



**HAL**  
open science

# Microbial iron and carbon metabolism as revealed by taxonomy-specific functional diversity in the Southern Ocean

Ying Sun, Pavla Debeljak, Ingrid Obernosterer

► **To cite this version:**

Ying Sun, Pavla Debeljak, Ingrid Obernosterer. Microbial iron and carbon metabolism as revealed by taxonomy-specific functional diversity in the Southern Ocean. *The International Society of Microbiological Ecology Journal*, 2021, 10.1038/s41396-021-00973-3 . hal-03253204

**HAL Id: hal-03253204**

**<https://hal.science/hal-03253204>**

Submitted on 10 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17

**Microbial iron and carbon metabolism as revealed by taxonomy-specific functional diversity in the Southern Ocean**

**Running title: Microbial iron and carbon strategies in the Southern Ocean**

Ying Sun<sup>1\*†</sup>, Pavla Debeljak<sup>1,2\*</sup>, Ingrid Obernosterer<sup>1</sup>

1 CNRS, Sorbonne Université, Laboratoire d’Océanographie Microbienne, LOMIC, F-66650, Banyuls/mer, France.

2 University of Vienna, Department of Functional and Evolutionary Ecology, A-1090 Vienna, Austria.

\* These authors contributed equally to this study.

† Corresponding author: Ying Sun, Microbial Oceanography Laboratory (LOMIC), CNRS-Sorbonne University, 1 Avenue Pierre Fabre, 66650 Banyuls sur mer, France. Tel: (+33) 04 68 88 73 53. Email: [ying.sun@obs-banyuls.fr](mailto:ying.sun@obs-banyuls.fr).

18 **Abstract**

19 Marine microbes are major drivers of all elemental cycles. The processing of organic carbon by  
20 heterotrophic prokaryotes is tightly coupled to the availability of the trace element iron in large  
21 regions of the Southern Ocean. However, the functional diversity in iron and carbon metabolism  
22 within diverse communities remains a major unresolved issue. Using novel Southern Ocean  
23 meta-omics resources including 133 metagenome-assembled genomes (MAGs), we show a  
24 mosaic of taxonomy-specific ecological strategies in naturally iron-fertilized and high nutrient  
25 low chlorophyll (HNLC) waters. Taxonomic profiling revealed apparent community shifts across  
26 contrasting nutrient regimes. Community-level and genome-resolved metatranscriptomics  
27 evidenced a moderate association between taxonomic affiliations and iron and carbon-related  
28 functional roles. Diverse ecological strategies emerged when considering the central metabolic  
29 pathways of individual MAGs. Closely related lineages appear to adapt to distinct ecological  
30 niches, based on their distribution and gene regulation patterns. Our in-depth observations  
31 emphasize the complex interplay between the genetic repertoire of individual taxa and their  
32 environment and how this shapes prokaryotic responses to iron and organic carbon availability in  
33 the Southern Ocean.

34 **Introduction**

35 Heterotrophic prokaryotes play a key role in the cycling of elements in the ocean [1-3]. Efforts to  
36 decipher prokaryotes-environment interactions, facilitated by meta-omics technologies [4, 5],  
37 have revealed highly complex dynamics in marine microbial community structure and function  
38 [6-18]. Recent studies from global ocean expeditions provided insights into microbial diversity  
39 coupled with significant changes in gene repertoire and expression along temperature gradients  
40 and depths [13, 14, 18]. However, there are still fundamental gaps in understanding how  
41 individual prokaryotic taxa regulate their metabolic capacities in response to changes of other  
42 environmental factors, such as nutrient availability, and how their adaptation strategies may  
43 influence energy flows and nutrient cycling in their ecological niches.

44         The Southern Ocean remains one of the least explored ocean regions. The perennially  
45 cold waters present the largest high-nutrient, low-chlorophyll (HNLC) region of the global ocean,  
46 where iron is the primary limiting factor of biological productivity [19]. Heterotrophic  
47 prokaryotes experience a double constraint due to low concentrations of bioavailable iron and  
48 dissolved organic carbon (DOC) [20, 21]. The availability of these nutrients affects prokaryotic  
49 heterotrophic activities, particularly growth and respiration [22, 23], and as a consequence their  
50 functions in microbial food webs. Further, iron is present in various chemical forms in the ocean  
51 [24] and a multitude of substrates constitute the pool of organic matter [25], challenging the  
52 exploration of the functional roles of diverse prokaryotic taxa in accessing these essential  
53 resources.

54         Southern Ocean islands are a source of iron to the surrounding seawater, leading to  
55 localized spring phytoplankton blooms [26]. Kerguelen Island, located in the Indian Sector of the  
56 Southern Ocean, sustains the largest iron-fertilized region [27-30]. The annually occurring

57 diatom-dominated phytoplankton blooms east of the island have important consequences on  
58 microbial communities. Field studies and onboard experiments identified diatom-derived DOC  
59 as a key driver for prokaryotic diversity, activity, and seasonal community succession [21, 31-  
60 34]. The naturally-iron fertilized region off Kerguelen Island within HNLC waters provides a  
61 natural laboratory to examine phylogenetic and functional diversity of microbial lineages that  
62 mediate iron and carbon cycling.

63         In this study, we provide a comprehensive survey of the structure, genetic repertoire and  
64 expression pattern of the free-living (< 0.8  $\mu\text{m}$  size fraction) prokaryotic community in  
65 contrasting Southern Ocean productivity regions. We sampled three stations during the Marine  
66 Ecosystem Biodiversity and Dynamics of Carbon around Kerguelen (MOBYDICK) cruise in late  
67 austral summer (18th February to 30th March 2018), including one located in the naturally iron-  
68 fertilized waters and two off-plateau ones within HNLC waters. Metagenomic assembly and  
69 curation recovered a novel Southern Ocean meta-omics resource with 3 million protein-coding  
70 genes and characterized 133 metagenome-assembled genomes (MAGs) complementary to  
71 existing oceanic databases. Our main objective was to explore the distribution of prokaryotic  
72 functions related to iron and carbon metabolism in contrasting nutrient regimes and their links to  
73 taxonomy. We addressed this objective on the community and taxon-specific level by  
74 considering both the functional potential and the gene expression patterns.

75 **Materials and methods**

76 *Sample collection, metagenome and metatranscriptome sequencing*

77 Surface seawater (10 m) was collected at three stations in contrasting oceanic regions during the  
78 MOBYDICK cruise (**Supplementary Fig. 1A**). Station M2 was located in the naturally iron-  
79 fertilized waters above the central Kerguelen Plateau and stations M3 and M4 were located in  
80 off-plateau HNLC waters. The timing of the cruise covered the demise of the summer  
81 phytoplankton blooms (**Supplementary Fig. 1B**), as reflected in enhanced concentrations of  
82 dissolved organic carbon, prokaryotic abundance and heterotrophic production in on-plateau  
83 surface waters as compared to HNLC waters (**Supplementary Fig. 1C, Supplementary Table 1**  
84 **and Supplementary Methods**) [35]. Station M2 was visited three times at an 8-day interval, and  
85 station M3 and M4 were visited twice at a two-week interval. For metagenomes, triplicate 6 L  
86 seawater samples, collected by Niskin bottles, were each filtered through 0.8  $\mu\text{m}$  Polycarbonate  
87 filters (PC, Nuclepore). The cells in the  $< 0.8 \mu\text{m}$  fraction were concentrated in 0.2  $\mu\text{m}$  Sterivex  
88 filter units (Millipore). Total genomic DNA was extracted from the Sterivex filter units using the  
89 AllPrep DNA/RNA kit (Qiagen, Hiden, Germany) with modifications (**Supplementary**  
90 **Methods**). The DNA was extracted from the triplicate seawater samples collected during each of  
91 the repeated visits per station. Triplicate DNA extracts were pooled in equimolar amounts  
92 providing 1 pooled DNA extract per visit and station. The DNA extracts from the repeated visits  
93 (3 at M2 and 2 at each M3 and M4) were then pooled for each station to achieve 1  $\mu\text{g}$  in 30  $\mu\text{L}$   
94 Tris for sequencing purposes. Three metagenomic libraries (one per station) were prepared using  
95 the Illumina Nano library preparation kit. For metatranscriptomes, 10 L seawater samples were  
96 immediately pre-filtered through 0.8  $\mu\text{m}$  PC filters (Nucelpore) and the cells in the  $< 0.8 \mu\text{m}$   
97 fraction concentrated on 0.22  $\mu\text{m}$  Express Plus Polyethersulfate (PES) filters (Millipore). RNA

98 was extracted from the samples collected during the first visit at each site using the NucleoSpin®  
99 RNA Midi kit (Macherey-Nagel, Düren, Germany). Two internal standard RNA molecules were  
100 synthesized and added to each sample with known copy numbers. Technical details are provided  
101 in **Supplementary Methods**. Nine metatranscriptomic libraries (3 triplicates × 3 stations) were  
102 prepared using the Illumina TruSeq Stranded mRNA Library Prep kit. Paired-end sequencing (2  
103 × 150 bp) was performed on Illumina HiSeq 4000 platform at FASTER SA, Inc. (Switzerland).

#### 104 *Metagenome assembly and binning*

105 Quality control (QC) passed reads from each sample were co-assembled using MEGAHIT  
106 (v1.0.4) [36] with default settings, resulting 949,228 contigs with a minimum length of 1,000 bp  
107 (**Supplementary Methods**). Sequencing and assembly statistics are summarized in  
108 **Supplementary Fig. 2** and **Supplementary Table 2**. MetaWRAP (v1.1.3) [37] was used to  
109 assign contigs ( $\geq 2,500$  bp) into metagenome-assembled genomes (MAGs) with the aid of three  
110 binning tools, including CONCOCT (v1.0.0) [38], MaxBin (v2.2.5; -markerset 40 -  
111 prob\_threshold 0.5) [39] and MetaBAT (v2.12.1) [40]. Further refinement was implemented,  
112 based on read coverage and GC content of each contig in a MAG, by multivariate outlier  
113 detection using the aq.plot function in the mvoutlier package (v2.0.9) from R (v3.6.1). MAGs  
114 after contig outlier removal were reassessed using CheckM (v1.1.2) [41] for completeness and  
115 redundancy. Sequence-discrete populations closely related to the MAGs were identified as  
116 previously described using BMap (v38.22) [42] (**Supplementary Fig. 3**). Comparisons  
117 between our Southern Ocean assemblies and existing databases, including the NCBI and the  
118 TARA Ocean Global Expedition Project, were performed to evaluate the novelty of our data  
119 (**Supplementary Fig. 2** and **4**). Technical details are thoroughly described in **Supplementary**  
120 **Methods**.

121 *Metagenome functional profiling*

122 A total of 3,003,586 protein-coding genes were identified from the 949,228 contigs by Prodigal  
123 (v2.6.3) [43] under meta mode (-p meta). Functional annotation was carried out against eggnog  
124 [44] using eggNOG-mapper (v1.0.3) [45], Pfam [46] using HMMER (v3.2.1) [47], KEGG [48]  
125 using GhostKOALA (v2.2) [49] and KofamKOALA (v1.0.0) [50], TCDB [51] using BLASTP  
126 (v2.7.1) [52], CAZy [53] using dbCAN2 (v2.0.1) [54], and MEROPS [55] using BLASTP  
127 (v2.7.1) [52]. Iron-related genes were further examined by FeGenie [56], and Fe-containing  
128 domains were characterized using Superfamily (v1.75) [57].

129 *Metagenome taxonomic profiling*

130 Taxonomy classification of the 133 MAGs was determined using the classify\_wf function of the  
131 GTDB-Tk toolkit [58] based on the Genome Taxonomy Database (v0.3.0) (**Supplementary**  
132 **Table 3**). For phylogeny inference, 218 single-copy orthologous gene families shared by at least  
133 20 (out of 133) MAGs were identified by OrthoFinder (v2.2.3) [59], aligned with MAFFT  
134 (v7.313) [60] and filtered by trimAl (v1.4) [61]. Maximum Likelihood (ML) phylogenetic  
135 reconstruction was performed based on the concatenation of the proteins using IQ-Tree (v1.6.8; -  
136 m TESTMERGE -bb 1000 -bnni) [62] (**Fig. 1**). Metagenome-assembled genes which were not  
137 included in the MAGs were subjected to taxonomic classification using Kaiju (v1.7.0) [63] with  
138 its precompiled nr databases. We also quantified taxonomic diversity and relative abundance in  
139 each sample by using SSU reconstruction and assembly-free taxonomic classifiers  
140 (**Supplementary Fig. 5**). Technical details are thoroughly described in **Supplementary**  
141 **Methods**.

142 *Metatranscriptome transcript abundance and gene expression profiling*



143 Read counts of each gene were generated using featureCounts (v2.0.0) [64] with the BAM files  
 144 produced by mapping QC-passed meta-genomic/transcriptomic reads to the 949,228 annotated  
 145 contigs using Bowtie2 (v2.3.5) [65] (**Supplementary Methods**). Besides common shared  
 146 options including “-Q 1 --primary -p -B -P -C”, different settings were used for metagenomic (“-  
 147 O --fracOverlap 0.25 --ignoreDup -s 0”) and metatranscriptomic (“-s 2”) reads. Based on internal  
 148 standard recoveries (**Supplementary Methods**), we estimated the quantitative inventories of  
 149 transcripts per liter of each gene, and enumerated transcripts mediating key iron uptake and  
 150 carbon metabolism pathways (**Fig. 2, Supplementary Fig. 6-8, and Supplementary Table 4**).  
 151 Functional diversity was measured by Shannon index based on the abundance matrix of  
 152 functional groups in each sample using the “diversity” function of the vegan package in R  
 153 (v3.6.1). The abundance of each functional group in a sample (M2, M3 or M4) was defined as  
 154 the sum of all transcripts of genes assigned with the corresponding function,

155  $f_{Abundance}(Func_i) = \sum_{gene \in Func_i} f_{Transcripts L^{-1}}(gene)$ . The taxonomic composition of a  
 156 functional group was assessed by the ratio of  $\frac{\sum_{gene \in Func_i \cap Tax_j} f_{Transcripts L^{-1}}(gene)}{\sum_{gene \in Func_i} f_{Transcripts L^{-1}}(gene)}$ .

157 Further, differential gene expression analyses were performed in two ways, including on  
 158 the original metatranscriptomic read counts and on the metagenome-normalized  
 159 metatranscriptomic profile. The normalization was performed to minimize the influence of  
 160 genome abundance on the assessment of gene expression levels, given that fluctuations in  
 161 transcript abundance could be a result of shifting genome copies rather than changes in  
 162 expression levels (**Supplementary Fig. 9A**) [14]. We normalized the metatranscriptomic profile  
 163 by relative gene abundance through the division of variance-stabilizing transformed count tables,  
 164 and then converted the ratios to integer pseudo-counts in a range of 0 to  $10^6$  (**Supplementary**  
 165 **Fig. 9B**). DESeq2 (v1.24.0) was applied to identify significantly differentially expressed genes

166 (SDEGs) across contrasting oceanic regions (on- vs. off-plateau) at a false discovery rate (FDR)  
167 threshold of 0.1 [66]. Considering that during the MOBYDICK cruise the two HNLC sites (M3  
168 and M4) were located at distinct water bodies separated by the Antarctic Polar Front  
169 (**Supplementary Fig. 1A**), we included in our design formula, besides the factor of iron  
170 concentration gradients, a term representing the influence of the Antarctic Polar Front. Genes,  
171 identified as significantly differentially expressed, were further summarized according to their  
172 functional categories (**Fig. 3, Supplementary Fig. 10-11 and Supplementary Table 4-5**).

173 **Results and Discussion**

174 *A novel Southern Ocean meta-omics resource*

175 An average of 316.4 million (M) pairs of high-quality metagenomic reads were obtained from  
176 each station, achieving approximately 95% average coverage of the sampled communities  
177 (**Supplementary Fig. 2A** and **Supplementary Table 2**) [67]. Combined with another pre-  
178 sequenced metagenome from station M2 in early spring, a total of 1,286.5 M pairs of reads were  
179 assembled into 949,228 non-redundant contigs ( $\geq 1,000$  bp), on which 3,003,586 protein-coding  
180 genes were identified. Most contigs and genes in the assembled metagenome had low similarity  
181 to sequences in the NCBI nt database (**Supplementary Fig. 2B-E** and **Supplementary**  
182 **Methods**), underlying the novelty of our data. Although almost half of the predicted proteins  
183 displayed homology with high similarity to sequences in the NCBI nr database (bitscore  $\geq 200$   
184 and E-value  $< 1e-10$ ) [68], the amount of near identical matches (percentage of identity  $\geq 90\%$ )  
185 does not exceed 15.32% and another 717,088 (23.87%) proteins have no significant homologs  
186 found in the nr database (**Supplementary Fig. 2FG**).

187 A total of 133 Southern Ocean MAGs were recovered, among which 116 have a  
188 completeness  $\geq 50\%$  and a redundancy  $< 5\%$  (**Fig. 1** and **Supplementary Table 3**). The  
189 Southern Ocean MAGs represent a wide range of taxonomic groups, including 4 archaeal and  
190 129 bacterial genomes. The classes of Alphaproteobacteria (n=34), Gammaproteobacteria (n=35)  
191 and Bacteroidia (n=39) dominated the bacterial Southern Ocean MAGs, while other members  
192 belonging to Planctomycetota (n=5), Myxococcota (n=3), Verrucomicrobiota (n=4) and  
193 Actinobacteriota (n=3) were also present. Metagenomic read recruitment revealed variable  
194 abundance for some taxonomic groups across sampling sites, including the Pelagibacterales  
195 order (also known as the SAR11 clade) and Flavobacteriaceae family (**Supplementary Fig. 3**).

196           Only 13 of our Southern Ocean MAGs conform to > 95% intra-species ANI values with  
197 counterparts from the TARA Ocean Global Expedition (**Supplementary Fig. 4A** and  
198 **Supplementary Table 3**) [69]. Metagenomic read recruitment analysis further confirmed that  
199 the novelty of our Southern Ocean assemblies is not derived from biases introduced during  
200 metagenome assembly and binning (**Supplementary Fig. 4B-F** and **Supplementary**  
201 **Information**). The protein-level comparison indicated more functional similarity than diversity,  
202 given that around 90.51% of the proteins in the Southern Ocean MAGs were assigned with  
203 orthologs in the TARA assemblies (**Supplementary Fig. 4GH**).

#### 204 *Taxonomic profiling of prokaryotic communities*

205 Regarding the considerable amount of metagenomic reads that could not be assembled into  
206 MAGs (**Supplementary Table 2**), we carried out additional taxonomic profiling analyses using  
207 both 16S rRNA reconstruction-based and assembly-free methods (**Supplementary Methods**), in  
208 order to obtain a less biased estimate of the microbial community composition in our samples.  
209 Overall, the dominance of Alphaproteobacteria, Gammaproteobacteria and Bacteroidia, as well  
210 as the variations in diversity and abundance of individual taxa across samples (**Supplementary**  
211 **Fig. S5ABE**), agreed with the observations in the 133 Southern Ocean MAGs. The SAR11 clade  
212 was ubiquitous and abundant across all sampling sites, and its 16S rRNA assemblies displayed  
213 high phylogenetic diversity (**Supplementary Fig. 5A**). Species diversity, measured by Shannon  
214 index, was higher for the microbial communities in the off-plateau HNLC M3 and M4 sites than  
215 those in the on-plateau M2 site (**Supplementary Fig. S5C**). Microbial community variability  
216 among stations was explored with double principal coordinate analysis (DPCoA) followed by  
217 Monte Carlo permutation tests, incorporating not only information on abundance patterns but  
218 also phylogenetic structures (**Supplementary Methods**). Ordination of communities by DPCoA

219 revealed a significant clustering of taxonomic groups along the first principal component  
220 correlated with contrasting nutrient regimes (p-value  $\approx 0.001$ ; **Supplementary Fig. S5D** and  
221 **Supplementary Methods**). However, the statistical significance of categorical explanatory  
222 variable (on- vs. off-plateau waters) could not be appropriately assessed due to the small number  
223 of representative samples per environment (**Supplementary Methods**).

#### 224 *Community functional potential and gene expression patterns*

225 To assess the functional potential and gene expression patterns at the community level, we  
226 examined key functions in carbon and iron metabolism and the contribution of prokaryotic taxa  
227 to the respective functional groups across the metagenomes and metatranscriptomes obtained  
228 from the different sites (**Supplementary Fig. 6-7** and **Supplementary Table 4**). The overall  
229 functional potential, based on gene presence and absence in metagenomes, was similar across  
230 sampling sites at the community level (**A-D** in **Supplementary Fig. 6-7**). Key metabolic genes  
231 were universally present in all samples, including those involved in iron uptake and  
232 carbohydrate-active enzymes (CAZymes). We further examined the community-level transcript  
233 abundance based on the normalized per-liter transcripts estimated following the internal  
234 standards protocol (**E-J** in **Supplementary Fig. 6-7**) [70]. We did not observe an overall  
235 enrichment of functional groups related to iron and carbon metabolism in either the on- or off-  
236 plateau prokaryotic communities, but the patterns were quite patchy. For example, the  
237 siderophore transporters belonging to ExbD (K03559), ExbB (K03561), TonB (K03832), TonB-  
238 dependent outer membrane receptors (K16087 and K02014), ferrous iron transporter FeoA  
239 (K04758) and FeoB (K04759), as well as two transcriptional regulators Fur (K09823) and TroR  
240 (K03709), had higher abundance in the iron-fertilized waters. To the contrary, the vitamin B12  
241 transporter (K16092), putative heme transport protein HmuS (K07225), heme iron utilization

242 protein HugZ (K07226), vacuolar iron transporter VIT (K22736), ferredoxin/ flavodoxin switch  
243 relate HemG (K00230), ferric transporters FbpA (K02012), a siderophore transporter (K02016),  
244 and another Fur transcriptional regulator Irr (K09826) was more abundant in the off-plateau  
245 waters.

246         To better explain this mosaic pattern, we explored the possible link between taxonomy  
247 and function. We measured the functional diversity using the Shannon index based on the  
248 abundance matrix of functional groups across samples. In contrary to the species diversity  
249 (**Supplementary Fig. S5C**), the functional diversity of the on-plateau M2 site was no less than  
250 the average of the off-plateau M3 and M4 sites (**Fig. 2A** and **Supplementary Fig. 8A**). That is,  
251 the evidently lower species diversity in the on-plateau waters was decoupled from the  
252 community functional structure. The taxonomic compositions within functional groups across  
253 study sites provided a complementary perspective (**Supplementary Fig. 6-7**). Overall, the  
254 SAR11 clade contributed slightly higher to the gene pool and transcript inventories in the off-  
255 plateau waters, whilst Flavobacteriales made up a larger share in the iron-fertilized on-plateau  
256 zone (**Fig. 2B** and **Supplementary Fig. 8B**). The Roseobacterales and Gammaproteobacteria  
257 adopted more flexible ecological strategies, as their contributions to the functional pool were  
258 similar in different waters. This pattern is consistent with the clear separation in community  
259 taxonomic composition across divergent environmental conditions (**Supplementary Fig. 5D**)  
260 and suggests the aforementioned mosaic transcript abundance as a result of environmental  
261 nutrient availability and microbial life strategies. That is to say, when the variation in  
262 environmental conditions leads to the selection for specific metabolic functions (e.g., DOC  
263 degradation, access to iron), the taxonomic variation within functional groups would be a result

264 of both the importance of the specific function and the phylogenetic distribution of those  
265 functions [71].

266 We further recovered the gene expression profiles by normalizing the metatranscriptomic  
267 transcript abundance using the metagenomic gene abundance (**Supplementary Fig. 9**). The  
268 SDEGs obtained with and without the metagenome-based normalization were partially  
269 overlapped (**Supplementary Fig. 10-11**), confirming that prokaryotic community transcripts  
270 vary as a function of shifts in both community composition and gene expression levels [14]. We  
271 classified SDEGs according to their functional groups and taxonomic affiliations, and confirmed  
272 that gene expression patterns were not fully determined by nutrient regimes, but more taxonomy-  
273 resolved with microenvironmental considerations (**Fig. 3** and **Supplementary Fig. 10-11**).

274 The Flavobacteriales group and Gammaproteobacteria constituted the majority of the  
275 SDEGs belonging to the glycoside hydrolysis (GH) and glycosyltransferase (GT) families, which  
276 primarily have higher expression levels in the on-plateau iron-fertilized waters. Among them, the  
277 most corresponding GH families included GH16 and GH17 responsible for the decomposition of  
278 glucans and galactans, GH92 for the degradation of mannoses, and the  $\beta$ -1,3-D-glucan  
279 phosphorylases GH149 with inconclusive roles [72]. Gammaproteobacteria and Flavobacteriales  
280 were also enriched in the SDEGs of the GH3 family, which facilitates the utilization of glucose,  
281 arabinose and xylose. These GH3 SDEGs were higher expressed in either the on- or off-plateau  
282 region. This intra-taxonomy difference in expression patterns across different nutrient regimes  
283 suggests a mixture of copio- and oligotrophic life strategists applied within these two taxonomic  
284 groups [73]. On the contrary, the SDEGs of the GH23 and GH73 families were mainly from the  
285 SAR11 clade. GH23 (lytic transglycosylases) and GH73 ( $\beta$ -N-acetylglucosaminidases) are both  
286 involved in peptidoglycan degradation, an essential macromolecule of the bacterial outer cell

287 wall. The higher expression of these families in the off-plateau HLNC waters could indicate the  
288 use of peptidoglycan as carbon source or the accelerated growth of SAR11.

289 The SDEGs involved in Fe-uptake and Fe-related pathways exhibited similar patterns  
290 (**Fig. 3** and **Supplementary Fig. 11**). SAR11 constituted a great proportion of the SDEGs coding  
291 for two iron-related transcriptional regulators (IscR and Irr; **Fig. 3**). IscR monitors Fe-S cluster  
292 homeostasis and is responsible for the autorepression of genes involved in Fe-S cluster  
293 biogenesis, such as the *sufBCD* operon [74]. Under oxidative stress and iron starvation, IscR is in  
294 its apoform and relieves its repression of the *suf* operon [75, 76]. Irr, a global regulator of iron  
295 homeostasis, functions as a sensor of the cellular heme biosynthesis and accumulates under iron  
296 limitation to control target genes [77, 78]. It is reported to be conserved in the SAR11 subgroup  
297 Ia and maintained by selection due to fitness advantage [79, 80]. The induction of the glyoxylate  
298 shunt (GS) is an efficient strategy for heterotrophic prokaryotes to maintain growth and  
299 respiration rates under iron stress [23, 81]. We examined three key enzymes related to the GS,  
300 including isocitrate lyase (K01637; ICL encoded by *aceA*) and malate synthase (K01638; MS  
301 encoded by *aceB*) within the GS pathway, as well as isocitrate dehydrogenase that catalyses the  
302 oxidative decarboxylation of isocitrate (K00031; IDH encoded by *icd*) (**Supplementary Fig. 11**).  
303 The upregulation of *aceA* and *aceB* indicates the elevation of the GS under stress conditions,  
304 whereas the variant expression of *icd* may provide a clue to the competition between IDH and  
305 ICL for the substrate isocitrate. The SAR11 clade accounts for a large share of the *aceA* and  
306 *aceB* SDEGs which have higher expression levels in off-plateau waters. The community-level  
307 abundance of GS-related transcripts generally agree with the observations of SDEGs  
308 (**Supplementary Fig. 6**). For instance, the amount of SAR11 *aceA* transcripts L<sup>-1</sup> were more  
309 than doubled in the off-plateau samples. The enrichment of both *aceA/aceB* and *icd* suggests that



310 the SAR11 GS system functions supplementary to the classic TCA cycle in response to iron  
311 and/or carbon limitation. Gammaproteobacteria and Flavobacteriales dominated the SDEGs  
312 encoding the TonB-ExbB-ExbD complex for siderophore uptake. Related functional groups were  
313 generally enriched with SDEGs in both the iron-fertilized and the HNLC waters at the  
314 community level, however associated to different taxa.

### 315 *Taxon-specific ecological roles*

316 To resolve ecological roles of prokaryotic taxa across contrasting oceanic waters, we proceeded  
317 our data mining effort at a finer resolution with the 133 MAGs (**Fig. 1**). Initially, we performed a  
318 systematic survey for metabolically active prokaryotes through transcript abundances. Ribosomal  
319 proteins (RP) are critical for protein synthesis and levels of RP transcripts have been proposed as  
320 an indicator for prokaryotic growth rates [82-85]. We surveyed 93 prokaryotic RP KEGG  
321 Orthology groups (KOs) through all our assemblies (**Supplementary Table 4**). Generally, taxa  
322 with higher grow rates (more RP transcripts) had also high cell metabolism (more total  
323 transcripts) in the off-plateau HNCL waters ( $R^2 > 0.9$ ), whereas in the on-plateau zone several  
324 MAGs showed an all-vs.-RP ratio depart from the fitted line ( $R^2 < 0.5$ ; **Fig. 4**). This provided us  
325 with interesting insights. First, environmental properties in the iron-fertilized on-plateau zone  
326 lead to a decoupling between cell metabolism and growth of several MAGs. Secondly, individual  
327 species from closely related taxonomic groups revealed diverse ecological strategies. For  
328 example, MAG\_91, although it forms a monophyletic clade with MAG\_126 on the phylogeny  
329 tree (**Fig. 1**), were better adapted to the HNLC environment at site M4 as compared to the on-  
330 plateau zone (**Fig. 4**).

331 The diversity in ecological strategies became more evident when we examined central  
332 metabolic pathways gene by gene (**Fig. 5, Supplementary Fig. 12 and Supplementary Table 4-**

333 5). While the SAR11 SDEGs tuned their expression in a relatively consistent manner among  
334 individual MAGs, we observed diverse expression patterns of SDEGs belonging to  
335 Flavobacteriales and Gammaproteobacteria. Generally, the Flavobacteriales MAGs constituted  
336 more genes that were significantly higher expressed in the iron-fertilized on-plateau zone,  
337 whereas a limited number of them displayed the opposite pattern by downregulating the  
338 expression of genes responsible for iron uptake and carbon metabolism in the same water.

339         Particularly, with respect to polysaccharide metabolism, two MAGs (MAG\_51 and 99),  
340 belonging to the same Flavobacteriales UA16 genus but sharing an inter-species ANI value of  
341 79.05% [69], exhibited contrasting expression patterns. Moreover, four MAGs (MAG\_78, 3, 134  
342 and 73) from the Flavobacteriales 1G12 family but distinct genera, which formed a monophyletic  
343 clade on the phylogenomic tree, also reflected niche divergence. For instance, polysaccharide  
344 utilization loci (PULs) are operon-like gene structures that encode co-regulated proteins that  
345 specialize in polysaccharide detection, uptake and hydrolysis [86]. PULs are prevalent in the  
346 Bacteroidota phylum, and typically feature with SusD-like substrate-binding proteins, TonB-  
347 dependent receptors (TBDR) and various CAZymes. We identified several PUL-like loci in four  
348 Flavobacteriales UA16 MAGs (**Supplementary Fig. 13A-D** and **Supplementary Methods**). All  
349 the UA16 MAGs consisted of a “GH149 + GH30 + GH16 (and/or GH17) + transporters” PUL  
350 structure, indicating a general utilization of glucans and galactans, but MAG\_73 contained a  
351 unique “GH92 + GH78 (+ CBM67 domain) + transporter” locus. MAG\_73 was ubiquitous in all  
352 sampling sites but more abundant in the off-plateau waters (**Fig. 1** and **Supplementary Table 3**)  
353 and exhibited an opposite expression pattern to the other three (**Fig. 5**). The GH92 family exo- $\alpha$ -  
354 mannosidases function on  $\alpha$ -linked mannose residues in an exo-acting manner and therefore are  
355 responsible for the depolymerization of  $\alpha$ -linked mannans [87]. Algal mannans are widely

356 distributed in marine ecosystems, and  $\alpha$ -mannans were identified in red seaweed [88, 89] and the  
357 diatom *Phaeodactylum tricornutum* [90]. Recent studies demonstrated that marine bacteria,  
358 especially Bacteroidota, can degrade mannans [91-93]. Although whether the GH92 PUL  
359 facilitates the growth of MAG\_73 across all sampling sites requires further study due to the  
360 incompleteness of MAGs. The utilization of mannans elucidates the metabolic potential of  
361 MAG\_73 and suggests that it could occupy distinct niches as compared to MAG\_78, 3 and 134.  
362 Fucose is another bioavailable monosaccharide common in marine waters and released by diatoms  
363 [94]. We identified candidate gene clusters specific for fucose utilization among  
364 Verrucomicrobiae MAGs, showing orthology to the recently discovered functional loci in  
365 *Lentimonas* sp. CC4 (**Supplementary Fig. 13E**) [95].

366 We also observed MAGs with similar expression patterns in late summer above the  
367 plateau, but contrasting abundances in samples from early spring at the same site  
368 (**Supplementary Table 3**) providing subtle clues to seasonal adaptation. This was specifically  
369 the case for the metabolically active on-plateau MAG\_103 and MAG\_62 belonging to the  
370 Pseudomonadales HTCC2089 family but different genera (**Fig. 5**). To explore the potential  
371 genetic reasons behind this observation, we constructed pan-genomes by using both our MAGs  
372 and their closely related reference genomes. The incorporation of reference genomes into our  
373 analysis is to compensate the incompleteness of our MAGs. For the comparison between the  
374 MAG\_103 and MAG\_62, a total of 19 Pseudomonadales HTCC2089 draft genomes were  
375 retrieved from NCBI GenBank database based on the phylogenetic information provided by  
376 GTDB [58], including 15 from UBA4421 genus and 4 from UBA9926 (**Supplementary**  
377 **Methods**). The two MAGs shared most of their polysaccharide degradation and proteolysis  
378 genes and perceived to be competitors for similar resources (**Supplementary Table 6**). We

379 identified one singleton chitinase (GH18) unique to MAG\_103 but missing from all other  
380 Pseudomonadales HTCC2089 draft genomes, and three GHs (GH17, GH149 and GH158)  
381 conserved in MAG\_62 but absent from the UBA4421 genus. However, as discussed above,  
382 GH149 is with indecisive function, and GH17 and GH158 share similar substrate specificities  
383 with GH16, which is common in HTCC2089. The number of peptidases-encoding genes was  
384 also comparable in MAG\_103 (n=102) and MAG\_62 (n=104).

385         We made an unexpected observation that could explain the different abundance patterns  
386 of these MAGs in early spring and late summer. We identified a gene cluster in MAG\_103  
387 related to light-induced energy acquisition, which was conserved in the UBA4421 genus but  
388 absent from the UBA9926 genomes (**Supplementary Fig. 14AB** and **Supplementary Table 6**).  
389 This gene cluster consisted of 6 genes, encoding a bacteriorhodopsin (PF01036.18), a synthase  
390 (PF00348.17), a phytoene desaturase (*crtI*; K10027), a 15-cis-phytoene synthase (*crtB*; K02291),  
391 a lycopene beta-cyclase (*crtLI*; K06443), and a beta-carotene 15,15'-dioxygenase (*blh*; K21817).  
392 Bacteriorhodopsin could facilitate MAG\_103 with the capability to use light as a supplemental  
393 energy source [96, 97]. The MAG\_103 bacteriorhodopsin sequence contained a blue light  
394 absorbing glutamine “Q” and a proton pumping motif “DTE” (**Supplementary Fig. 14C**) [98].  
395 The *crtI*, *crtB* and *crtLI* genes are involved in the internal retinal biosynthesis system and  
396 responsible for beta-carotene biosynthesis. Further, the beta-carotene dioxygenase encoded by  
397 *blh* cleaves the beta-carotene to produce all-trans retinal, which could be used by MAG\_103 as  
398 its photoreactive chromophore [99]. The activity of beta-carotene dioxygenase was reported to  
399 be iron dependent [100], and the expression of the *blh* gene was slightly upregulated, but not  
400 significantly, in the iron-fertilized region. All the other four genes involved in light harvest were  
401 significantly higher expressed in the on-plateau region as compared to the off-plateau waters.

402 The extra energy supplied from light might be the reason that MAG\_103 became more  
403 competitive in stratified summer surface waters.

404 MAG\_103 and MAG\_62 differed in their potential to use and resist to antibiotics  
405 (**Supplementary Results** and **Supplementary Table 6**). While competing with other  
406 prokaryotes in the late summer surface waters when bulk abundances reached  $1.18 \times 10^9$  cells L<sup>-1</sup>  
407 (**Supplementary Table 1**), the production of antibiotics might greatly facilitate MAG\_103's  
408 dominance over other species. MAG\_103 genes involved in antibiotic production were higher  
409 expressed in the on-plateau region. The expression level of MAG\_103 genes encoding the  
410 general secretory pathway proteins (*gspC*, *gspD*, *gspE*, *gspF*, *gspG* and *gspL*), as well as the *sec*  
411 translocase system (*secA*, *secB* and *secD*), were significantly upregulated in the iron-fertilized  
412 on-plateau water. The type II secretion (T2S) pathway, coupled with Sec translocon, is regarded  
413 as the main protein secretion pathway of bacteria, which is capable of transporting a wide range  
414 of substrates, including proteases, lipases, phosphatases, carbohydrates-degrading enzymes and  
415 toxins [101]. The translocation of antibiotics produced by MAG\_103 through the outer  
416 membrane might be mainly facilitated by the T2S. Within the two-component system (TCS),  
417 MAG\_103 genes encoded a phosphate regulon sensor histidine kinase PhoR, a phosphate  
418 regulon response regulator OmpR, an osmolarity sensor histidine kinase EnvZ, and an invasion  
419 response regulator UvrY were all significantly upregulated, possibly due to stress specific  
420 responses. Further, MAG\_103 genes related to iron uptake, encoding a ferric transporter (*fbpA*),  
421 a ferredoxin (*fdx*), the TonB-ExbB-ExbD system, a bacterioferritin (*bfr*), a vitamin B12  
422 transporter (*btuB*), a bacteriorhodopsin, and the Fe-S cluster assembly proteins (*nfuA* and  
423 *sufBCD*), were also significantly higher expressed in the iron-fertilized water. The significant  
424 increase in expression levels of TCA cycle enzymes, CAZymes, more than half of the peptidases

425 [55], and the aerobic carbon-monoxide dehydrogenase subunits (*coxS* and *coxL*) indicated an  
426 enhanced carbon flux between phytoplankton and MAG\_103 represented Pseudomonadales  
427 population during the bloom decline. Although these accessory genomic features and their  
428 corresponding expression patterns are not direct evidence related to iron and carbon metabolism,  
429 the enhanced competence may facilitate the survival and growth of the microbes, whose  
430 abundances influence their roles in nutrient cycles.

431         Deciphering the many unknowns regarding the ecological roles of marine prokaryotes  
432 inhabiting the Southern Ocean, undoubtedly a region of key importance in ongoing global  
433 warming, remains profoundly challenging. Here we provide a comprehensive investigation of  
434 prokaryotic functional activities from the community level to individual taxa, targeting in situ  
435 responses linked to iron and carbon cycling. Despite remarkable shifts in community  
436 composition across contrasting nutrient regimes, we observed conservation of functional  
437 diversity through functional redundancy among community members inhabiting each ecosystem.  
438 The distinct gene expression patterns of individual taxa illustrate the link between the genetic  
439 repertoire of prokaryotic taxa and their diverse responses to the multitude of environmental  
440 factors. Our observations of a mosaic of taxonomy-specific ecological strategies in the cycling of  
441 iron and organic carbon provides insights how the habitat shapes microbial diversity in the ocean.

442 **Acknowledgements**

443 We thank B. Quéguiner, the PI of the MOBYDICK project, for providing us the opportunity to  
444 participate to this cruise, the captain and crew of the R/V Marion Dufresne for their enthusiasm  
445 and support aboard during the MOBYDICK–THEMISTO cruise  
446 (<https://doi.org/10.17600/18000403>). This work was supported by the French oceanographic  
447 fleet (“Flotte océanographique française”), the French ANR (“Agence Nationale de la  
448 Recherche”, AAPG 2017 program, MOBYDICK Project number: ANR-17-CE01-0013), the  
449 French Research program of INSU-CNRS-LEFE/CYBER (“Les enveloppes fluides et  
450 l’environnement” - “Cycles biogéochimiques, environnement et ressources”) and the  
451 Austrian FWF grant under the number P28781-B21. The authors thank the Roscoff Analyses and  
452 Bioinformatics for Marine Sciences Platform (ABiMS; <http://abims.sb-roscoff.fr/>) and the  
453 French Institute of Bioinformatics (IFB; <https://www.france-bioinformatique.fr>) for providing  
454 computational facilities and technical supports. We thank S. Blain for providing the satellite-  
455 based chlorophyll a data for the period 1998-2017. We also thank M. A. Moran and M. Landa  
456 for providing the internal standards for further *in vitro* transcription. Two reviewers provided  
457 detailed and insightful comments that helped to improve previous versions of our manuscript.

458

459 **Author Contributions**

460 I.O. conceived the project and designed the experiments. P.D. and I.O. participated to the cruise.  
461 P.D. collected the samples and carried out the nucleic acid extraction. P.D. performed  
462 metagenome read processing and contig assembly. Y.S. carried out bioinformatics analysis. Y.S.  
463 and I.O. wrote the manuscript, and P.D. provided input to the results and commented on the  
464 manuscript.

465

466 **Competing Interests**

467 The authors declare no competing interests.

468

469 **Data Availability**

470 The data sets generated and analysed during the current study are available in the European

471 Nucleotide Archive (ENA) repository at <https://www.ebi.ac.uk/ena> under the project ID

472 PRJEB37465 (metagenome) and PRJEB37466 (metatranscriptome). The metagenome reads are

473 under the accession number ERR4234198- 4234200. The metatranscriptome reads are under the

474 accession number ERR4234183-4234191. The 949,228 contigs are under the accession number

475 ERZ1694383. The 133 MAGs are under the accession number ERZ1694384-1694516.



476 **References**

- 477 1. Azam F, Malfatti F. Microbial structuring of marine ecosystems. *Nat Rev Microbiol.*  
478 2007;5(10):782-91.
- 479 2. Lechtenfeld OJ, Hertkorn N, Shen Y, Witt M, Benner R. Marine sequestration of carbon  
480 in bacterial metabolites. *Nat Commun.* 2015;6:6711.
- 481 3. Buchan A, LeClerc GR, Gulvik CA, Gonzalez JM. Master recyclers: features and  
482 functions of bacteria associated with phytoplankton blooms. *Nat Rev Microbiol.*  
483 2014;12(10):686-98.
- 484 4. DeLong E (ed). *Microbial Metagenomics, Metatranscriptomics, and Metaproteomics*, 1st  
485 edn. San Diego, CA, USA: Academic Press; 2013.
- 486 5. White III RA, Callister SJ, Moore RJ, Baker ES, Jansson JK. The past, present and future  
487 of microbiome analyses. *Nat Protoc.* 2016;11(11):2049.
- 488 6. Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JB. Beyond biogeographic  
489 patterns: processes shaping the microbial landscape. *Nat Rev Microbiol.* 2012;10(7):497-506.
- 490 7. Gilbert JA, Steele JA, Caporaso JG, Steinbruck L, Reeder J, Temperton B, et al. Defining  
491 seasonal marine microbial community dynamics. *ISME J.* 2012;6(2):298-308.
- 492 8. Needham DM, Fuhrman JA. Pronounced daily succession of phytoplankton, archaea and  
493 bacteria following a spring bloom. *Nat Microbiol.* 2016;1:16005.
- 494 9. Lindh MV, Sjostedt J, Andersson AF, Baltar F, Hugerth LW, Lundin D, et al.  
495 Disentangling seasonal bacterioplankton population dynamics by high-frequency sampling.  
496 *Environ Microbiol.* 2015;17(7):2459-76.
- 497 10. Fuhrman JA, Cram JA, Needham DM. Marine microbial community dynamics and their  
498 ecological interpretation. *Nat Rev Microbiol.* 2015;13(3):133-46.

- 499 11. Ruiz-González C, Logares R, Sebastián M, Mestre M, Rodríguez-Martínez R, Galí M, et  
500 al. Higher contribution of globally rare bacterial taxa reflects environmental transitions across  
501 the surface ocean. *Mol Ecol.* 2019;28(8):1930-45.
- 502 12. Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, et al. Ocean  
503 plankton. Structure and function of the global ocean microbiome. *Science.*  
504 2015;348(6237):1261359.
- 505 13. Ibarbalz FM, Henry N, Brandao MC, Martini S, Busseni G, Byrne H, et al. Global trends  
506 in marine plankton diversity across kingdoms of life. *Cell.* 2019;179(5):1084-97 e21.
- 507 14. Salazar G, Paoli L, Alberti A, Huerta-Cepas J, Ruscheweyh HJ, Cuenca M, et al. Gene  
508 expression changes and community turnover differentially shape the global ocean  
509 metatranscriptome. *Cell.* 2019;179(5):1068-83 e21.
- 510 15. Boeuf D, Edwards BR, Eppley JM, Hu SK, Poff KE, Romano AE, et al. Biological  
511 composition and microbial dynamics of sinking particulate organic matter at abyssal depths in  
512 the oligotrophic open ocean. *Proc Natl Acad Sci U S A.* 2019;116(24):11824-32.
- 513 16. Ghiglione JF, Galand PE, Pommier T, Pedros-Alio C, Maas EW, Bakker K, et al. Pole-  
514 to-pole biogeography of surface and deep marine bacterial communities. *Proc Natl Acad Sci U S*  
515 *A.* 2012;109(43):17633-8.
- 516 17. Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW, et al.  
517 Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci U S A.*  
518 2008;105(10):3805-10.
- 519 18. Acinas SG, Sánchez P, Salazar G, Cornejo-Castillo FM, Sebastián M, Logares R, et al.  
520 Metabolic architecture of the deep ocean microbiome. *bioRxiv.* 2019:635680.  
521 <https://doi.org/10.1101/635680>.

- 522 19. Martin JH, Gordon RM, Fitzwater SE. Iron in Antarctic waters. *Nature*.  
523 1990;345(6271):156-8.
- 524 20. Church MJ, Hutchins DA, Ducklow HW. Limitation of bacterial growth by dissolved  
525 organic matter and iron in the Southern ocean. *Appl Environ Microbiol*. 2000;66(2):455-66.
- 526 21. Obernosterer I, Fourquez M, Blain S. Fe and C co-limitation of heterotrophic bacteria in  
527 the naturally fertilized region off the Kerguelen Islands. *Biogeosciences*. 2015;12(6):1983-92.
- 528 22. Fourquez M, Obernosterer I, Blain S. A method for the use of the radiotracer <sup>55</sup>Fe for  
529 microautoradiography and CARD-FISH of natural bacterial communities. *FEMS Microbiol Lett*.  
530 2012;337(2):132-9.
- 531 23. Koedooder C, Gueneugues A, Van Geersdaële R, Vergé V, Bouget F-Y, Labreuche Y, et  
532 al. The role of the glyoxylate shunt in the acclimation to iron limitation in marine heterotrophic  
533 bacteria. *Front Mar Sci*. 2018;5:435.
- 534 24. Blain S, Tagliabue A (eds). *Iron Cycle in Oceans*, 1st edn. London, UK: ISTE Ltd and  
535 John Wiley & Sons, Inc.; 2016.
- 536 25. Dittmar T, Arnosti C. An inseparable liaison: marine microbes and nonliving organic  
537 matter. In: Gasol JM, Kirchman DL, editors. *Microbial Ecology of the Oceans*, 3rd edn.  
538 Hoboken NJ, USA: John Wiley and Sons, Inc.; 2018, pp 189–229.
- 539 26. Blain S, Queguiner B, Armand L, Belviso S, Bombled B, Bopp L, et al. Effect of natural  
540 iron fertilization on carbon sequestration in the Southern Ocean. *Nature*. 2007;446(7139):1070-4.
- 541 27. Lasbleiz M, Leblanc K, Armand LK, Christaki U, Georges C, Obernosterer I, et al.  
542 Composition of diatom communities and their contribution to plankton biomass in the naturally  
543 ironfertilized region of Kerguelen in the Southern Ocean. *FEMS Microbiol Ecol*.  
544 2016;92:fiw171.

- 545 28. Obernosterer I, Catala P, Lebaron P, West NJ. Distinct bacterial groups contribute to  
546 carbon cycling during a naturally iron fertilized phytoplankton bloom in the Southern Ocean.  
547 *Limnol Oceanogr.* 2011;56(6):2391-401.
- 548 29. Blain S, Capparos J, Guéneuguès A, Obernosterer I, Oriol L. Distributions and  
549 stoichiometry of dissolved nitrogen and phosphorus in the iron-fertilized region near Kerguelen  
550 (Southern Ocean). *Biogeosciences.* 2015;12(2):623-35.
- 551 30. d'Ovidio F, Della Penna A, Trull TW, Nencioli F, Pujol M-I, Rio M-H, et al. The  
552 biogeochemical structuring role of horizontal stirring: Lagrangian perspectives on iron delivery  
553 downstream of the Kerguelen Plateau. *Biogeosciences.* 2015;12(19):5567-81.
- 554 31. Landa M, Blain S, Christaki U, Monchy S, Obernosterer I. Shifts in bacterial community  
555 composition associated with increased carbon cycling in a mosaic of phytoplankton blooms.  
556 *ISME J.* 2016;10(1):39-50.
- 557 32. Landa M, Blain S, Harmand J, Monchy S, Rapaport A, Obernosterer I. Major changes in  
558 the composition of a Southern Ocean bacterial community in response to diatom-derived  
559 dissolved organic matter. *FEMS Microbiol Ecol.* 2018;94(8).
- 560 33. Fourquez M, Beier S, Jongmans E, Hunter R, Obernosterer I. Uptake of Leucine, chitin,  
561 and iron by prokaryotic groups during spring phytoplankton blooms induced by natural iron  
562 fertilization off Kerguelen Island (Southern Ocean). *Front Mar Sci.* 2016;3:256.
- 563 34. Debeljak P, Toulza E, Beier S, Blain S, Obernosterer I. Microbial iron metabolism as  
564 revealed by gene expression profiles in contrasted Southern Ocean regimes. *Environ Microbiol.*  
565 2019;21(7):2360-74.

- 566 35. Christaki U, Gueneugues A, Liu Y, Blain S, Catala P, Colombet J, et al. Seasonal  
567 microbial food web dynamics in contrasting Southern Ocean productivity regimes. *Limnol*  
568 *Oceanogr.* 2021;66(1):108-22.
- 569 36. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, et al. MEGAHIT v1.0: A fast  
570 and scalable metagenome assembler driven by advanced methodologies and community  
571 practices. *Methods.* 2016;102:3-11.
- 572 37. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP-a flexible pipeline for genome-  
573 resolved metagenomic data analysis. *Microbiome.* 2018;6(1):158.
- 574 38. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning  
575 metagenomic contigs by coverage and composition. *Nat Methods.* 2014;11(11):1144-6.
- 576 39. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to  
577 recover genomes from multiple metagenomic datasets. *Bioinformatics.* 2016;32(4):605-7.
- 578 40. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive  
579 binning algorithm for robust and efficient genome reconstruction from metagenome assemblies.  
580 *PeerJ.* 2019;7:e7359.
- 581 41. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the  
582 quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome*  
583 *Res.* 2015;25(7):1043-55.
- 584 42. Bendall ML, Stevens SL, Chan LK, Malfatti S, Schwientek P, Tremblay J, et al.  
585 Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. *ISME*  
586 *J.* 2016;10(7):1589-601.

587 43. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal:  
588 prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*.  
589 2010;11:119.

590 44. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, et al.  
591 eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for  
592 eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res*. 2016;44(D1):D286-93.

593 45. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, et al.  
594 Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper.  
595 *Mol Biol Evol*. 2017;34(8):2115-22.

596 46. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam  
597 protein families database in 2019. *Nucleic Acids Res*. 2019;47(D1):D427-D32.

598 47. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol*. 2011;7(10):e1002195.

599 48. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids*  
600 *Res*. 2000;28(1):27-30.

601 49. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for  
602 functional characterization of genome and metagenome sequences. *J Mol Biol*. 2016;428(4):726-  
603 31.

604 50. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al.  
605 KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score  
606 threshold. *Bioinformatics*. 2019;36:2251–2.

607 51. Saier MH, Jr., Reddy VS, Tsu BV, Ahmed MS, Li C, Moreno-Hagelsieb G. The  
608 transporter classification database (TCDB): recent advances. *Nucleic Acids Res*.  
609 2016;44(D1):D372-9.

- 610 52. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search  
611 tool. *J Mol Biol.* 1990;215(3):403-10.
- 612 53. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The  
613 Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics.  
614 *Nucleic Acids Res.* 2009;37(Database issue):D233-8.
- 615 54. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. dbCAN2: a meta server  
616 for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 2018;46(W1):W95-  
617 W101.
- 618 55. Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS  
619 database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with  
620 peptidases in the PANTHER database. *Nucleic Acids Res.* 2018;46(D1):D624-D32.
- 621 56. Garber AI, Neelson KH, Okamoto A, McAllister SM, Chan CS, Barco RA, et al. FeGenie:  
622 a comprehensive tool for the identification of iron genes and iron gene neighborhoods in genome  
623 and metagenome assemblies. *Front Microbiol.* 2020;11:37.
- 624 57. Dupont CL, Yang S, Palenik B, Bourne PE. Modern proteomes contain putative imprints  
625 of ancient shifts in trace metal geochemistry. *Proc Natl Acad Sci U S A.* 2006;103(47):17822-7.
- 626 58. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify  
627 genomes with the Genome Taxonomy Database. *Bioinformatics.* 2019;36:1925-7.
- 628 59. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative  
629 genomics. *Genome Biol.* 2019;20(1):238.
- 630 60. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7:  
631 improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772-80.

- 632 61. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated  
633 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 2009;25(15):1972-3.
- 634 62. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective  
635 stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*.  
636 2015;32(1):268-74.
- 637 63. Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for  
638 metagenomics with Kaiju. *Nat Commun*. 2016;7:11257.
- 639 64. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for  
640 assigning sequence reads to genomic features. *Bioinformatics*. 2014;30(7):923-30.
- 641 65. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*.  
642 2012;9(4):357-9.
- 643 66. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for  
644 RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
- 645 67. Rodriguez RL, Gunturu S, Tiedje JM, Cole JR, Konstantinidis KT. Nonpareil 3: fast  
646 estimation of metagenomic coverage and sequence diversity. *mSystems*. 2018;3(3).
- 647 68. Pearson WR. An introduction to sequence similarity ("homology") searching. *Curr Protoc*  
648 *Bioinformatics*. 2013;Chapter 3:Unit3 1.
- 649 69. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI  
650 analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*.  
651 2018;9(1):5114.
- 652 70. Satinsky BM, Gifford SM, Crump BC, Moran MA. Use of internal standards for  
653 quantitative metatranscriptome and metagenome analysis. *Methods Enzymol*. 2013;531:237-50.



- 654 71. Louca S, Polz MF, Mazel F, Albright MBN, Huber JA, O'Connor MI, et al. Function and  
655 functional redundancy in microbial systems. *Nat Ecol Evol.* 2018;2(6):936-43.
- 656 72. Kuhaudomlarp S, Patron NJ, Henrissat B, Rejzek M, Saalbach G, Field RA.  
657 Identification of *Euglena gracilis* beta-1,3-glucan phosphorylase and establishment of a new  
658 glycoside hydrolase (GH) family GH149. *J Biol Chem.* 2018;293(8):2865-76.
- 659 73. Ho A, Di Lonardo DP, Bodelier PL. Revisiting life strategy concepts in environmental  
660 microbial ecology. *FEMS Microbiol Ecol.* 2017;93(3).
- 661 74. Rodionov DA, Gelfand MS, Todd JD, Curson AR, Johnston AW. Computational  
662 reconstruction of iron- and manganese-responsive transcriptional networks in alpha-  
663 proteobacteria. *PLoS Comput Biol.* 2006;2(12):e163.
- 664 75. Rincon-Enriquez G, Crete P, Barras F, Py B. Biogenesis of Fe/S proteins and  
665 pathogenicity: IscR plays a key role in allowing *Erwinia chrysanthemi* to adapt to hostile  
666 conditions. *Mol Microbiol.* 2008;67(6):1257-73.
- 667 76. Py B, Barras F. Building Fe-S proteins: bacterial strategies. *Nat Rev Microbiol.*  
668 2010;8(6):436-46.
- 669 77. Zappa S, Bauer CE. Iron homeostasis in the *Rhodobacter* genus. *Adv Bot Res.*  
670 2013;66:289-326.
- 671 78. Jaggavarapu S, O'Brian MR. Differential control of *Bradyrhizobium japonicum* iron  
672 stimulon genes through variable affinity of the iron response regulator (Irr) for target gene  
673 promoters and selective loss of activator function. *Mol Microbiol.* 2014;92(3):609-24.
- 674 79. Grote J, Thrash JC, Huggett MJ, Landry ZC, Carini P, Giovannoni SJ, et al. Streamlining  
675 and core genome conservation among highly divergent members of the SAR11 clade. *mBio.*  
676 2012;3(5):e00252-12.

- 677 80. McAdams HH, Srinivasan B, Arkin AP. The evolution of genetic regulatory systems in  
678 bacteria. *Nat Rev Genet.* 2004;5(3):169-78.
- 679 81. Fourquez M, Devez A, Schaumann A, Guéneuguès A, Jouenne T, Obernosterer I, et al.  
680 Effects of iron limitation on growth and carbon metabolism in oceanic and coastal heterotrophic  
681 bacteria. *Limnol Oceanogr.* 2014;59(2):349-60.
- 682 82. Wilson DN, Nierhaus KH. The weird and wonderful world of bacterial ribosome  
683 regulation. *Crit Rev Biochem Mol Biol.* 2007;42(3):187-219.
- 684 83. Wei Y, Lee JM, Richmond C, Blattner FR, Rafalski JA, LaRossa RA. High-density  
685 microarray-mediated gene expression profiling of *Escherichia coli*. *J Bacteriol.* 2001;183(2):545-  
686 56.
- 687 84. Hendrickson EL, Liu Y, Rosas-Sandoval G, Porat I, Soll D, Whitman WB, et al. Global  
688 responses of *Methanococcus maripaludis* to specific nutrient limitations and growth rate. *J*  
689 *Bacteriol.* 2008;190(6):2198-205.
- 690 85. Gifford SM, Sharma S, Booth M, Moran MA. Expression patterns reveal niche  
691 diversification in a marine microbial assemblage. *ISME J.* 2013;7(2):281-98.
- 692 86. Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firkbank SJ, Bolam DN, et al.  
693 Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced  
694 microbiota alterations. *Cell.* 2010;141(7):1241-52.
- 695 87. Gregg KJ, Zandberg WF, Hehemann JH, Whitworth GE, Deng L, Vocadlo DJ, et al.  
696 Analysis of a new family of widely distributed metal-independent alpha-mannosidases provides  
697 unique insight into the processing of N-linked glycans. *J Biol Chem.* 2011;286(17):15586-96.

- 698 88. Matulewicz M, Cerezo A. Water-soluble sulfated polysaccharides from the red seaweed  
699 *Chaetangium fastigiatum*. Analysis of the system and the structures of the  $\alpha$ -D-(1 $\rightarrow$ 3)-linked  
700 mannans. *Carbohydr Polym.* 1987;7(2):121-32.
- 701 89. Kolender AA, Pujol CA, Damonte EB, Matulewicz MC, Cerezo AS. The system of  
702 sulfated  $\alpha$ -(1 $\rightarrow$ 3)-linked D-mannans from the red seaweed *Nothogenia fastigiata*: structures,  
703 antiherpetic and anticoagulant properties. *Carbohydr Res.* 1997;304(1):53-60.
- 704 90. Le Costaouëc T, Unamunzaga C, Mantecon L, Helbert WJAr. New structural insights  
705 into the cell-wall polysaccharide of the diatom *Phaeodactylum tricornutum*. *Algal Res.*  
706 2017;26:172-9.
- 707 91. Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, et al.  
708 Substrate-controlled succession of marine bacterioplankton populations induced by a  
709 phytoplankton bloom. *Science.* 2012;336(6081):608-11.
- 710 92. Teeling H, Fuchs BM, Bennke CM, Kruger K, Chafee M, Kappelmann L, et al.  
711 Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *Elife.*  
712 2016;5:e11888.
- 713 93. Chen J, Robb CS, Unfried F, Kappelmann L, Markert S, Song T, et al. Alpha- and beta-  
714 mannan utilization by marine Bacteroidetes. *Environ Microbiol.* 2018;20(11):4127-40.
- 715 94. Biersmith A, Benner R. Carbohydrates in phytoplankton and freshly produced dissolved  
716 organic matter. *Mar Chem.* 1998;63(1):131-44.
- 717 95. Sichert A, Corzett CH, Schechter MS, Unfried F, Markert S, Becher D, et al.  
718 *Verrucomicrobia* use hundreds of enzymes to digest the algal polysaccharide fucoidan. *Nat*  
719 *Microbiol.* 2020; 5:1026–39.

- 720 96. Beja O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, et al. Bacterial  
721 rhodopsin: evidence for a new type of phototrophy in the sea. *Science*. 2000;289(5486):1902-6.
- 722 97. DeLong EF, Beja O. The light-driven proton pump proteorhodopsin enhances bacterial  
723 survival during tough times. *PLoS Biol*. 2010;8(4):e1000359.
- 724 98. Olson DK, Yoshizawa S, Boeuf D, Iwasaki W, DeLong EF. Proteorhodopsin variability  
725 and distribution in the North Pacific Subtropical Gyre. *ISME J*. 2018;12(4):1047-60.
- 726 99. Kim SY, Waschuk SA, Brown LS, Jung KH. Screening and characterization of  
727 proteorhodopsin color-tuning mutations in *Escherichia coli* with endogenous retinal synthesis.  
728 *Biochim Biophys Acta*. 2008;1777(6):504-13.
- 729 100. von Lintig J, Vogt K. Filling the gap in vitamin A research. Molecular identification of an  
730 enzyme cleaving beta-carotene to retinal. *J Biol Chem*. 2000;275(16):11915-20.
- 731 101. Korotkov KV, Sandkvist M, Hol WG. The type II secretion system: biogenesis,  
732 molecular architecture and mechanism. *Nat Rev Microbiol*. 2012;10(5):336-51.
- 733

734 **Figure Legends**

735 **Fig. 1** Genomic features of the 133 Southern Ocean (SO) metagenome-assembled genomes  
736 (MAGs) visualized using the circlize package (v0.4.9) in R (v3.6.1). The outmost circle shows  
737 the phylogenetic tree derived from the concatenation of 218 single-copy orthologous genes  
738 shared by the MAGs. The tip labels and branches are coloured according to their taxonomic  
739 affiliations determined by GTDB-Tk [58]. The 50 MAGs whose genes were excluded from the  
740 DGE analysis due to low read counts in either metagenomes or metatranscriptomes are marked  
741 with asterisks. The second circle (“Cov.”) is a heatmap displaying the average coverage of depth  
742 (per million reads) of each MAG in each sample. The third to seventh circles are heatmaps  
743 showing the number of total transcripts (“Trans L<sup>-1</sup>”), transcripts from genes encoding ribosomal  
744 proteins (“Ribo.”), genes involved in Fe-related metabolic activities (“Fe”) and TCA cycle  
745 (“TCA”), as well as genes encoding carbohydrate-active enzymes (“CAZy”) in one liter of  
746 sampled seawaters (L<sup>-1</sup>). The value of transcripts L<sup>-1</sup> of each MAG was further normalized by the  
747 length of the MAG (Mbp). The colour schemes are given at the left bottom. The 8th circle  
748 illustrates the number of significantly differentially expressed genes at contrasting oceanic  
749 regions (on-plateau iron-fertilized vs. off-plateau HNLC waters). The orange bars represent the  
750 number of genes that are significantly higher expressed in the on-plateau M2 site, as compared to  
751 the off-plateau M3 and M4 sites. The blue bars summarize genes that are significantly higher  
752 expressed in the off-plateau HNLC waters.

753

754 **Fig. 2** Community functional diversity and taxonomic composition within functional groups.  
755 Shannon index based on the abundance of functional groups (A) and shifts in taxonomic  
756 composition within functional groups (B) across sampling sites were calculated based on the

757 community-level transcript abundance represented by the normalized per-liter transcripts  
758 estimated following the internal standards protocol [70] (**E-J** in **Supplementary Fig. 6-7**). In **B**,  
759 the relative contribution (%) of a specific taxonomic category (e.g., Gammaproteobacteria) to a  
760 functional group (e.g., ferrous iron transporter FeoA) in each sampling station was calculated  
761 (Materials and methods). Shifts in the relative contribution across stations were estimated using  
762 the ratio of the relative contribution in M2 to that in M3 (or M4) and visualized by violin plots. A  
763 ratio value less than 1 indicates that the taxonomic category accounts for a larger share of the  
764 transcripts ( $L^{-1}$ ) of a functional group in the off-plateau HNLC waters, and vice versa. Multiple  
765 databases were considered, including CAZy, FeGenie, KEGG, Pfam, Superfamily and TCDB.  
766 Five dominant taxonomic groups in gene pool and transcript inventories across all sampling sites  
767 were shown. Colour code is the same as **Supplementary Fig. 6-7**. Only functional groups  
768 consisting of at least 50 genes, out of the 3,003,586 protein-coding genes predicted from the  
769 metagenome assemblies, were used in the calculation.

770

771 **Fig. 3** Statistics of significantly differentially expressed genes (SDEGs) involved in glycoside  
772 hydrolysis and key iron metabolic pathways. Panels from top to bottom represent glycoside  
773 hydrolase (GH), iron uptake regulators (Reg.), ferrous uptake ( $Fe^{2+}$ ), ferric uptake ( $Fe^{3+}$ ),  
774 siderophore biosynthesis and uptake (Sid. Syn./Upt.), heme uptake (Heme), iron storage (Sto.)  
775 and Ferredoxin/Flavodoxin switch (F/F). The full list of KEGG Orthology groups (KOs) related  
776 to iron metabolism examined in this study could be found in **Supplementary Table 4**. Each row  
777 represents one functional group. The two vertical panels show statistics of the SDEGs based on  
778 the metagenome-normalized metatranscriptomic pseudo counts and the corresponding log<sub>2</sub>-  
779 based fold changes. In the bi-direction bar plots, the bars pointing to the left indicate the number

780 of genes that are significantly higher expressed in the on-plateau iron-fertilized M2 site, as  
781 compared to the off-plateau HNLC M3 and M4 sites. To the contrary, the bars pointing to the  
782 right represent genes that are significantly higher expressed in the off-plateau HNLC waters. The  
783 colour scheme of taxonomy is shown on top.

784

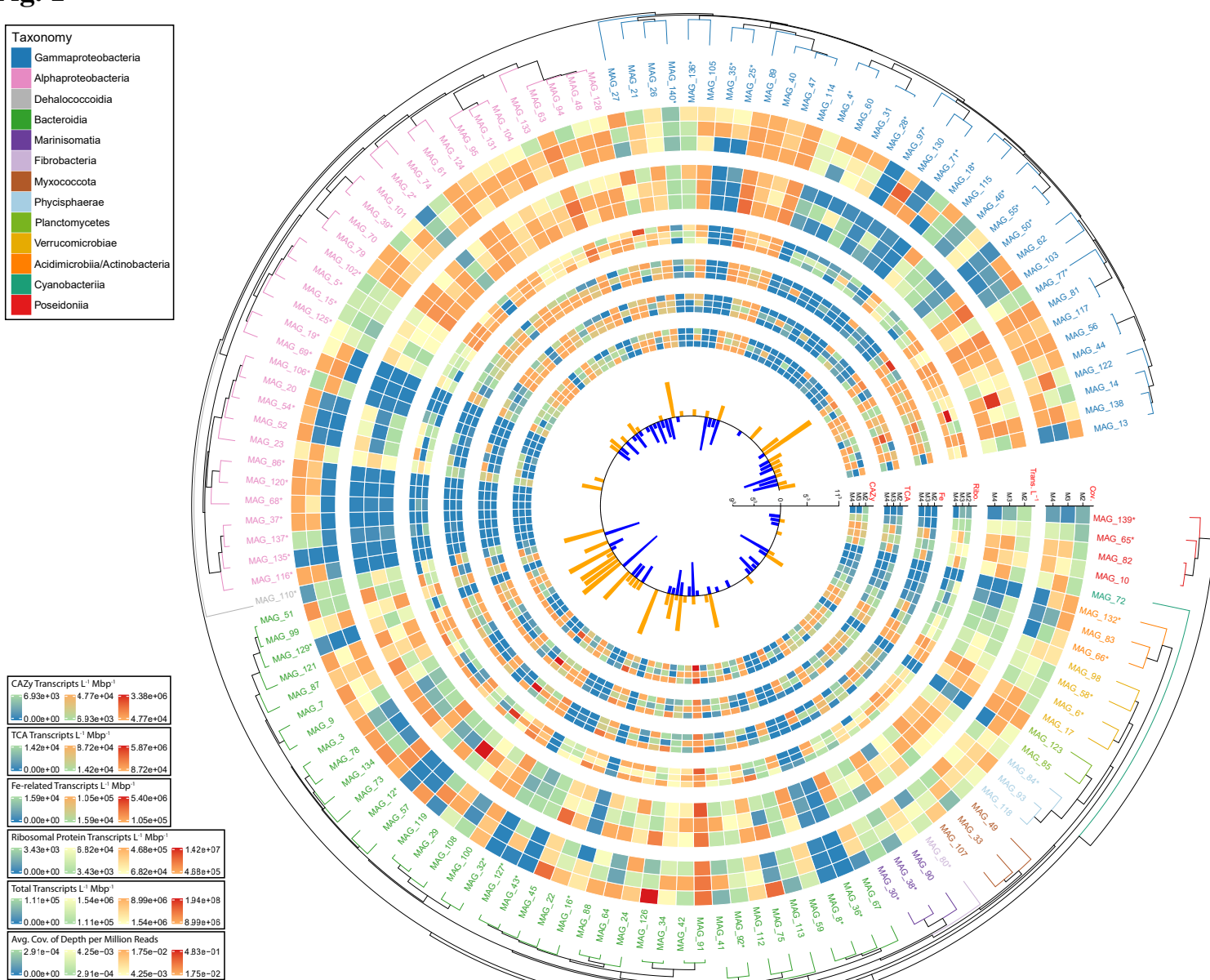
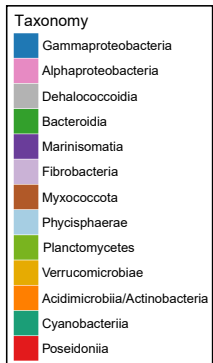
785 **Fig. 4** Ratios of ribosomal-protein versus all transcripts ( $L^{-1} \text{ Mbp}^{-1}$ ) from 133 MAGs. Lines  
786 were calculated from Model II linear regression analyses. The corresponding formulas and R-  
787 squared measures are shown at the bottom.

788

789 **Fig. 5** The distribution of significantly differentially expressed genes (SDEGs) in the MAGs  
790 among diverse functional categories related to iron uptake and carbon metabolism. Only 47  
791 MAGs with SDEGs are shown here. From left to right, the panels represent the phylogenetic tree  
792 (the same as shown in **Fig. 1**), the iron-related KEGG Orthology groups (KOs), the KOs  
793 involved in the tricarboxylic acid (TCA) cycle, and carbohydrate-active enzymes (CAZymes).  
794 Each square block describes the statistics of a protein family in a MAG. An empty square  
795 suggests that no genes in the MAG (y axis) are classified into the corresponding functional group  
796 (x axis). A circle in the square block indicates the identification of homologs to a protein family  
797 in the MAG, with its size proportional to the number of genes assigned to that family. The square  
798 blocks are coloured according to the differential expression patterns of its gene(s). As illustrated  
799 in **Fig. 1**, genes, which are significantly higher expressed in the iron-fertilized site M2 as  
800 compared to the HNLC M3 and M4 sites, are highlighted in orange; vice versa, in blue. Given  
801 that genes belonging to the same functional group might not be synchronized in their expression  
802 patterns, the transparency of each square block shows the percentage of genes that are

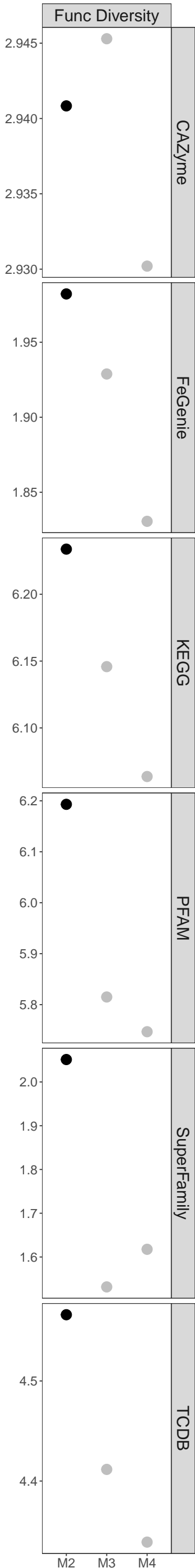
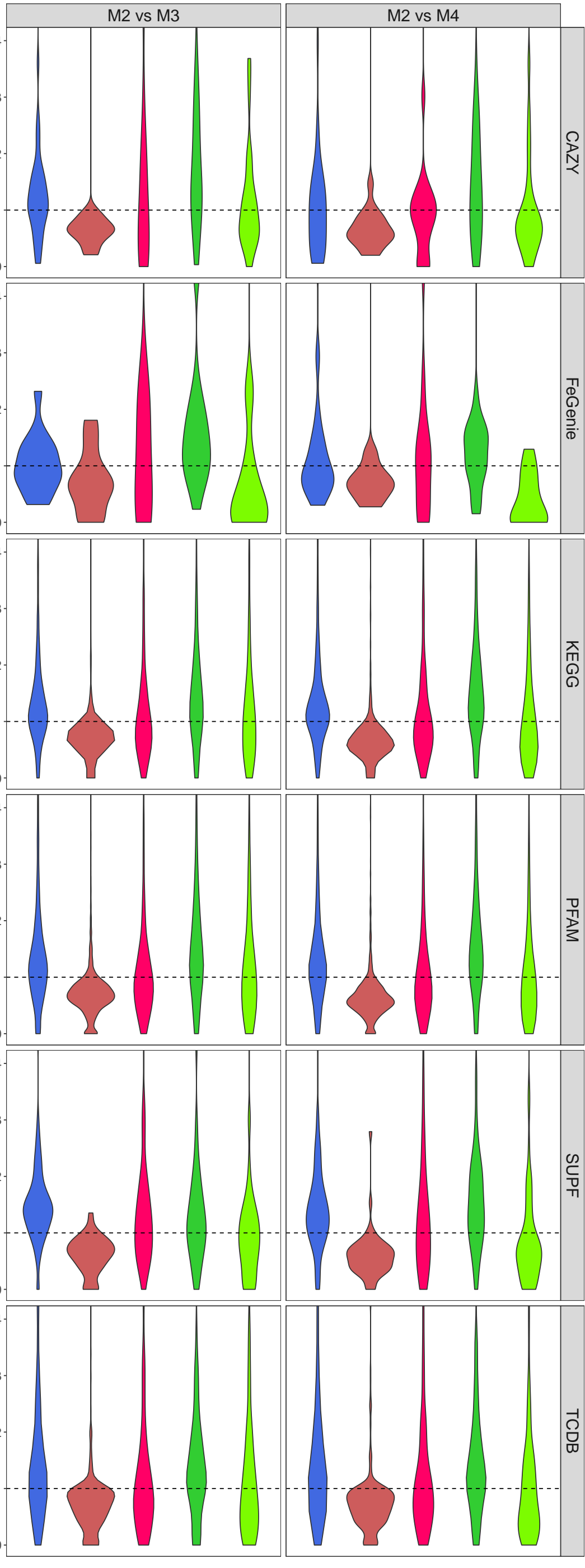
803 significantly differentially expressed. We have not detected protein families whose genes were  
804 significantly shifting their expression levels in opposite directions (e.g., parts of the genes in the  
805 same family significantly upregulate their expression levels whilst others significantly  
806 downregulate theirs). The KO “K00240” (marked with an asterisk) is shown twice, because it is  
807 a Fe-S protein family and also involved in the TCA cycle. Among the pathways involved in  
808 carbon metabolism, KOs shared by multiple pathways are only shown once. All the information  
809 illustrated in this graph is summarized based on the differential expression analysis performed on  
810 the metagenome-normalized metatranscriptomic profile (see **Methods**).



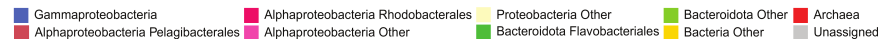
**Fig. 1**

**Fig. 2**

■ Alphaproteobacteria Pelagibacterales    ■ Alphaproteobacteria Rhodobacterales  
■ Gammaproteobacteria    ■ Bacteroidota Flavobacteriales    ■ Bacteroidota Other

**A. Diversity****B. Ratios between sample pairs**

# Fig. 3



**Fig. 4**