
Bacteria
Archaea
Bootstrap ≥ 50%
Tree scale: 0.5

A


Phylum
Hydrogenedentota
OLB16
Sumerlaeota
Eisenbacteria
Krumholzibacteriota
Fibrobacterota
Zixibacteria
Gemmatimonadota_A
Marinisomatota
Calditrichota
KSB1
Bipolaricaulota
UBP6
FEN-1099
RBG-13-61-14
Others
Genome size
0.5 M
2.5 M
4.5 M
6.5 M
8.5 M
10.5 M
GC content (%)
20
30
40
50
60
70
Completeness
50
60
70
80
90
100
Contamination
0
2
4
6
8
10
Tree scale: 1
Bootstrap >= 50%
Genome.size
GC(%)
Completeness
Contamination
narGHI
nirBD
narB
nirA
nirSK
norBC
nosZ
nif
amo
hao
hzs
hdh
RdhA
PLA
PBAT
P3HV_PHA_PHB
PCL
PVA
PBSA
Other biodegradables
PEG
PE
PET
Nylon
PS
PU
aromatic
phenylpropionate
benzoate
Phylum
GCA-001730085
Hydrogenedentota
OLB16
Sumerlaeota
WOR-3
Fermentibacterota
Eisenbacteria
Krumholzibacteriota
Cloacimonadota
Fibrobacterota
Zixibacteria
Latescibacterota
Gemmatimonadota_A
Delongbacteria
Marinisomatota
Calditrichota
QNDG01
KSB1
Ratteibacteria
Omnitrophota
UBA3054
Aureabacteria_A
Goldbacteria
UBA6262
CG03
Elusimicrobiota
Caldatribacteriota
Muirbacteria
Riflebacteria
Margulisbacteria
Armatimonadota
Patescibacteria
Bipolaricaulota
B130-G9
Moduliflexota
UBA10199
UBP6
FEN-1099
J088
FEN-1099
SAR324
FEN-1099
RBG-13-61-14


B
Phylum
Thorarchaeota
Lokiarchaeota
Hadarchaeota
Undinarchaeota
Altarchaeota
Micrarchaeota
Iainarchaeota
Aenigmatarchaeota
Nanohaloarchaeota
Genome size
0.5 M
2.5 M
4.5 M
6.5 M
8.5 M
10.5 M
GC content (%)
20
30
40
50
60
70
Completeness
50
60
70
80
90
100
Contamination
0
2
4
6
8
10
Tree scale: 1
Bootstrap >= 50%
Genome.size
GC(%)
Completeness
Contamination
narGHI
nirBD
narB
nirA
nirSK
norBC
nosZ
nif
amo
hao
hzs
hdh
RdhA
PLA
PBAT
P3HV_PHA_PHB
PCL
PVA
PBSA
Other biodegradables
PEG
PE
PET
Nylon
PS
PU
phenylpropionate
Phylum
blk s 991
blk s 925
adh s 900
adh s 120
blk s 86
blk s 71
blk s 909
blk s 552
blk s 540
blk s 904
blk s 275
blk s 847
blk s 825
blk s 21
adh s 864
adh s 766
adh s 155
adh s 366
adh s 130
adh s 661
adh s 595
blk s 1014
blk s 883
blk s 245
blk s 582
blk s 729
blk s 357
adh s 839
adh s 377
adh s 430
adh s 17
qjj s 151
qjj w 131
qjj s 31
qjj w 41
qjj w 2
qjj w 9
qjj w 114
qjj s 103
qjj s 39
qjj w 24
qjj s 212
Nanoarchaeota
Thorarchaeota
Lokiarchaeota
Hadarchaeota
Undinarchaeota
Altarchaeota
Micrarchaeota
Iainarchaeota
Aenigmatarchaeota
Nanohaloarchaeota