Controls of SAR11 subclade abundance, diversity, and growth in two Mid-Atlantic estuaries

Running Title: SAR11 in two Mid-Atlantic estuaries

Barbara J. Campbell[1#], Shen Jean Lim[1*], David L. Kirchman[2]

[1]Department of Biological Sciences, Clemson University, Clemson SC, 29634

[2]School of Marine Science and Policy, University of Delaware, Lewes, DE 19958

[#]corresponding author: bcampb7@clemson.edu

*current address: University of Miami Cooperative Institute of Marine and Atmospheric Studies, 4600 Rickenbacker Cswy, Key Biscayne, FL 33149

Abstract word count: 244

Importance word count: 136

Text word count: 4915 (without methods, references, and figure legends)

Key words: growth rate, transcriptomics, Metagenome assembled genome, estuarine, nutrients, genomospecies

Abstract

SAR11 is a dominant bacterial clade in marine oligotrophic ecosystems. SAR11 can also be dominant in estuarine systems, where they are not well-studied. We examined the effects of season, nutrient concentrations, and salinity in shaping SAR11 subclade abundance, diversity, function, and growth in two Mid-Atlantic estuaries, the Delaware and Chesapeake Bays. Using metagenome-assembled genomes, we identified twelve distinct genomospecies within the Ia, II, IIIa, and V subclades, which made up to 60% of the total bacterial community. The functional potential of all SAR11 genomospecies varied, especially in carbohydrate metabolism, transporters, and one-carbon metabolic pathways. Predicted growth rates, estimated by the Peak to Trough method, varied by season and genomospecies. SAR11 growth rates negatively correlated in the spring but positively correlated in the summer with chlorophyll *a* concentrations and bacterial production, as well as phosphate and ammonium concentrations. Genomospecies in Ia.1, IIIa.2, and IIIa.4 subclades had low growth rates, while genomospecies in Ia.3, Ia.5, Ia.6, II, and V subclades had higher and more variable growth rates that were positively correlated with phosphate concentrations and temperature. Growth rate variation between subclades was associated with carbohydrate metabolic gene repertoires, especially glycolysis and number of transporters. While total transcript to genome ratios generally mirrored growth rates, transcription of genes involved in phosphate and nitrogen transport were negatively associated with growth rates. These data suggest that SAR11 genomospecies abundance varies in these estuaries because of differences in growth rates and metabolic capacities in response to changes in environmental conditions.

Importance

The SAR11 clade is one of the most abundant bacterial groups in marine systems, including many estuaries. From the Delaware Bay and Chesapeake Bay environmental metagenomes, we reconstructed nearly complete SAR11 metagenome-assembled genomes representing ten

genomospecies in four subclades, of which at least one is novel. Growth rate estimates of genomospecies correlated with functional gene repertoires of carbohydrate transporter and metabolism. Different SAR11 genomospecies dominated among the seasons, depending on their growth rates, biological productivity, and nutrient concentrations. Our RNAseq approach facilitated an understanding of the environmental controls on the abundance of SAR11 genomospecies in their natural habitat. This study is the first to combine multiple measures of diversity, abundance, functional potential, growth rates and activity of this important group, demonstrating a direct link between SAR11 genomospecies abundance and growth in the context of its environment.

Introduction

The SAR11 clade of bacteria (*Pelagibacterales*) represents about 25% of all marine planktonic cells in the stratified, oligotrophic open ocean (1, 2). The clade comprises at least nine oceanic subclades (Ia.1, Ia.3, Ib, Ic, IIa, IIb, IIIa, IV and V) and one freshwater subclade (IIIb) (3-5). Although typically less than 5% of the microbial community in freshwater systems, SAR11 members unexpectedly may comprise up to 60% of free-living bacteria in estuaries (6). Estuaries typically are hot spots of production with nutrient concentrations ten-fold higher or more than the open ocean. Various SAR11 subclades such as Ia.1, II and IIIa are found in estuaries including the Delaware Bay and some taxa are estuary-specific (2, 7-10). Nevertheless, SAR11 subclade dominance in estuaries is not universal and may depend on season and salinity (6, 7, 9, 11).

SAR11 are aerobic heterotrophs with small, streamlined genomes that are diverse across subclades (12, 13). Analyses of seven published genomes in four marine SAR11 subclades (Ia.1, Ia.3, IIIa, V) revealed common genomic features, including low G+C content, low numbers of paralogs, a large hypervariable region, and high gene synteny, as well as core genes that comprised about 50% of their genomes (13). Transport genes make up approximately 15% of these genomes, which is around twice

48 that of other bacteria, including the marine *Roseobacter* clade from the same alphaproteobacterial class
49 (13-15). Core SAR11 genes encode ATP-binding cassette (ABC) transport systems for amino acids, iron
50 transport, and tripartite ATP-independent periplasmic (TRAP) transporter(s) for C4 compounds. Other
51 common genes participate in photoheterotrophy (13, 16), electron transport, and the TCA cycle (13).

52 Despite their commonalities, several important distinctions exist among SAR11 subclades. Their
53 abundance differs among geographic locations (2, 5) and with different environmental parameters,
54 including light (17-19), temperature, and nutrient concentrations, which also vary with depth (17, 18, 20,
55 21). Accordingly, SAR11 subclades and genomes differ in the presence of ecologically important genes,
56 such as those involved in the transport of organic carbon, sulfur, and phosphorus, as well as the
57 oxidation of carbon (e.g., C1) compounds (13, 22). However, the number of sequenced genomes used to
58 make these observations is low compared to the abundance and diversity of the SAR11 clade. There is
59 little known about the genomic makeup of several subclades, as well as SAR11 subclade distribution in
60 estuaries.

61 In addition to the lack of information about SAR11 subclade ecogenomics in estuaries, the
62 differences in SAR11 activity or growth between subclades is not well studied. Some of this is because of
63 the paucity of cultured SAR11 representatives (3). The slow growth rate of *P. ubique* and relatives in the
64 laboratory, averaging about 0.4 per day (12, 23-25), raises questions about growth rates and activity of
65 this group in the ocean. Previous estimates by our group using 16S rRNA gene:rRNA ratios in the coastal
66 Mid-Atlantic suggest that about half of the SAR11 taxa had lower growth rates than average, whereas
67 the other half had average or above average growth rates, which varied by season (26). Differential use
68 of glucose and phosphate was also observed within the SAR11 clade, likely contributing to differences in
69 ocean productivity (27, 28) and perhaps also to growth rate (29). In addition, data from a technique
70 combining microautoradiography and fluorescence in situ hybridization (MAR-FISH) suggest that SAR11
71 bacteria are in fact growing as fast, if not faster than other bacteria in surface waters of the North

72 Atlantic (30, 31). However, oceanographic studies found that global SAR11 gene transcription was
73 generally lower than expected based on abundance (32) and varied in response to light and the activity
74 of other microbes in oligotrophic open oceans (33-35).

75 Estuaries are ideal environments to study the abundance, activity, and functional dynamics of
76 SAR11 subclades since these ecosystems predictably vary in time or space. The goal of this study was to
77 determine SAR11 taxonomic, genomic, and functional diversity as well as growth and activity along
78 environmental gradients in two contrasting but geographically close estuaries, the Delaware and
79 Chesapeake Bays. The Delaware Bay has a steep salinity gradient, is well-mixed with little to no
80 stratification. Primary production is limited by light in the winter, phosphorus in the late spring, and
81 nitrogen in the summer (36). In contrast, the Chesapeake Bay is highly stratified, especially in the
82 summer, has higher primary productivity, and contains much more mesohaline waters because it is less
83 influenced by tides due to its length (37-39). Similar to the Delaware Bay, primary productivity in the
84 Chesapeake varies by light, phosphorus, and nitrogen, depending on the season and location within the
85 estuary (40). We hypothesized that these environmental gradients would structure and influence the
86 abundance, taxonomic, and functional variability of SAR11 subclades as well as their growth and activity.
87 To test this hypothesis, we utilized medium and high-quality assembled MAGs and reference genomes
88 to examine the taxonomic relationships, abundances, and functional potentials of the *Pelagibacterales*
89 in these estuaries. We also investigated differences in genomospecies growth and activity analyzing
90 differences in metagenomes (MG) and metatranscriptomes (MT) between environmental conditions.

91

92 **Results**

93 *MAG characteristics*

94 Forty-four MAGs classified as *Pelagibacterales* were identified from 36 metagenome (MG)
95 samples from five cruises to the Delaware or Chesapeake Bays in 2014 and 2015 (41) (**Table S1**). They
96 were, on average, 88.6% complete with 0.4% redundancy. Only one MAG contained partial 16S or 23S
97 rRNA genes. Based on phylogenomic and AAI analyses, these MAGs clustered into four larger clades (Ia,
98 II, IIIa and V) with other known *Pelagibacterales* genomes, SAGs, or MAGs (**Fig. 1, see Fig. S1 at**
99 **10.6084/m9.figshare.19701379**). Within Ia, five different genomospecies (Ia.1, Ia.3, Ia.4, Ia.,5, and Ia.6)
100 were identified. Most genomospecies had representative MAGs in the Delaware and Chesapeake Bays.
101 We did not recover MAGs related to 1a.4, Ib, Ic, IIa.B or IIIb from the two bays. Novel to this study are
102 well-represented genomospecies from II and Ia.6, as well as a much-expanded IIIa (genomospecies
103 IIIa.4). The subclades are examined in more detail in the supplemental text.

104 *SAR11 abundance in the Chesapeake and Delaware bays*

105 To estimate the relative abundance of SAR11 genomospecies in our samples, we mapped each
106 MG and metatranscriptome (MT) to SAR11 MAGs and genomes dereplicated at a threshold of 95%
107 average nucleotide identity (ANI). We then compared their predicted abundances to Kaiju-based
108 estimates of total *Pelagibacterales* or bacteria in their corresponding MG and MT. *Pelagibacterales*
109 dominated the microbial communities from mid- and high-salinity < 0.8 µm samples (21-61%
110 abundance) and were less abundant in the > 0.8 µm samples (8-25% abundance) (**Fig. 2a, see Fig. S2 at**
111 **10.6084/m9.figshare.19701460**). The relative abundance of the genomospecies varied extensively and
112 is detailed in the supplemental text. Briefly, SAR11 genomospecies Ia.1 and IIIa.2 were most abundant in
113 the spring, genomospecies Ia.1 and IIIa.4 dominated Chesapeake spring samples, while genomospecies
114 II, Ia.5, Ia.6, IIIa.4 and V dominated in summer in both bays (**Fig. 2b**). The average ratio of
115 Pelagibacterales reads from recovered MAGs to total Pelagibacterales reads in their corresponding
116 metagenomes was 75% and ranged from 60-95% in non-freshwater samples (**Fig. 2c**).

117 *Pangenome characteristics*

118 Since SAR11 genomes/MAGs from these estuaries have not been comprehensively compared,
119 we used Anvi'o to perform pangenome analyses of gene clusters, annotated KO (KEGG) genes, and COG
120 genes. From 104 SAR11 reference genomes and medium to high-quality MAGs (>70% completeness and
121 <5% contamination) (**Table S1**), 13,457 gene clusters were identified (see **SI at**
122 **10.6084/m9.figshare.19701478**). No gene cluster was shared between all 104 genomes, likely due to
123 the incompleteness of the MAGs. However, 107 gene clusters, 254 KEGG Orthology (KO) terms and 349
124 clusters of orthologous genes (COGs) present in at least 50% of the genomes within a genomospecies
125 were shared between genomospecies.

126 Since transporters make up a large percentage of SAR11 genomes, we next assessed their
127 differences between subclades. We identified 97 transporter genes based on Pfam annotations (**Fig. S3,**
128 see **SI at 10.6084/m9.figshare.19701478**). Transporter genes classified into 18 COG categories and their
129 abundance varied by subclade. The number of genes in the G (carbohydrate transport and metabolism)
130 COG category in each subclade varied in the order of highest to lowest number: V, Ia.3, II, 1a.5/Ia.6, Ia.1,
131 respectively. The abundance of the ABC-type proline/glycine betaine transport system (*pro*VWX) was
132 most represented in subclades Ia.3, Ia.5 and V (**Fig. 3**). Both ABC and TRAP transporters for
133 carbohydrates were well represented in most subclades (**Fig. 3,** see **SI at**
134 **10.6084/m9.figshare.19701478**), although there were some notable differences. Most subclades had
135 variable numbers of genes in the ABC-type glycerol-3-phosphate (*ugp*ABE) and ABC-type sugar transport
136 systems (e.g., *mal*K). The supplemental text discusses other transporter genes in more detail.

137 As with transporter genes, carbohydrate metabolism genes were differentially present between
138 subclades and genomospecies. These included genes in the classical Embden-Meyerhof-Parnas (EMP)
139 glycolysis pathway, Entner–Doudoroff (ED), a variant of ED pathway, Pentose Phosphate Pathway (PPP),

140 tricarboxylic acid (TCA) cycle, and the glyoxylate shunt, which are known to differ between
141 representative SAR11 genomospecies (13, 28). The two V and all II MAGs from the Delaware and
142 Chesapeake, as well as the two IIIb genomes, contained all but one of the genes in the EMP pathway.
143 None contained genes for the full ED pathway. However, MAGS from genomospecies Ia.5 and II and NP1
144 (Ia.1) contained all but one of the expected genes in the variant ED described in *Pelagibacter ubique*
145 HTCC1062, and most genes were on the same contig (**Fig. 4, see SI at 10.6084/m9.figshare.19701478**).
146 Ia.1 MAGs from the Delaware and Chesapeake were missing at least three genes previously described in
147 the variant of the ED pathway, and the gene content and synteny was identical to that of HTCC1016.
148 Members of the 1a.3, 1a.6, and IIIa subclades were missing 4 of 11 genes in this pathway (**see SI at**
149 **10.6084/m9.figshare.19701478**). Only MAGs and genomes within subclades IIIa, IIIb and V contained all
150 the genes in the parallel glycolysis PPP. We next explored the genetic potential of SAR11 genomospecies
151 for utilizing the TCA cycle or the alternate glyoxylate shunt. Almost half of the MAGs in the Ia.I subclade
152 had identical glyoxylate shunt genes present in reference Ia.1 genomes. The two V MAGs contained
153 most or all genes in the TCA cycle but not the glyoxylate shunt. Most representatives in IIIa and IIIb
154 contained most or all genes in the TCA cycle and glyoxylate shunt. However, MAGs within II and Ia.3
155 genomospecies from the Delaware and Chesapeake bays were missing genes in glycolysis, the TCA cycle
156 and the glyoxylate shunt, possibly due to incomplete genomes. Additionally, MAGs in genomospecies II,
157 IIIa.2 and IIIa.4 contained, phasin and *pha*C genes, involved in carbon storage via polyhydroxyalkanoate
158 (PHA) metabolism, used when there is an imbalance between carbon and nutrients in several bacteria
159 (42). Across genomospecies, we also observed a clear difference between the presence of genes in
160 many known C1 metabolism pathways in SAR11 (13, 22). The Ia.1 and V subclades contained most
161 genes, except for carbon monoxide oxidation genes (*cut*MLS-*cox*G), which were present in some Ia.3,
162 Ia.4, Ia.5, Ia.6, Ib, Ic, and IIIa subclades (**Fig. S4, see SI at 10.6084/m9.figshare.19701478**). Further details
163 of the SAR11 pangenome is discussed in the supplemental text.

164 *Growth estimates*

165 We used two independent methods to gain an understanding of the microbial growth rates
166 across the representative MAGs/genomes of the genomospecies in our dataset. First, estimates of
167 maximal growth rates for representatives based on codon usage patterns were below the minimum
168 values in gRodon of doubling times > 5 hours, a known bias for slow growing organisms (43). Second,
169 patterns of sequence coverage across the representative MAGs as assessed by CoPTR Peak to Trough
170 ratios that reflect actual growth rates (44), varied between samples and representative genomospecies
171 (**Table S2**). To broadly look at the entire SAR11 community in relation to transcript to genome ratios and
172 nutrient concentrations, we first averaged CoPTR values in a sample. Average CoPTR values of all
173 genomospecies within each metagenomic sample were correlated ($r = 0.93$, $n = 11$, $p < 0.001$) with the
174 independently derived ratios of the percentage of total *Pelagibacterales* in the metatranscriptome in
175 relation to the metagenome (MT/MG) from each sample (**Fig. 2c**). Additionally, both average CoPTR
176 values and MT/MG were negatively or positively correlated with chlorophyll a, phosphate and
177 ammonium in the spring and summer, respectively (**Figs. 5a-b, S5-S6**). Average CoPTR values were
178 negatively correlated with nitrate in the spring, but not in the summer. We next evaluated the
179 relationship between all CoPTR values and environmental variables (**Table S2, Figs. 5c-f, S7a**). All CoPTR
180 values were significantly correlated with bay (Delaware>Chesapeake) and season (fall+summer>spring)
181 ($p<0.05$, Wilcoxon rank), but were not with salinity. SAR11 genomospecies clustered into two separate
182 groups based on their CoPTR values (**Fig. 6**). Group 1, with low, relatively constant values, included
183 representatives from 1a.1, IIIa.2 and IIIa.4, while group 2, with higher and more variable values, included
184 representatives from Ia.3, Ia.5, Ia.6, II and V. Average CoPTR values from multiple representatives within
185 a genomospecies did not vary significantly (**Fig. S7b**). When the representative MAG CoPTR values were
186 split into the two groups described above, none were significantly correlated with any Chesapeake
187 environmental variable and only the second group, with representatives from genomospecies Ia.3, Ia.5,

188 Ia.6, II and V, had values that were either positively (phosphate and silicate) or negatively (salinity)
189 correlated with variables from the Delaware (**Fig. 5, S7**).

190 *Activity estimates*

191 We first examined activity by looking at the relationship between the relative abundances of
192 representative MAGs in MTs and MGs (**Fig. 7, see Fig. S8 at 10.6084/m9.figshare.19701463**).
193 Representatives from genomospecies Ia.1 (NP1 only), Ia.3, Ia.5, and V had MT/MG ratios significantly
194 greater than one, suggesting that they were more active than expected while the opposite was the case
195 for IIIa.4 ($p<0.02$, Wilcon-signed rank test). The ratios of the other genomospecies were not significantly
196 different.

197 Next, differences in transcription between conditions were evaluated from MT sequences
198 mapped to gene clusters of representative MAGs listed in **Table S1.** Most of the observed variation was
199 in the first axis (92%), which separated spring from summer samples (**Fig. 8**). The rest of the variation
200 was related to bay and then salinity, especially with the Chesapeake MTs. We next compared transcript
201 profiles from the representative MAGs between various conditions to examine whether there were
202 commonalities in gene expression changes. The total percentage of genes from each MAG that either
203 increased or decreased in each comparison ranged from 3.6-25.3% and averaged 11.6% (**see Table S3 at**
204 **10.6084/m9.figshare.19701457**). Gene expression differences due to salinity, especially in the
205 Chesapeake, are detailed in the supplemental text (see **Fig. S9 at 10.6084/m9.figshare.19701469**).

206 Because we saw the most difference in growth rates of MAGs between the two bays, we
207 concentrated on activity differences between the Chesapeake and Delaware, controlling for salinity and
208 season. We compared the transcript counts mapped to representatives from genomospecies within
209 subclades Ia.3, Ia.5, Ia.6, II, IIIa.4 and V between the Delaware and Chesapeake in the summer,
210 separating medium salinity samples from high salinity samples to control for salinity (**Fig. 9, see SI at**

211 **10.6084/m9.figshare.19701484**). Expression of most genes in COG categories C (energy production and
212 conversion), J (translation), O (post-translational modification, protein turnover, chaperone functions)
213 and S (function unknown), including ribosomal protein genes, energy conservation genes and co-
214 chaperonin genes, was higher in the Chesapeake than in the Delaware. In contrast, genes within the
215 COG categories for amino acid transport and metabolism (E) and secondary metabolites biosynthesis,
216 transport, and catabolism (Q) were generally upregulated in the Delaware compared to the Chesapeake.
217 In general, expression of phosphate transport and regulatory protein encoding genes (*pst*SCBA, *pho*BU)
218 and the ammonia channel protein encoding gene *amt*B was higher in the Chesapeake than in the
219 Delaware, except for the phosphate transporters in genomospecies Ia.5, where the opposite was true.
220 Expression of other genes previously associated with nitrogen limitation (45) besides *amt*B, namely *tau*A
221 (taurine transport), *yhd*W (amino acid transport), and *gln*A, were also increased in the Chesapeake
222 compared to the Delaware in the Ia.3 representative MAG (**see SI at 10.6084/m9.figshare.19701484**).
223 While ammonium concentrations were much less in the Chesapeake compared to the high salinity
224 waters in the Delaware, there was not much difference in the concentrations from the medium salinity
225 samples (**Fig. S5**). Representative MAGs from genomospecies II and IIIa.4 had higher *pha*C and phasin
226 gene expression in Chesapeake than in the Delaware. No differentially expressed genes in the glycolysis
227 or the pentose phosphate pathway were observed from representative MAGs between the bays.
228 However, expression of several TCA cycle genes was higher in the Chesapeake than in the Delaware in
229 the V representative. Similar gene expression patterns in the Delaware summer L08 samples were
230 observed when comparing high salinity and medium day salinity samples to the medium night salinity
231 samples. These samples also had differences in growth rates and phosphate levels and are detailed in
232 the supplemental text (**Fig. S10**).

233 **Discussion**

234 Few estimates of growth rates of taxonomic groups below order level directly in natural
235 environments exist beyond a few studies using MAR-FISH or rRNA:rRNA gene ratios (6, 25, 46, 47). Here,
236 we found that growth rates of *Pelagibacterales* varied between and within genomospecies depending
237 on environmental conditions, seasons, and bays. No other study has estimated SAR11 genomospecies *in*
238 *situ* growth rates and compared them to environmental factors. Importantly, the average growth rate of
239 genomospecies that dominated in the spring was negatively correlated with productivity measures and
240 many environmental factors as one would expect with bacteria that dominate oligotrophic
241 environments. However, the average growth rate of genomospecies that dominated in the summer was
242 positively correlated with those same measures indicating different controls of growth rates. Phosphate
243 concentrations and temperature, as well as transcripts involved in the cycling of phosphate and nitrogen
244 compounds, were also positively and negatively correlated, respectively, to growth rates within the
245 genomospecies that dominated summer samples. In addition to nutrients, variations in carbohydrate
246 gene repertoires also helped explain differences in growth rates between these genomospecies. Finally,
247 some genomospecies may be more suited to estuaries than open ocean systems. For example, at least
248 one genomospecies in subclade II was not previously described and others in Ia.6 and IIIa not well
249 represented in previous SAR11 ecogenomics studies done in oceanic waters and were found in estuaries
250 (7, 48-50).

251 Past studies demonstrated generally low ratios of SAR11 rRNA to rRNA genes in coastal and
252 estuarine environments (6, 26, 51-53), with some potential variation in estimated growth rates between
253 seasons or salinities (6, 26). Others have developed algorithms using metagenomics or genomics to
254 measure maximum potential growth rates, based on codon usage, as well as estimates of actual growth
255 rates, based on differences between read depth at the origin of replication versus the terminus (PTR)
256 (43, 44, 54). These would be ideal to use in environmental settings because they require input MAGs
257 with high completeness and low contamination. However, both maximal potential growth rates and

258 some measures of PTR are not sensitive, especially with slow growing organisms like those in the SAR11
259 clade (43). We assembled SAR11 MAGs with high completeness and low contamination using many
260 metagenomic samples for use in our growth rate estimates (41). Our average estimates of growth rate
261 for SAR11 Ia.1 MAGs was about 0.5 per day using CoPTR (44), similar to estimates of *P. ubique* in culture
262 (~0.4 per day) (25, 55) and of the total clade in situ, which ranged from 0.1-1.8 per day (31, 56).
263 Additionally, we found that our average PTR estimates per sample also correlated well with average
264 ratios of mRNA transcripts per genome reads. The results parallel what was found with total RNA:DNA
265 or rRNA:rRNA genes in many previous culture-based studies (57, 58), but were different from a previous
266 study with *P. ubique* (25).
267 SAR11 is typically known as an oligotroph that thrives in low nutrient conditions. SAR11
268 subclade abundances vary under different environmental conditions (20, 21, 59). Our results from the
269 spring indicate SAR11 has an oligotrophic lifestyle in both estuaries, where the average growth rates and
270 transcripts per genome reads both negatively correlated with chlorophyll a and nutrient concentrations.
271 As expected from previous studies in temperate coastal waters, slow-growing members of
272 genomospecies in the Ia.1 subclade dominated both the total microbial community and the
273 *Pelagibacterales* in spring (18, 20). However, we found the opposite was true in the summer, where
274 average SAR11 growth rates in the two estuaries were positively correlated with chlorophyll a,
275 phosphate, and ammonium, but negatively correlated with nitrate. The summer estuarine samples were
276 dominated by members of Ia.3, Ia.5, Ia.6, II, IIIa.4 and V. Earlier studies indicated that some SAR11
277 subclades prefer environments with higher nutrients. For instance, subclade II members dominated
278 mesopelagic zones at the Bermuda Time series station two months after mixing events, while many Ia.5
279 and IIIa populations dominate more eutrophic estuarine environments (6, 9, 18, 20, 49, 50). While the
280 differences in genomospecies presence might be expected under different nutrient regimes, our results

281 indicate that growth rates are also differentially affected by environmental conditions, at least in these
282 estuaries.
283 We expected slower growth rates in the spring versus the summer based solely on temperature
284 effects (60, 61). However, we did not expect SAR11 growth rates to be slower in the Chesapeake than in
285 the Delaware. This difference could be due to the amounts and types of carbon available in the
286 Delaware compared to the Chesapeake from late summer phytoplankton blooms of mixed eukaryotes,
287 especially diatoms. Our data suggest eukaryotic phytoplankton blooms were more of a factor in the
288 Delaware than the Chesapeake; chlorophyll a concentrations in the mid-salinity samples from the
289 Delaware were two times higher than that in the Chesapeake while Cyanobacteria were much more
290 abundant in the Chesapeake than in the Delaware. Additionally, carbohydrate transporter genes from
291 representative SAR11 MAGs were differentially expressed between the two bays. These genes encoded
292 proteins in the sn-glycerol 3-phosphate transport system, as well as other ABC and TRAP-type
293 carbohydrate transporters. These results suggest that the types of carbon utilized by these different
294 genomospecies varied with environmental conditions.
295 SAR11 genomospecies also differed in their carbohydrate transporter and metabolism genetic
296 repertoires that were related to growth rate differences. The total number of carbohydrate transporter
297 and metabolism genes, when corrected for genome size, was related to the average growth rate of the
298 representative MAG. This suggests that SAR11 genomes with more access to carbohydrates can
299 replicate faster. Furthermore, MAGs in genomospecies Ia.1 and IIIa had significantly lower growth rates
300 than MAGs in genomospecies Ia.3, Ia.5, II and V. The MAGs in the former group lack genes in the typical
301 EMP and ED glycolysis pathways and the alternate ED pathway found in coastal Ia.1 isolates (28), but
302 had a gene content and synteny identical to HTCC1016, isolated from coastal Oregon (23). However,
303 MAGs in the latter genomospecies have most or all the alternative ED pathway (Ia.3, Ia.5, II) or most
304 genes in the EMP pathway (II and V). The lack of an alternative ED pathway in our estuarine/coastal Ia.1

305 MAGs is surprising, given the pathway was found in our NP1 Ia.1 representative. Previous studies
306 indicated that this pathway was much more prevalent in coastal SAR11 than those found in the open
307 ocean (28).

308 Growth rates of individual SAR11 genomospecies estimated from CoPTR and activities estimated
309 from MT/MG ratios generally agreed with each other, especially for Ia.3, Ia.5, IIIa.4, and V, that were
310 dominant in the summer samples. Ia.1 and IIIa.2, both abundant in the spring, had higher MT/MG ratios
311 compared to CoPTR values. In contrast, representatives from Ia.6 and II dominant in the summer
312 samples, had lower ratios than expected compared to CoPTR values. Differences in phage infection,
313 defense mechanisms, or gene expression of highly abundant transcripts such as proteorhodopsin or
314 different transporters may explain some of the unexpected differences (28, 33, 62, 63). Pelagiphages
315 were abundant in the viral fraction from these samples, but there were no clear differences in
316 abundance that corresponded to SAR11 abundance, nor were the phages identified to subclade level
317 (64). Further characterization of the possible mechanisms leading to these differences may help explain
318 this disconnect between predicted growth rates and transcript to genome ratios seen in this study.

319 The observed differences in growth rate variation within a SAR11 subclade are more difficult to
320 discern. We observed a significant correlation between growth rate and a few environmental factors,
321 namely phosphate, silicate, temperature (all positive) and salinity (negative), but only with
322 representative MAGs from Ia.3, Ia.5, Ia.6, II and V subclades with higher growth rates. Phosphate
323 concentration best explained PTR estimates with subclades Ia.3, Ia.5, II and V, and there was a negative
324 relationship between expression of genes associated with phosphate uptake and PTR values. This
325 indicated that phosphate, along with access to carbohydrates mentioned above, are likely key factors in
326 regulating the variation in growth rates. Differences in phosphate and phosphonate transporter protein
327 abundance have been observed between the oligotrophic Sargasso Sea and nutrient-rich upwelling
328 areas (65, 66), but not on these finer temporal and spatial scales. A more recent study using incubations

329 of microbial communities demonstrated that SAR11 increased growth rates with the addition of
330 phosphate but not the elimination of grazers (29). Phosphate starvation or possibly increases in
331 phosphonate also may lead to large changes in SAR11 metabolism. SAR11 growth under phosphate-
332 limited but organophosphonate-replete conditions led to modifications of lipid content in response to
333 phosphate starvation (24).

334 Besides phosphate and ammonium transporter gene expression having a negative relationship
335 to growth rate estimates in subclades Ia.3, Ia.5, II and V, we did not observe any other genes that were
336 consistently differentially expressed in relation to PTR values. Transcription of some nitrogen-associated
337 genes (*gln*A, *glt*B) also generally increased with lower growth rates in our comparisons, but others
338 previously found associated with nitrogen limitation in cultured SAR11 (*yhd*W, *occ*T and *tau*A) were not
339 consistently increased with low growth rates (45). Our results support the use of PHA storage
340 mechanisms in two subclades (II and IIIa.4) under nutrient limiting conditions, where transcript levels of
341 a PHA polymerase (*pha*C) were negatively correlated with growth rates in the representative MAGs.
342 PhaC activity decreased during growth in *Pseudomonas putida* and in many bacteria increases under
343 nutrient-limiting conditions (67-69). Transporters for some sugars and amino acids were increased in
344 several MAGs with higher growth rates, This is in contrast to genes involved in translation, transcription,
345 post-translational modification, protein turnover and chaperones and intracellular trafficking and
346 secretion, which showed generally higher expression when growth rates were lower. However, several
347 similar changes in global transcriptional patterns were observed between salinities in the Chesapeake,
348 where no significant differences in growth rates were found nor any relationship between growth rates
349 and environmental conditions. This suggests that some transcriptional responses may not be specific to
350 growth rate.

351 Our results illustrate the different strategies that *Pelagibacterales* subclades utilize to
352 successfully dominate bacterial and archaeal communities in a variety of estuarine and ocean systems.

353 Not only do distinct members of the clade display niche specificity that is reflected in their genomic
354 repertoire, but they also respond differently to environmental conditions as reflected by their growth
355 rates. The differences were influenced by phosphate, ammonium, and carbohydrate transport and
356 storage mechanisms. Our abilities to track these differences relied on a seasonal and salinity gradient
357 study from two similar, but environmentally contrasting estuaries. This approach can be applied to
358 understand the relationships in other estuarine microbial populations. Further research is needed in
359 coastal or open ocean systems to examine SAR11 growth rates in relation to comprehensive space and
360 time gradients.

361 **Methods:**

362 *Environmental Characterization and Sample Collection*

363 Surface (~1.5 mbsf) water samples were collected using a rosette sampler with associated
364 conductivity, temperature, depth (CTD) profiles during three cruises in 2014 and two cruises in 2015 of
365 longitudinal transects of the Delaware and Chesapeake Bays, respectively, on the R/V Sharp
366 (https://www.bco-dmo.org/dataset/565451). Water quality data was measured using a Sea-Bird data
367 sonde, light intensity and attenuation were measured, and water samples were filtered for later
368 measurements of nutrients and bacterial production as previously described (70-73). Collections for
369 metagenomic and duplicate metatranscriptomic samples were identical to those previously described
370 (41, 74).

371 *MAG taxonomic diversity and abundance*

372 Surface water samples were collected from spring and summer salinity transects along the
373 Delaware and Chesapeake Bays in 2014 and 2015, respectively (Supplemental text, **Fig. S5**) (41, 74).
374 Metatranscriptome (MT) and metagenome (MG) sequencing, metagenome assembly, metagenome-
375 assembled genome (MAG) generation, MAG quality assessment (completeness and contamination), and

376 MAG taxonomic assignments are detailed elsewhere (41). Briefly, samples were collected, filtered to
377 either 0.8 µm or through 0.8 µm and onto 0.2 µm size fractions, and total nucleic acids were extracted.
378 Libraries were made from either DNA or RNA from single or replicate samples, respectively, at the Joint
379 Genome Institute and sequenced. 44 MAGs classified as *Pelagibacteriales* with at least 80%
380 completeness and less than 5% contamination determined by either CheckM (75) or Anvi'o (76) were
381 used in this study (41) (**Table S1**). The MAGs were determined to be at least of medium quality due to
382 the lack of rRNA genes in most of the genomes, according to the Minimum Information about a
383 Metagenome-Assembled Genome standard (77). The phylogenetic position of these MAGs in relation to
384 88 NCBI reference genomes (78) was determined with Anvi'o phylogenomic analyses (76) with the
385 Bacteria_71 dataset, where included genes needed to be present in at least 90% of the genomes.
386 Subsequent phylogenomic maximum likelihood tree generation with 500 replicates was generated using
387 the JTT model in MEGA X (79) and partial deletion from the resulting 38 ribosomal protein genes. The
388 font sizes and colors were modified in iTOL from the resulting Newick tree (80). The average nucleotide
389 identity (ANI) and average amino acid identity (AAI) between MAGs and reference genomes were
390 calculated using MiGA version 0.7.26.2 using the --fast and --haai-p options (81). The AAI similarity
391 matrix was plotted in heatmaply in R (82).

392 The above SAR11 MAGs and reference genomes were dereplicated at 95% identify by drep
393 v2.6.2 and verified with MiGA version 0.7.26.2 (81, 83). Ribosomal RNAs were subtracted from MTs
394 using SortMeRNA (84). Anvi'o v7 (76) MG and MT profiles were made by competitive mapping to
395 representative dereplicated MAGs (**Table S1**) using Bowtie2 v2.3.4.1 (85) in default end-to-end mode
396 (85). A count table of the resulting mapped reads and the Anvi'o gene-calls file reformatted to gff3 was
397 generated using featureCounts v1.5.2 (86). In addition, MG and MT classifications to Bacteria-Archaea
398 and order levels were performed with Kaiju v1.7.3 (87). The number of reads mapped to each group
399 were used to estimate abundances and to correct read counts between samples.

400 *Pangenome analyses*

401 Pangenomes were generated from MAGs and reference genomes/single amplified genomes
402 /MAGs that were at least 70% complete and less than 5% redundant (**Table S1**) using a gene clustering
403 approach (Diamond) in Anvi'o v7 (76). Protein-coding genes in each MAG were identified with Prodigal
404 v2.6.3 (88) and annotated internally with KEGG, COG and Pfam profiles in Anvi'o v7 (76) and externally
405 with Prokka v1.13 (89). Gene cluster, KEGG and COG counts for each MAG/reference genome were
406 extracted from the anvi-summarize gene cluster table and re-formatted using the reshape2 package in R
407 (90).

408 *Growth and Activity measures*

409 Maximum growth rate were estimated with gRodon in R (43) using representative MAGs listed
410 in **Table S1**. Rates for representative MAGs from each subclade from the fraction passing through a 0.8
411 µm filter (L08) were estimated via default CoPTR analyses (44), which estimates peak to trough ratios
412 from mapped MG reads. Additional MAGs (**Table S1**) and *Pelagibacter ubique* (HTCC1062) were used in
413 multiple replicate CoPTR analyses to validate the results.

414 The count table of MT reads mapped to representative MAGs described above was used for
415 RNA-seq analysis in the R DESeq2 package (v3.4.3) (91). The count table was first condensed to gene
416 cluster values (identified in Anvi'o) per MT sample, fitted with a negative binomial distribution and
417 transformed by regularized log transformation in DESeq2. DESeq2 was also used to generate the
418 principal component ordination plot. For differential expression analyses, subsets of the original count
419 table were used for each representative MAG of a SAR11 subclade. Samples with more than ~50,000
420 reads mapped to the MAG were used in DESeq2 (91) to identify differentially expressed genes, which
421 excluded particle-attached (cells captured on a 0.8 µm filter) samples because of the low abundance of
422 hits. To further reduce any potential effects of abundance differences, comparisons were limited to

423 within season, using SAR11 MAGs with less than two-fold change in relative abundance or coverage or
424 less than one-fold or no significant differences in *gyr*AB gene expression (92, 93) compared to one
425 another.

426 *Statistical analyses*

427 To test for significant differences in CoPTR values between genomospecies as well as the MT to
428 MG ratios that differed from 1, the Shapiro test (94) was first used to check for normality, followed by
429 the non-parametric Wilcoxon rank sum test (95). All statistical tests were performed in R version 4.1.0
430 (96). Pairwise correlation matrices between CoPTR and environmental variables were calculated using
431 the cor function in base R with Spearman as the model. The p-value for each pairwise correlation was
432 calculated using the cor.mtest function in R's corrplot library (97). The resulting correlation and p-value
433 matrices were visualized using the corrplot function in the same R library. Linear regression analyses
434 between CoPTR and environmental variables were performed using the ggscatter function in R's ggpubr
435 library (98). The non-parametric Spearman's correlation coefficient was used in all correlation analyses.

436 **Acknowledgements**

437 This work was supported by grants from the National Science Foundation (OCE-1261359) and the DOE
438 (CSP-1621) to BJC and DLK and an NSF grant (EF-2025541) to BJC. We thank Matt Cottrell and Liying Yu
439 for help with sampling and nutrient analyses and the crew of the R/V Hugh R. Sharp for sample
440 collection. We also thank Tijana Galvina del Rio at the Joint Genome Institute for project help.

441 BJC and DLK obtained funding for the study. All authors helped with sampling. BJC prepared the samples
442 and analyzed the data, with help from SJL. BJC wrote the paper with input from SJL and DLK.

443 **Competing Interests**

444 The authors declare no competing interests.

445

446 **Figure legends**

447 **Figure 1.** Phylogenomic analysis of SAR11 MAGs from the Delaware Bay (DE) and Chesapeake Bay (CP)
448 compared to previously characterized SAR11 and Rickettsia sp. (grey wedge) genomes via the Maximum
449 Likelihood method and JTT matrix-based model (79). A total of 38 shared genes were detected via Anvi'o
450 (76) with amino acid sequences derived from 135 genomes, SAGs or MAGs. There were 6532 positions
451 in the final dataset. The fraction of trees out of 100 in which the associated taxa clustered together are
452 indicated. The tree is drawn to scale, with branch lengths measured in the number of substitutions per
453 site. Information/accession number for each SAR11 genome is listed in **Table S1**.

454 **Figure 2.** Percentages of Pelagibacterales in total classified Bacteria and Archaea in the DE and CP Bay
455 metagenomes. A. Order level classification of metagenomic reads via Kajiu (87) in relation to total
456 bacterial and archaeal reads. B. Percentages of representative Pelagibacterales MAGs/SAGs/genomes in
457 the indicated metagenome in relation to total bacterial and archaeal reads. Abbreviations: CP=
458 Chesapeake Bay, DE = Delaware Bay, Spr = Spring, Sum = Summer, the numbers following indicate
459 salinity in PSU, G08 = >0.8 µm size fraction, L08 = <0.08 µm size fraction.

460 **Figure 3.** Average abundance of the indicated transporter gene in each SAR11 MAG. The transporters
461 per genomospecies was filtered and sorted by the COG category indicated on the right. Only
462 transporters found in 1-7 genomospecies are plotted. For a list of names and abundances of all
463 transporters see **SI at 10.6084/m9.figshare.19701478**.

464 **Figure 4.** Gene neighborhood map of genes related to the alternate ED glycolysis pathway described in
465 (28). Genomospecies and MAG or genome name listed on the left. Ia.5 and II MAGs also had genes on
466 other contigs not listed here. Full names of genes corresponding to gene symbols are listed in **SI at**
467 **10.6084/m9.figshare.19701478**.

468 **Figure 5.** Correlations between environmental variables, including average CoPTR values and
469 Pelagibacterales MT/MG ratios (A+B), and CoPTR values (C-F). A. Spring; B. Summer; C. Chesapeake,
470 genomospecies Ia.1+IIIa.2+IIIa.4; D. Chesapeake, genomospecies Ia.3+Ia.5+Ia.6+II+V; E. Delaware,
471 genomospecies Ia.1+IIIa.2+IIIa.4; F. Delaware, genomospecies Ia.3+Ia.5+Ia.6+II+V. * = $p<0.05$, ** =
472 $p<0.01$, *** = $p<0.001$.

473 **Figure 6.** Peak to trough ratios of representative MAGs from the indicated SAR11 genomospecies in
474 medium and high salinity L08 metagenome samples. Box and whisker plots of CoPTR values from the
475 indicated genomospecies. The number of metagenomes is indicated below each box and the a or b or ab
476 (not different) values above the boxes indicate significantly different groupings identified via the
477 Wilcoxon sum rank test (95).

478 **Figure 7.** Relative abundance of two SAR11 genomospecies in metatranscriptome and metagenomic
479 libraries. A. Genomospecies from Ia.3 subclade; B. Genomospecies from IIIa subclade. The p values are
480 from the Wilcoxon signed-rank test of the significance that the ratio of MT to MG is not equal to one. See
481 **Fig. S8 at 10.6084/m9.figshare.19701463** for data on other genomospecies.

482 **Figure 8.** Principle Coordinate Plot comparing transcripts from the indicated samples at the gene cluster
483 (GC) level. Representative SAR11 MAGs used are listed in Table S1. Transcripts were matched to gene
484 clusters identified in Anvi'o (76) and normalized in DESeq2 (91).

485 **Figure 9.** Common differentially expressed genes of representative MAGs from the indicated SAR11
486 genomospecies between Bay summer < 0.8 µm fraction (L08) metatranscriptome samples grouped by
487 COG category. Log2 Fold Change values listed as determined by DESeq2 (91). A. High salinity, B. Medium
488 salinity. Blue: increased in Delaware samples; red: increased in Chesapeake samples; white: not
489 significantly different; grey: not present. COG categories are listed on the right. Gene names are defined
490 in **SI at 10.6084/m9.figshare.19701478**.

491 **Supplemental Legend (for Supplemental Text, Tables S1 and S2, Figures S3, S4, S5, S6, S7 and S10, all**
492 **others linked to figshare)**

493 **Supplemental Text** – includes additional results and discussion ancillary to the main text.

494 **Table S1.** Properties of SAR11 MAGs, SAGs and genomes used in this study. Subclades were defined
495 based on identities to previously characterized MAGs, SAGs and genomes via phylogenomics and
496 AAI/ANI analyses. Genome/MAG properties were determined with Anvi'o.

497 **Table S2.** CoPTR values from a representative MAG in the indicated subclade from Delaware (DE) and
498 Chesapeake (CP) Bay metagenomes from the less than 0.8 µm size fraction. The number after Spr
499 (Spring), Sum (Summer or Fall is the salinity in PSU. N = night, D = day. Representative MAGs are listed in
500 **Table S1.**

501 **Figure S3.** Number of transporters in the indicated COG category in each SAR11 subclade. Values are
502 normalized to average genome size. A complete listing of all MAGs and reference genomes in each
503 subclade is listed in **Table S1**.

504 **Figure S4.** Average abundance of the indicated gene associated with C1 oxidative metabolism in each
505 SAR11 subclade. Heatmap was filtered and sorted by the C1 metabolic pathway indicated on the right.
506 For a full definition of gene symbols and list of gene names see **SI at 10.6084/m9.figshare.19701478**.

507 **Supplemental Figure 5**. Environmental Data for metagenomes and metatranscriptomes used in this
508 study. For further details, please see https://www.bco-dmo.org/dataset/565451. Average temperatures
509 (°C) for each bay and season are listed in the panels with production and PAR numbers. There are no
510 samples collected in the Fall in the Chesapeake Bay.

511 **Supplemental Figure 6.** Correlations between Pelagibacterales reads in MT/MG or average CoPTR values
512 of representative MAGs from each subclade (**Table S1**) and the indicated environmental variable in

513 spring or summer samples. p values were calculated by cor.mtest, after determining R values in cor via
514 the Spearman method.

515 **Supplemental Figure 7.** Growth rate estimates via CoPTR values. A. Correlations between CoPTR values
516 of representative MAGs from each subclade (**Table S1**) and the indicated environmental variable in
517 spring, summer and fall samples. B. Average CoPTR values from replicate MAGs from each indicated
518 subclade. p values were calculated by cor.mtest, after determining R values in cor via the Spearman
519 method.

520 **Supplemental Figure 10.** Common differentially expressed genes of representative MAGs from the
521 indicated SAR11 subclades between DE Bay summer less than 0.8 µm fraction (L08) metatranscriptome
522 samples grouped by COG category. Log2 Fold Change values listed as determined by DESeq2 (91). Blue:
523 increased in DE day medium and all high salinity samples; red: increased in DE medium night samples;
524 white: not significantly different; grey: not present. A. COG categories: C, C|T, E, E|G, E|T, G, H, I , P, Q
525 and S. B. COG categories: D, F, J, K, L, M and O. COG categories are listed on the right. Gene names are
526 defined in **SI at 10.6084/m9.figshare.19701478**.

527

528

529 **References**

530 1. Morris RM, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA, Giovannoni SJ. 2002.
531 SAR11 clade dominates ocean surface bacterioplankton communities. Nature 420:806-810.
532 2. Brown MV, Lauro FM, DeMaere MZ, Muir L, Wilkins D, Thomas T, Riddle MJ, Fuhrman JA,
533 Andrews-Pfannkoch C, Hoffman JM, McQuaid JB, Allen A, Rintoul SR, Cavicchioli R. 2012. Global
534 biogeography of SAR11 marine bacteria. Molecular Systems Biology 8.
535 3. Giovannoni SJ. 2017. SAR11 Bacteria: The Most Abundant Plankton in the Oceans. Annual
536 Review of Marine Science, Vol 9 9:231-255.
537 4. Henson MW, Lanclos VC, Faircloth BC, Thrash JC. 2018. Cultivation and genomics of the first
538 freshwater SAR11 (LD12) isolate. Isme Journal 12:1846-1860.

5. Kraemer S, Ramachandran A, Colatriano D, Lovejoy C, Walsh DA. 2019. Diversity and biogeography of SAR11 bacteria from the Arctic Ocean. The ISME Journal doi:10.1038/s41396-019-0499-4.
6. Campbell BJ, Kirchman DL. 2013. Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. The ISME journal 7:210-20.
7. Hugerth LW, Larsson J, Alneberg J, Lindh MV, Legrand C, Pinhassi J, Andersson AF. 2015. Metagenome-assembled genomes uncover a global brackish microbiome. Genome Biology 16.
8. Paver SF, Muratore D, Newton RJ, Coleman ML. 2018. Reevaluating the Salty Divide: Phylogenetic Specificity of Transitions between Marine and Freshwater Systems. Msystems 3.
9. Kan J, Evans SE, Chen F, Suzuki MT. 2008. Novel estuarine bacterioplankton in rRNA operon libraries from the Chesapeake Bay. Aquatic Microbial Ecology 51:55-66.
10. Kirchman DL, Cottrel MT, DiTullio GR. 2017. Shaping of bacterial community composition and diversity by phytoplankton and salinity in the Delaware Estuary, USA. Aquatic Microbial Ecology 78:93-106.
11. Dupont CL, Larsson J, Yooseph S, Ininbergs K, Goll J, Asplund-Samuelsson J, McCrow JP, Celepli N, Allen LZ, Ekman M, Lucas AJ, Hagstrom A, Thiagarajan M, Brindefalk B, Richter AR, Andersson AF, Tenney A, Lundin D, Tovchigrechko A, Nylander JAA, Brami D, Badger JH, Allen AE, Rusch DB, Hoffman J, Norrby E, Friedman R, Pinhassi J, Venter JC, Bergman B. 2014. Functional Tradeoffs Underpin Salinity-Driven Divergence in Microbial Community Composition. Plos One 9.
12. Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, Bibbs L, Eads J, Richardson TH, Noordewier M, Rappe MS, Short JM, Carrington JC, Mathur EJ. 2005. Genome streamlining in a cosmopolitan oceanic bacterium. Science 309:1242-1245.
13. Grote J, Thrash JC, Huggett MJ, Landry ZC, Carini P, Giovannoni SJ, Rappe MS. 2012. Streamlining and Core Genome Conservation among Highly Divergent Members of the SAR11 Clade. Mbio 3.
14. Moran MA, Belas R, Schell MA, Gonzalez JM, Sun F, Sun S, Binder BJ, Edmonds J, Ye W, Orcutt B, Howard EC, Meile C, Palefsky W, Goesmann A, Ren Q, Paulsen I, Ulrich LE, Thompson LS, Saunders E, Buchan A. 2007. Ecological genomics of marine roseobacters. Applied and Environmental Microbiology 73:4559-4569.
15. Ren QH, Paulsen IT. 2005. Comparative analyses of fundamental differences in membrane transport capabilities in prokaryotes and eukaryotes. Plos Computational Biology 1:190-201.
16. Beja O, Spudich EN, Spudich JL, Leclerc M, DeLong EF. 2001. Proteorhodopsin phototrophy in the ocean. Nature 411:786-789.
17. Morris RM, Vergin KL, Cho JC, Rappe MS, Carlson CA, Giovannoni SJ. 2005. Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-series Study site. Limnology And Oceanography 50:1687-1696.
18. Vergin KL, Beszteri B, Monier A, Thrash JC, Temperton B, Treusch AH, Kilpert F, Worden AZ, Giovannoni SJ. 2013. High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. Isme Journal 7:1322-1332.
19. Giovannoni SJ, Vergin KL. 2012. Seasonality in Ocean Microbial Communities. Science 335:671-676.
20. Carlson CA, Morris R, Parsons R, Treusch AH, Giovannoni SJ, Vergin K. 2009. Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. Isme Journal 3:283-295.
21. Eiler A, Hayakawa DH, Church MJ, Karl DM, Rappe MS. 2009. Dynamics of the SAR11 bacterioplankton lineage in relation to environmental conditions in the oligotrophic North Pacific subtropical gyre. Environmental Microbiology 11:2291-2300.
22. Sun J, Steindler L, Thrash JC, Halsey KH, Smith DP, Carter AE, Landry ZC, Giovannoni SJ. 2011. One Carbon Metabolism in SAR11 Pelagic Marine Bacteria. Plos One 6.

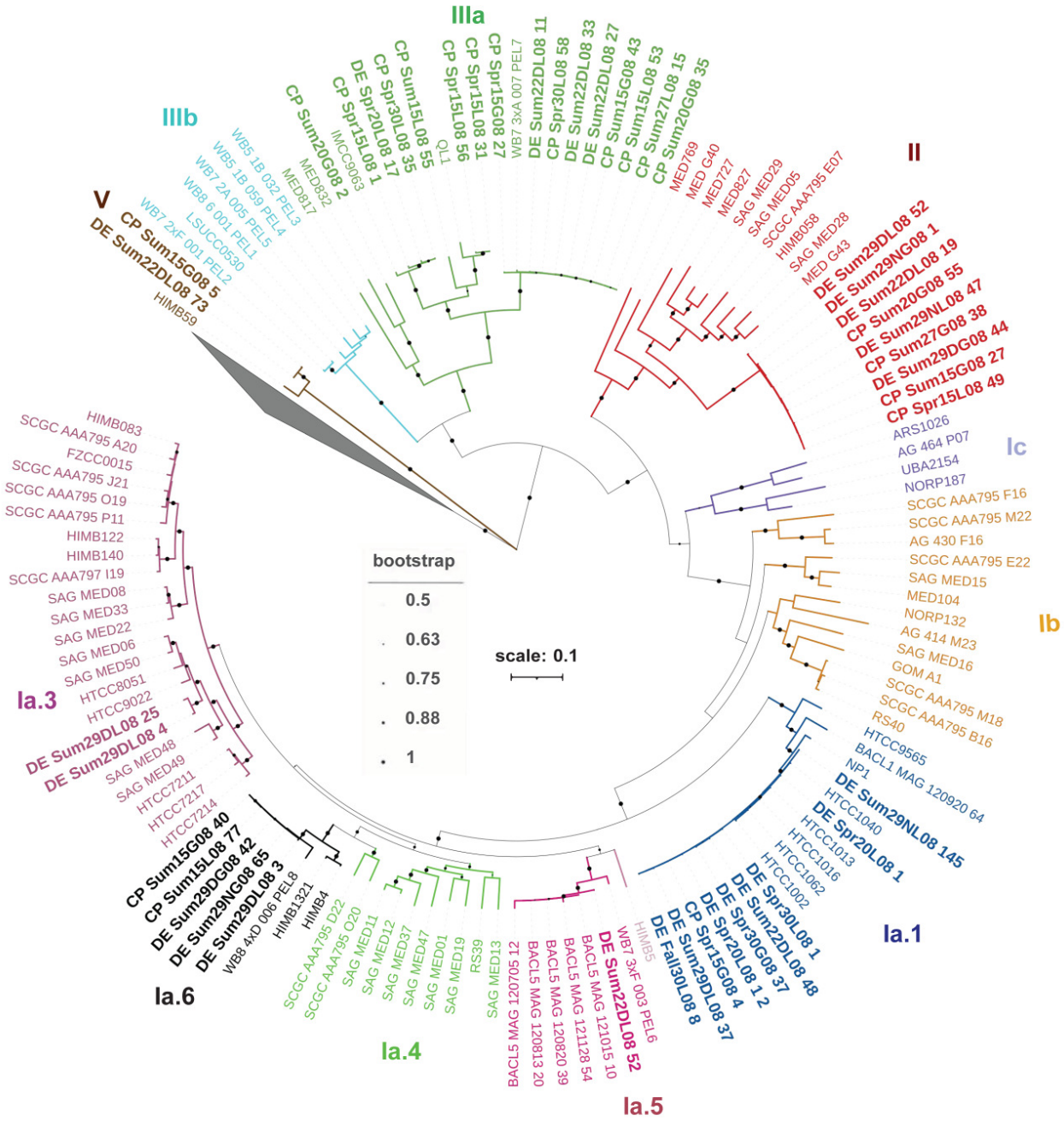
23. Rappé MS, Connon SA, Vergin KL, Giovannoni SJ. 2002. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. Nature 418:630-633.
24. Carini P, Van Mooy BAS, Thrash JC, White A, Zhao Y, Campbell EO, Fredricks HF, Giovannoni SJ. 2015. SAR11 lipid renovation in response to phosphate starvation. Proceedings of the National Academy of Sciences of the United States of America 112:7767-7772.
25. Lankiewicz TS, Cottrell MT, Kirchman DL. 2016. Growth rates and rRNA content of four marine bacteria in pure cultures and in the Delaware estuary. Isme Journal 10:823-832.
26. Campbell BJ, Yu L, Heidelberg JF, Kirchman DL. 2011. Activity of abundant and rare bacteria in a coastal ocean. Proc Natl Acad Sci USA 108:12776-81.
27. Coleman ML, Chisholm SW. 2010. Ecosystem-specific selection pressures revealed through comparative population genomics. Proceedings of the National Academy of Sciences of the United States of America 107:18634-18639.
28. Schwalbach MS, Tripp HJ, Steindler L, Smith DP, Giovannoni SJ. 2010. The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. Environmental Microbiology 12:490-500.
29. Fecskeova LK, Piwosz K, Santic D, Sestanovic S, Tomas AV, Hanusova M, Solic M, Koblizek M. 2021. Lineage-Specific Growth Curves Document Large Differences in Response of Individual Groups of Marine Bacteria to the Top-Down and Bottom-Up Controls. Msystems 6.
30. Malmstrom RR, Cottrell MT, Elifantz H, Kirchman DL. 2005. Biomass production and assimilation of dissolved organic matter by SAR11 bacteria in the Northwest Atlantic Ocean. Appl Environ Microbiol 71:2979-2986.
31. Malmstrom RR, Kiene RP, Cottrell MT, Kirchman DL. 2004. Contribution of SAR11 bacteria to dissolved dimethylsulfoniopropionate and amino acid uptake in the North Atlantic Ocean. Appl Environ Microbiol 70:4129-4135.
32. Dupont CL, McCrow JP, Valas R, Moustafa A, Walworth N, Goodenough U, Roth R, Hogle SL, Bai J, Johnson ZI, Mann E, Palenik B, Barbeau KA, Venter JC, Allen AE. 2015. Genomes and gene expression across light and productivity gradients in eastern subtropical Pacific microbial communities. ISME Journal 9:1076-1092.
33. Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA, DeLong EF. 2013. Pattern and synchrony of gene expression among sympatric marine microbial populations. Proceedings of the National Academy of Sciences of the United States of America 110:E488-E497.
34. Ottesen EA, Young CR, Gifford SM, Eppley JM, Marin R, 3rd, Schuster SC, Scholin CA, DeLong EF. 2014. Ocean microbes. Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. Science 345:207-12.
35. Aylward FO, Eppley JM, Smith JM, Chavez FP, Scholin CA, DeLong EF. 2015. Microbial community transcriptional networks are conserved in three domains at ocean basin scales. Proceedings of the National Academy of Sciences of the United States of America 112:5443-5448.
36. Pennock JR, Sharp JH. 1994. Temporal alternation between light-and nutrient-limitation of phytoplankton production in a coastal plain estuary. Marine Ecology Progress Series 111:275-288.
37. Boynton WR, Garber JH, Summers R, Kemp WM. 1995. Inputs, transformations, and transport of nitrogen and phosphorus in Chesapeake Bay and selected tributaries. Estuaries 18:285-314.
38. Malone TC, Conley DJ, Fisher TR, Glibert PM, Harding LW, Sellner KG. 1996. Scales of nutrient-limited phytoplankton productivity in Chesapeake Bay. Estuaries 19:371-385.
39. Fisher TR, Gustafson AB, Sellner K, Lacouture R, Haas LW, Wetzel RL, Magnien R, Everitt D, Michaels B, Karrh R. 1999. Spatial and temporal variation of resource limitation in Chesapeake Bay. Marine Biology 133:763-778.

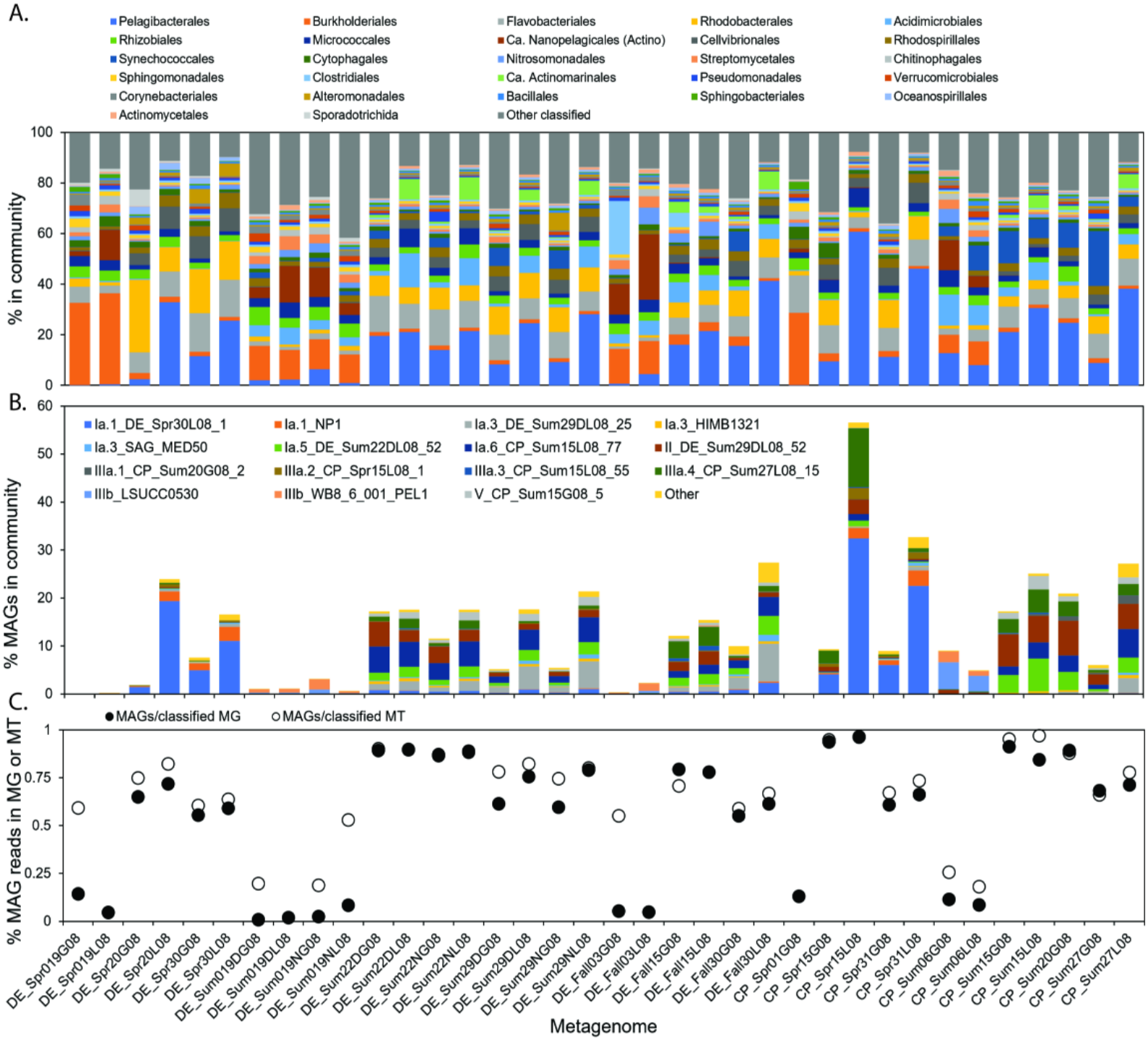
40. Kemp WM, Boynton WR, Adolf JE, Boesch DF, Boicourt WC, Brush G, Cornwell JC, Fisher TR, Glibert PM, Hagy JD, Harding LW, Houde ED, Kimmel DG, Miller WD, Newell RIE, Roman MR, Smith EM, Stevenson JC. 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. Marine Ecology-Progress Series 303:1-29.
41. Ahmed MA, Lim SJ, Campbell BJ. 2021. Metagenomes, Metatranscriptomes, and Metagenome-Assembled Genomes from Chesapeake and Delaware Bay (USA) Water Samples. Microbiology Resource Announcements 10.
42. Shang L, Jiang M, Chang HN. 2003. Poly(3-hydroxybutyrate) synthesis in fed-batch culture of Ralstonia eutropha with phosphate limitation under different glucose concentrations. Biotechnology Letters 25:1415-1419.
43. Weissman Jake L, Hou S, Fuhrman Jed A. 2021. Estimating maximal microbial growth rates from cultures, metagenomes, and single cells via codon usage patterns. Proceedings of the National Academy of Sciences 118:e2016810118.
44. Joseph TA, Chlenski P, Litman A, Korem T, Pe'er I. 2022. Accurate and robust inference of microbial growth dynamics from metagenomic sequencing reveals personalized growth rates. Genome Res doi:10.1101/gr.275533.121.
45. Smith DP, Thrash JC, Nicora CD, Lipton MS, Burnum-Johnson KE, Carini P, Smith RD, Giovannoni SJ. 2013. Proteomic and Transcriptomic Analyses of "Candidatus Pelagibacter ubique" Describe the First P-II-Independent Response to Nitrogen Limitation in a Free-Living Alphaproteobacterium. Mbio 4.
46. Kirchman DL. 2016. Growth Rates of Microbes in the Oceans. Annual Review of Marine Science, Vol 8 8:285-+.
47. Ducklow HW. 2000. Bacterial production and biomass in the oceans, p 85-120. *In* Kirchman DL (ed), Microbial ecology of the oceans. Wiley, New York.
48. Haro-Moreno JM, Rodriguez-Valera F, Rosselli R, Martinez-Hernandez F, Roda-Garcia JJ, Gomez ML, Fornas O, Martinez-Garcia M, Lopez-Perez M. 2020. Ecogenomics of the SAR11 clade. Environmental Microbiology 22:1748-1763.
49. Herlemann DPR, Woelk J, Labrenz M, Jurgens K. 2014. Diversity and abundance of "Pelagibacterales" (SAR11) in the Baltic Sea salinity gradient. Systematic and Applied Microbiology 37:601-604.
50. Tsementzi D, Rodriguez-R LM, Ruiz-Perez CA, Meziti A, Hatt JK, Konstantinidis KT. 2019. Ecogenomic characterization of widespread, closely-related SAR11 clades of the freshwater genus "Candidatus Fonsibacter" and proposal of Ca. Fonsibacter lacus sp. nov. Systematic and Applied Microbiology 42:495-505.
51. Campbell BJ, Yu L, Straza TRA, Kirchman DL. 2009. Temporal changes in bacterial rRNA and rRNA genes in Delaware (USA) coastal waters. Aquatic Microbial Ecology 57:123-135.
52. Hunt DE, Lin YJ, Church MJ, Karl DM, Tringe SG, Izzo LK, Johnson ZI. 2013. Relationship between Abundance and Specific Activity of Bacterioplankton in Open Ocean Surface Waters. Applied and Environmental Microbiology 79:177-184.
53. Zhang Y, Zhao ZH, Dai MH, Jiao NZ, Herndl GJ. 2014. Drivers shaping the diversity and biogeography of total and active bacterial communities in the South China Sea. Molecular Ecology 23:2260-2274.
54. Brown CT, Olm MR, Thomas BC, Banfield JF. 2016. Measurement of bacterial replication rates in microbial communities. Nature Biotechnology 34:1256-1263.
55. Carini P, Campbell EO, Morre J, Sanudo-Wilhelmy SA, Thrash JC, Bennett SE, Temperton B, Begley T, Giovannoni SJ. 2014. Discovery of a SAR11 growth requirement for thiamin's pyrimidine precursor and its distribution in the Sargasso Sea. Isme Journal 8:1727-1738.

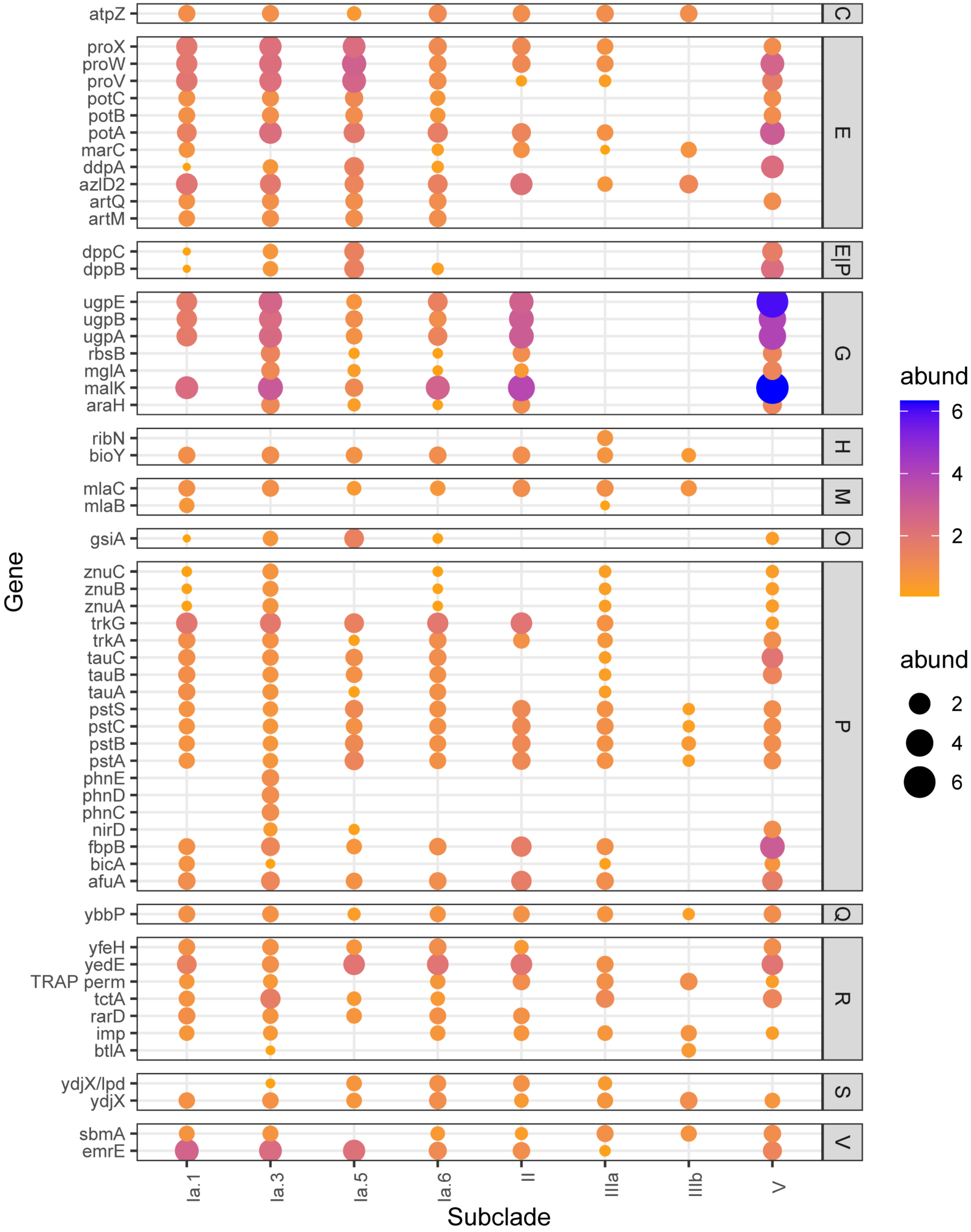
681 56. Ferrera I, Gasol JM, Sebastian M, Hojerova E, Koblizek M. 2011. Comparison of Growth Rates of
682 Aerobic Anoxygenic Phototrophic Bacteria and Other Bacterioplankton Groups in Coastal
683 Mediterranean Waters. Applied and Environmental Microbiology 77:7451-7458.
684 57. Kerkhof L, Kemp P. 1999. Small ribosomal RNA content in marine Proteobacteria during non-
685 steady-state growth. Fems Microbiology Ecology 30:253-260.
686 58. Kerkhof L, Ward BB. 1993. Comparison of Nucleic-Acid Hybridization and Fluorometry for
687 Measurement of the Relationship between Rna/DNA Ratio and Growth-Rate in a Marine
688 Bacterium. Applied and Environmental Microbiology 59:1303-1309.
689 59. Brown MV, Schwalbach MS, Hewson I, Fuhrman JA. 2005. Coupling 16S-ITS rDNA clone libraries
690 and automated ribosomal intergenic spacer analysis to show marine microbial diversity:
691 development and application to a time series. Environmental Microbiology 7:1466-1479.
692 60. Tsai AY, Gong GC, Mukhanov V. 2021. Experimental Warming Effects on Prokaryotic Growth and
693 Viral Production in Coastal Waters of the Northwest Pacific during the Cold Season. Diversity-
694 Basel 13.
695 61. Tuyet DTA, Tanaka T, Sohrin R, Hao DM, Nagaosa K, Kato K. 2015. Effects of warming on
696 microbial communities in the coastal waters of temperate and subtropical zones in the Northern
697 Hemisphere, with a focus on Gammaproteobacteria. Journal of Oceanography 71:91-103.
698 62. Thingstad TF, Lignell R. 1997. Theoretical models for the control of bacterial growth rate,
699 abundance, diversity and carbon demand. Aquatic Microbial Ecology 13:19-27.
700 63. Zhao YL, Temperton B, Thrash JC, Schwalbach MS, Vergin KL, Landry ZC, Ellisman M, Deerinck T,
701 Sullivan MB, Giovannoni SJ. 2013. Abundant SAR11 viruses in the ocean. Nature 494:357-360.
702 64. Sun MQ, Zhan YC, Marsan D, Paez-Espino D, Cai LL, Chen F. 2021. Uncultivated Viral Populations
703 Dominate Estuarine Viromes on the Spatiotemporal Scale. Msystems 6.
704 65. Sowell SM, Abraham PE, Shah M, Verberkmoes NC, Smith DP, Barofsky DF, Giovannoni SJ. 2011.
705 Environmental proteomics of microbial plankton in a highly productive coastal upwelling system.
706 Isme Journal 5:856-865.
707 66. Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, Barofsky DF, Carlson CA, Smith RD,
708 Giovanonni SJ. 2009. Transport functions dominate the SAR11 metaproteome at low-nutrient
709 extremes in the Sargasso Sea. Isme Journal 3:93-105.
710 67. La Rosa R, de la Peña F, Prieto MA, Rojo F, Rojo F. 2014. The Crc protein inhibits the production
711 of polyhydroxyalkanoates in Pseudomonas putida under balanced carbon/nitrogen growth
712 conditions. Environmental microbiology 16:278-290.
713 68. Ren Q, de Roo G, Witholt B, Zinn M, Thöny-Meyer L. 2010. Influence of growth stage on
714 activities of polyhydroxyalkanoate (PHA) polymerase and PHA depolymerase in Pseudomonas
715 putida U. BMC Microbiology 10:254.
716 69. Rehm BHA. 2010. Bacterial polymers: biosynthesis, modifications and applications. Nature
717 Reviews Microbiology 8:578-592.
718 70. Preen KE, Kirchman DL. 2004. Microbial respiration and production in the Delaware Estuary.
719 Aquatic Microbial Ecology 37:109-119.
720 71. Michelou VK, Cottrell MT, Kirchman DL. 2007. Light-stimulated bacterial production and amino
721 acid assimilation by cyanobacteria and other microbes in the North Atlantic Ocean. App Environ
722 Microbiol 73 5539-5546.
723 72. Cottrell MT, Mannino A, Kirchman DL. 2006. Aerobic anoxygenic phototrophic bacteria in the
724 Mid-Atlantic Bight and the North Pacific Gyre. Appl Environ Microbiol 72:557-564.
725 73. Kirchman DL, Dittel AI, Malmstrom RR, Cottrell MT. 2005. Biogeography of major bacterial
726 groups in the Delaware Estuary. Limnology and Oceanography 50:1697-1706.

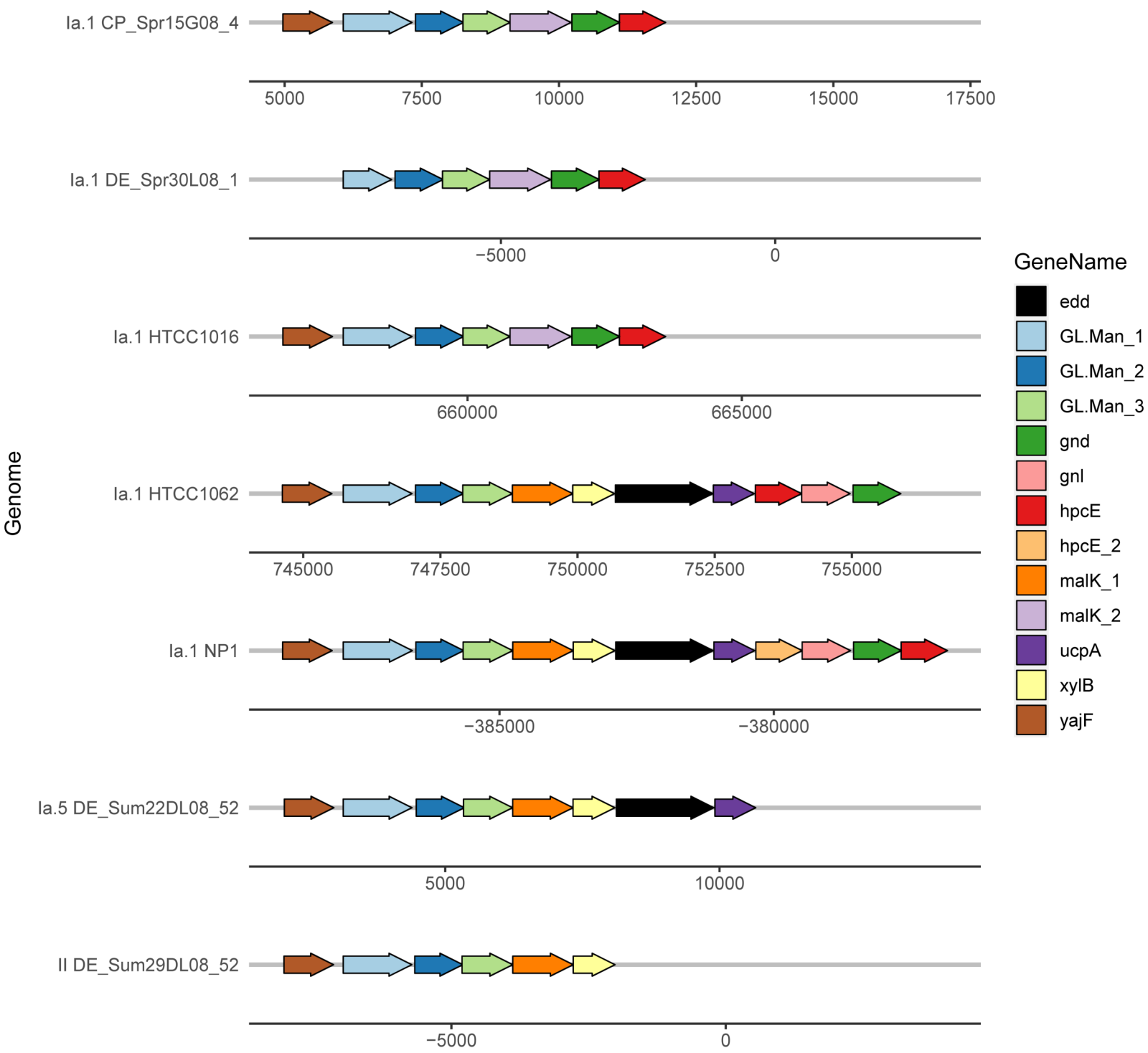
74. Maresca JA, Miller KJ, Keffer JL, Sabanayagam CR, Campbell BJ. 2018. Distribution and Diversity of Rhodopsin-Producing Microbes in the Chesapeake Bay. Applied and Environmental Microbiology 84.
75. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research 25:1043-1055.
76. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, Trigodet F, Watson AR, Esen ÖC, Moore RM, Clayssen Q, Lee MD, Kivenson V, Graham ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjödin A, Scott JJ, Vázquez-Campos X, McKay LJ, McDaniel EA, Stevens SLR, Anderson RE, Fuessel J, Fernandez-Guerra A, Maignien L, Delmont TO, Willis AD. 2021. Community-led, integrated, reproducible multi-omics with anvi'o. Nature Microbiology 6:3-6.
77. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooseph S, Sutton G, Glockner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu WT, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Lapidus A, Meyer F, Yilmaz P, Parks DH, Eren AM, et al. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nature Biotechnology 35:725-731.
78. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2006. GenBank. Nucleic Acids Research 34:D16-D20.
79. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution 30:2725-2729.
80. Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Research 49:W293-W296.
81. Rodriguez-R LM, Gunturu S, Harvey WT, Rossello-Mora R, Tiedje JM, Cole JR, Konstantinidis KT. 2018. The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. Nucleic Acids Research 46:W282-W288.
82. Galili T, O'Callaghan A, Sidi J, Sievert C. 2018. heatmaply: an R package for creating interactive cluster heatmaps for online publishing. Bioinformatics 34:1600-1602.
83. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. Isme Journal 11:2864-2868.
84. Kopylova E, Noe L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28:3211-3217.
85. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357-U54.
86. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923-930.
87. Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nature Communications 7.
88. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. Bmc Bioinformatics 11.
89. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-2069.
90. Wickham H. 2007. Reshaping Data with the reshape Package. Journal of Statistical Software 21:1 - 20.

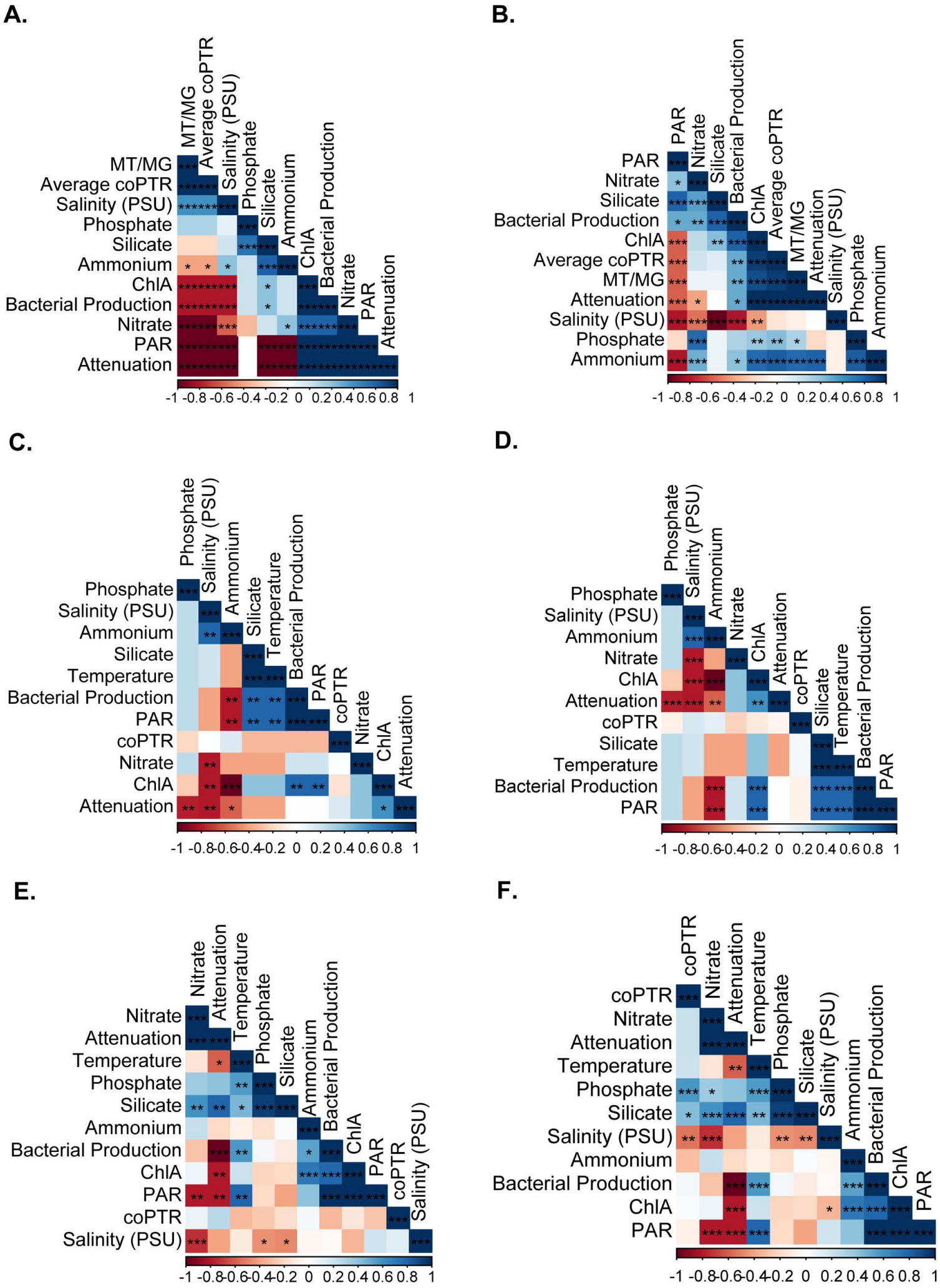
91. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15.
92. Nuccio EE, Starr E, Karaoz U, Brodie EL, Zhou JZ, Tringe SG, Malmstrom RR, Woyke T, Banfield JF, Firestone MK, Pett-Ridge J. 2020. Niche differentiation is spatially and temporally regulated in the rhizosphere. Isme Journal doi:10.1038/s41396-019-0582-x.
93. Klingenberg H, Meinicke P. 2017. How to normalize metatranscriptomic count data for differential expression analysis. Peerj 5.
94. Royston JP. 1982. An Extension of Shapiro and Wilk-W Test for Normality to Large Samples. Journal of the Royal Statistical Society Series C-Applied Statistics 31:115-124.
95. Bauer DF. 1972. Constructing Confidence Sets Using Rank Statistics. Journal of the American Statistical Association 67:687-690.
96. Team RC. 2020. R: A language and environment for statistical computing. https://www.R-project.org/. Accessed
97. Wei T, Simko V. 2021. R package 'corrplot': Visualization of a Correlation Matrix (Version 0.92), https://github.com/taiyun/corrplot.
98. Kassambara A. 2020. ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. , https://CRAN.R-project.org/package=ggpubr.

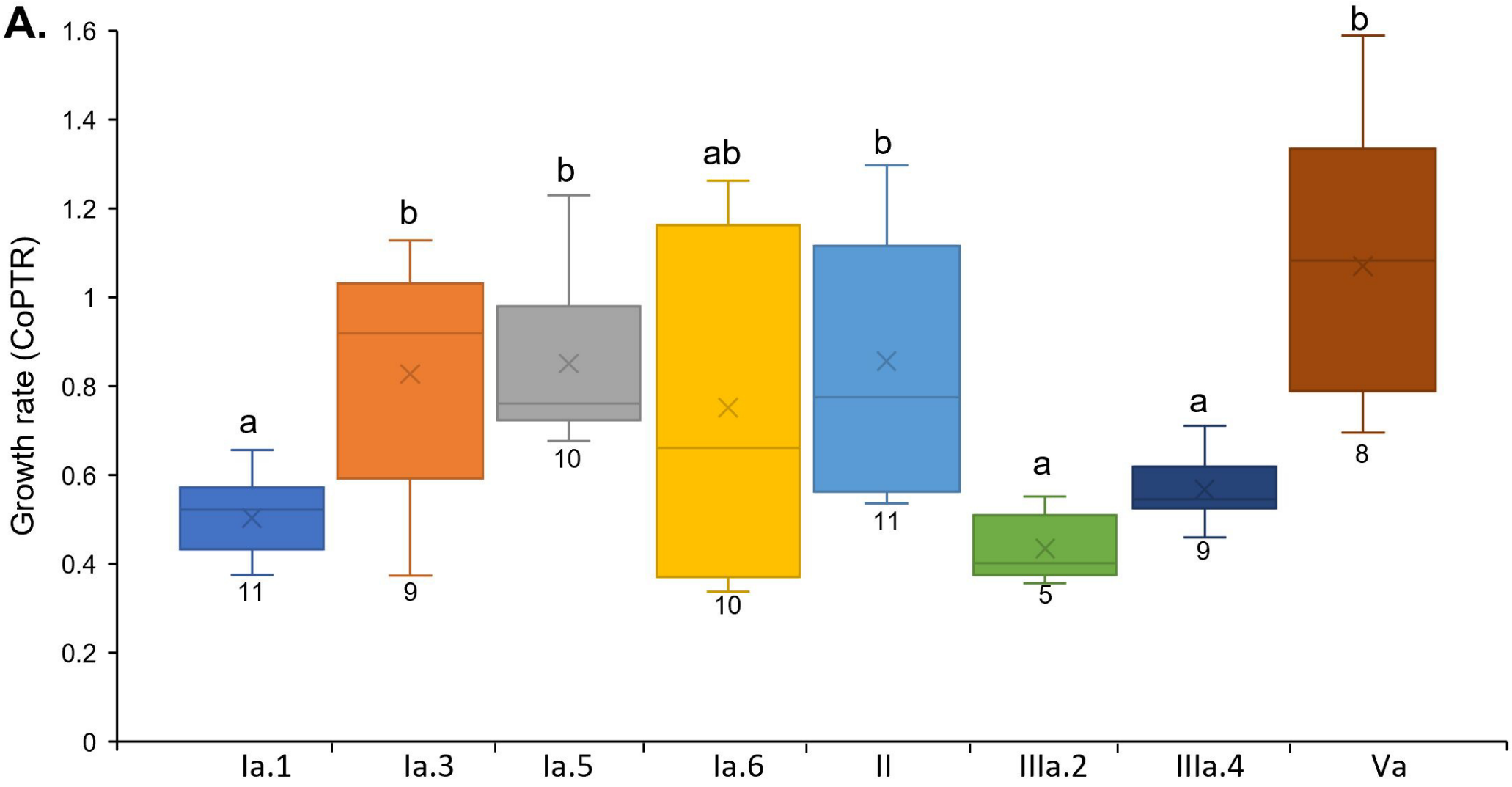

IIIa
IIIb
V
II
Ic
Ib
Ia.1
Ia.5
Ia.4
Ia.6
Ia.3
bootstrap
0.5
0.63
0.75
0.88
1
scale: 0.1
WB7 3xA 007 PEL7
DE Sum22DL08 11
CP Spr30L08 58
DE Sum22DL08 33
DE Sum22DL08 27
CP Sum15G08 43
CP Sum15L08 53
CP Sum27L08 15
CP Sum20G08 35
CP Spr15G08 27
CP Spr15L08 31
CP Spr15L08 56
CP Sum15L08 55
CP Spr30L08 35
CP Spr20L08 17
DE Spr15L08 1
CP Sum20G08 2
IMCC9063
QL1
MED832
MED817
WB5 1B 032 PEL3
WB5 1B 059 PEL4
WB7 2A 005 PEL5
WB8 6 001 PEL1
LSUCC0530
WB7 2xF 001 PEL2
CP Sum15G08 5
DE Sum22DL08 73
HIMB59
MED769
MED G40
MED727
MED827
SAG MED29
SAG MED05
SCGC AAA795 E07
HIMB058
SAG MED28
MED G43
DE Sum29DL08 52
DE Sum29NG08 1
DE Sum22DL08 19
DE Sum20G08 55
CP Sum29NL08 47
DE Sum27G08 38
CP Sum29DG08 44
DE Sum29DG08 27
CP Sum15G08 27
CP Spr15L08 49
ARS1026
AG 464 P07
UBA2154
NORP187
SCGC AAA795 F16
SCGC AAA795 M22
AG 430 F16
SCGC AAA795 E22
SAG MED15
MED104
NORP132
AG 414 M23
SAG MED16
GOM A1
SCGC AAA795 M18
SCGC AAA795 B16
RS40
HTCC9565
BACL1 MAG 120920 64
NP1
DE Sum29NL08 145
DE Spr20L08 1
HTCC1040
HTCC1013
HTCC1016
HTCC1062
HTCC1002
DE Spr30L08 1
DE Sum22DL08 48
DE Spr30G08 37
DE Spr20L08 1 2
CP Spr15G08 4
DE Sum29DL08 37
DE Fall30L08 8
HIMB5
WB7 3xF 003 PEL6
DE Sum22DL08 52
BACL5 MAG 121015 10
BACL5 MAG 121128 54
BACL5 MAG 120820 39
BACL5 MAG 120813 20
BACL5 MAG 120705 12
SAG MED13
RS39
SAG MED19
SAG MED01
SAG MED47
SAG MED37
SAG MED12
SAG MED11
SAG MED20
SCGC AAA795 D22
SCGC AAA795 O20
HIMB4
HIMB1321
WB8 4xD 006 PEL8
DE Sum29DL08 3
DE Sum29NG08 65
CP Sum15L08 77
DE Sum29DG08 42
CP Sum15G08 40
DE Sum29DL08 25
DE Sum29DL08 4
SAG MED48
SAG MED49
HTCC7211
HTCC7217
HTCC7214
HTCC9022
HTCC8051
SAG MED50
SAG MED06
SAG MED22
SAG MED33
SAG MED08
SCGC AAA797 I19
HIMB140
HIMB122
SCGC AAA795 P11
SCGC AAA795 O19
SCGC AAA795 J21
FZCC0015
SCGC AAA795 A20
HIMB083

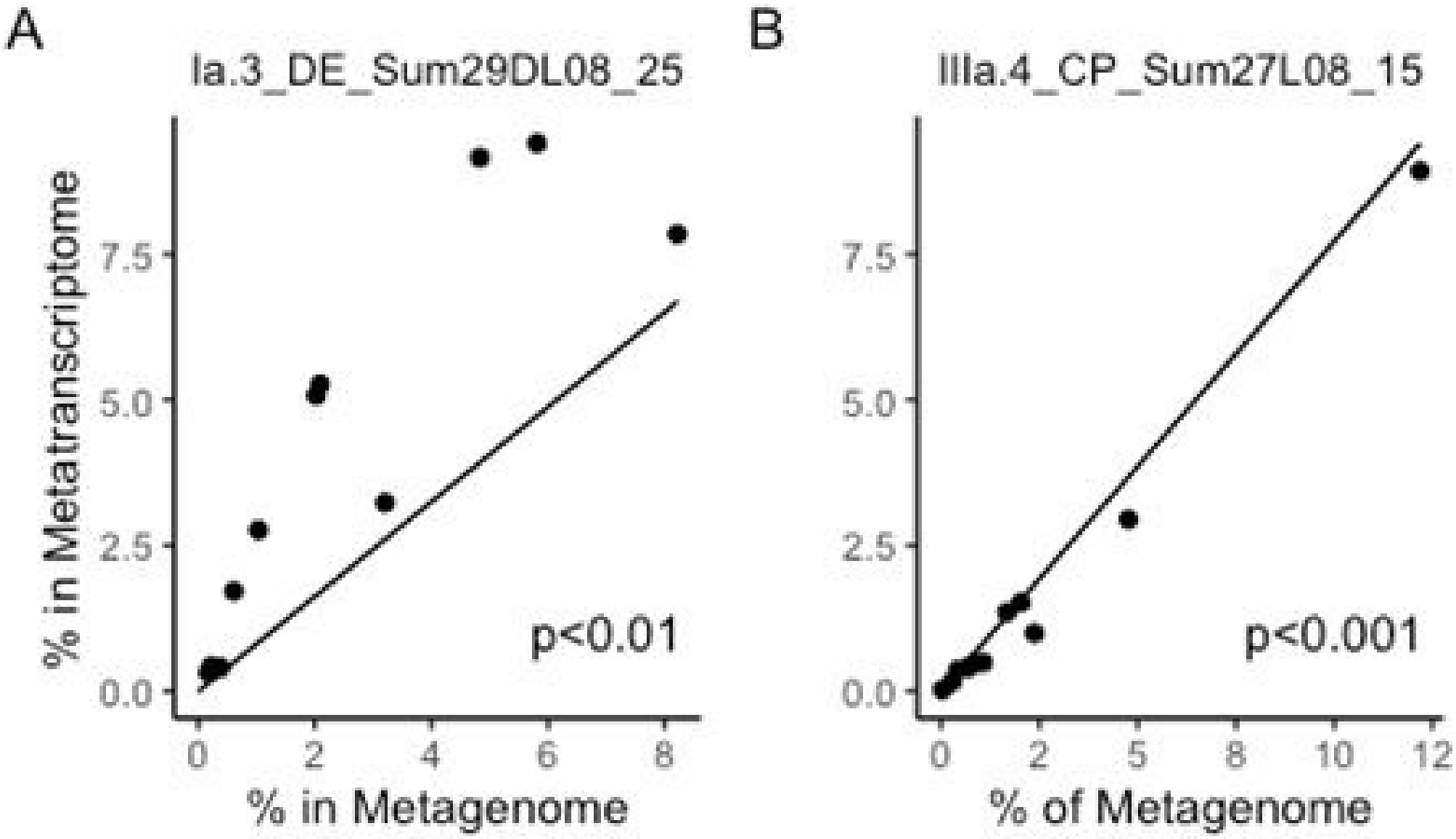

A.
Pelagibacterales
Burkholderiales
Flavobacteriales
Rhodobacterales
Acidimicrobiales
Rhizobiales
Micrococcales
Ca. Nanopelagicales (Actino)
Cellvibrionales
Rhodospirillales
Synechococcales
Cytophagales
Nitrosomonadales
Streptomycetales
Chitinophagales
Sphingomonadales
Clostridiales
Ca. Actinomarinales
Pseudomonadales
Verrucomicrobiales
Corynebacteriales
Alteromonadales
Bacillales
Sphingobacteriales
Oceanospirillales
Actinomycetales
Sporadotrichida
Other classified
% in community
B.
Ia.1_DE_Spr30L08_1
Ia.1_NP1
Ia.3_DE_Sum29DL08_25
Ia.3_HIMB1321
Ia.3_SAG_MED50
Ia.5_DE_Sum22DL08_52
Ia.6_CP_Sum15L08_77
II_DE_Sum29DL08_52
IIIa.1_CP_Sum20G08_2
IIIa.2_CP_Spr15L08_1
IIIa.3_CP_Sum15L08_55
IIIa.4_CP_Sum27L08_15
IIIb_LSUCC0530
IIIb_WB8_6_001_PEL1
V_CP_Sum15G08_5
Other
% MAGs in community
C.
MAGs/classified MG
MAGs/classified MT
% MAG reads in MG or MT
DE_Spr019G08
DE_Spr019L08
DE_Spr20G08
DE_Spr20L08
DE_Spr30G08
DE_Spr30L08
DE_Sum019DG08
DE_Sum019DL08
DE_Sum019NG08
DE_Sum019NL08
DE_Sum22DG08
DE_Sum22DL08
DE_Sum22NG08
DE_Sum22NL08
DE_Sum29DG08
DE_Sum29DL08
DE_Sum29NG08
DE_Sum29NL08
DE_Fall03G08
DE_Fall03L08
DE_Fall15G08
DE_Fall15L08
DE_Fall30G08
DE_Fall30L08
CP_Spr01G08
CP_Spr15G08
CP_Spr15L08
CP_Spr31G08
CP_Spr31L08
CP_Sum06G08
CP_Sum06L08
CP_Sum15G08
CP_Sum15L08
CP_Sum20G08
CP_Sum27G08
CP_Sum27L08
Metagenome

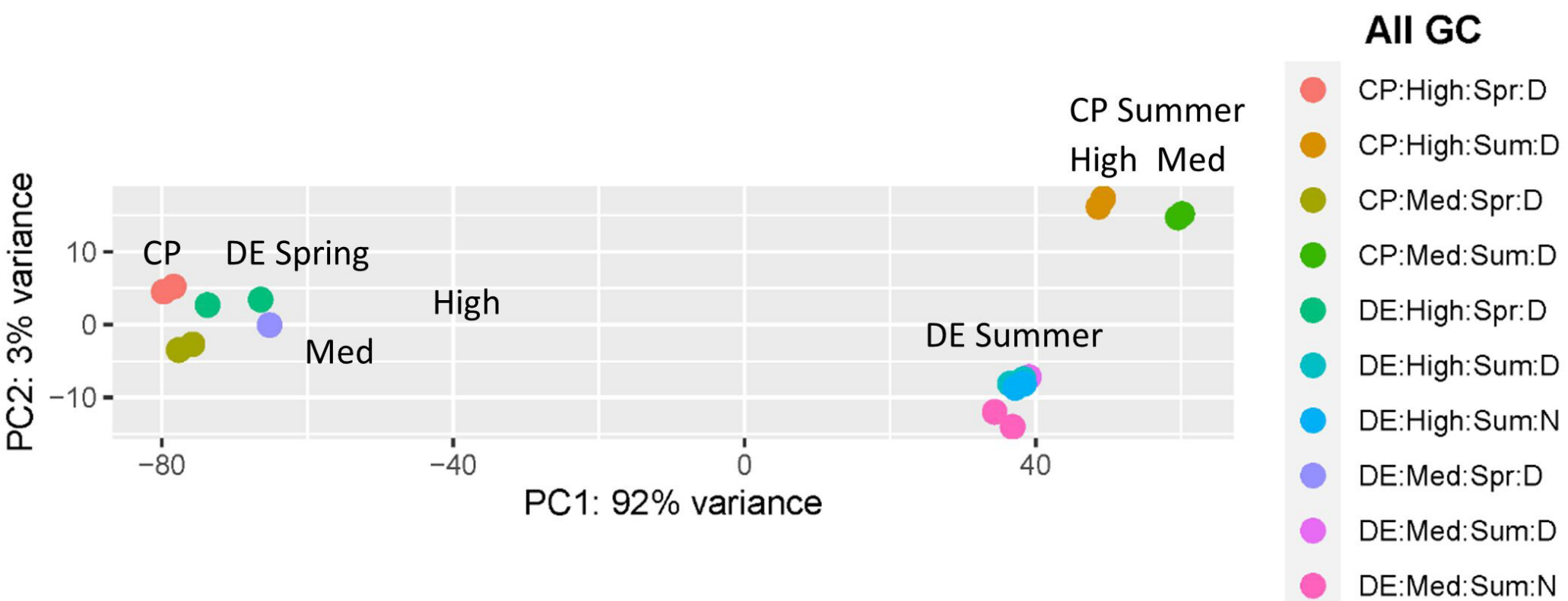

Gene
atpZ
C
proX
proW
proV
potC
potB
potA
marC
ddpA
azlD2
artQ
artM
E
dppC
dppB
E|P
ugpE
ugpB
ugpA
rbsB
mglA
malK
araH
G
ribN
bioY
H
mlaC
mlaB
M
gsiA
O
znuC
znuB
znuA
trkG
trkA
tauC
tauB
tauA
pstS
pstC
pstB
pstA
phnE
phnD
phnC
nirD
fbpB
bicA
afuA
P
ybbP
Q
yfeH
yedE
TRAP perm
tctA
rarD
imp
btlA
R
ydjX/lpd
ydjX
S
sbmA
emrE
V
Ia.1
Ia.3
Ia.5
Ia.6
II
IIIa
IIIb
V
Subclade
abund
6
4
2
abund
2
4
6

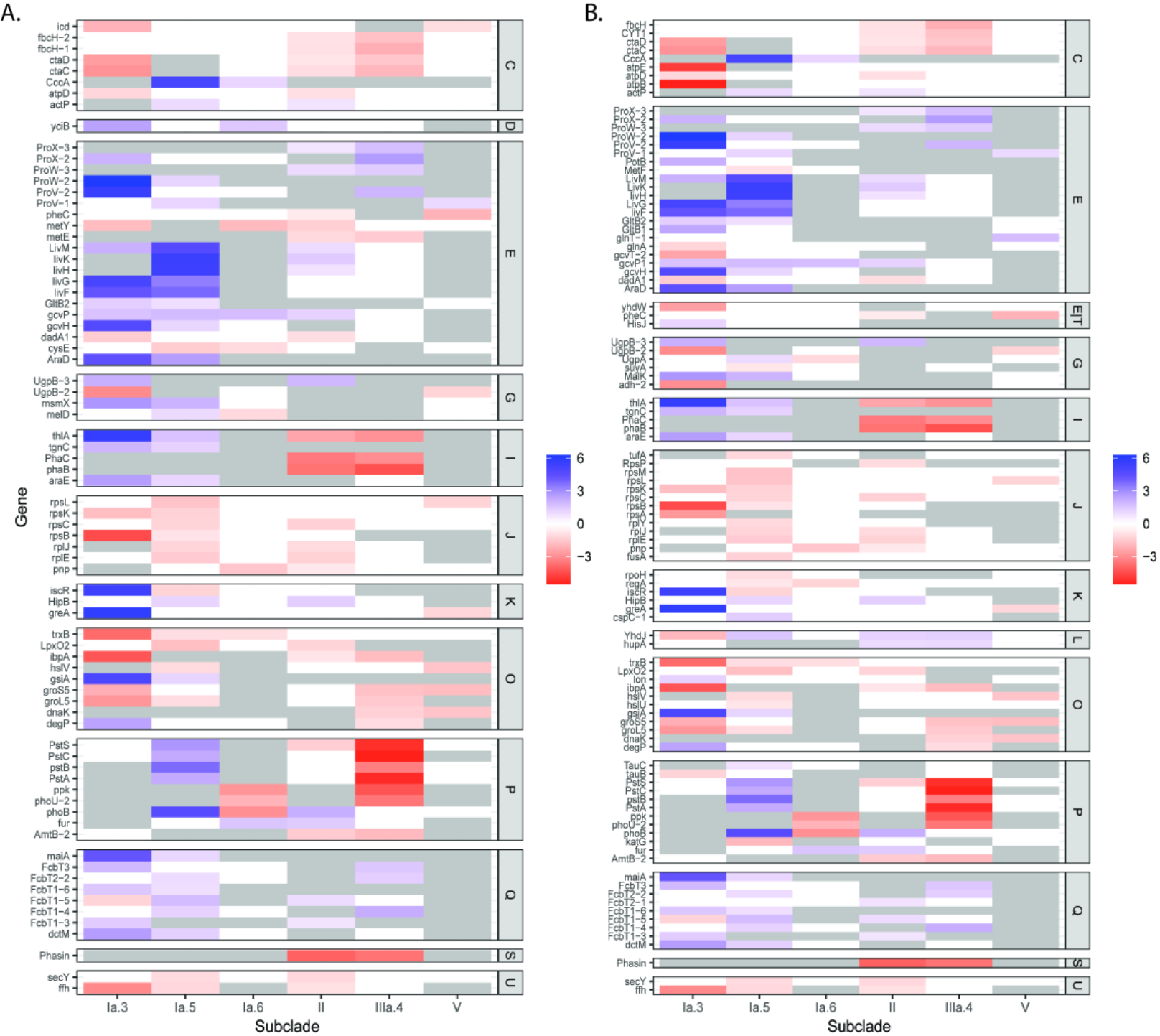

Ia.1 CP_Spr15G08_4
5000
7500
10000
12500
15000
17500
Ia.1 DE_Spr30L08_1
−5000
0
Ia.1 HTCC1016
660000
665000
Ia.1 HTCC1062
745000
747500
750000
752500
755000
Ia.1 NP1
−385000
−380000
Ia.5 DE_Sum22DL08_52
5000
10000
II DE_Sum29DL08_52
−5000
0
Genome
GeneName
edd
GL.Man_1
GL.Man_2
GL.Man_3
gnd
gnl
hpcE
hpcE_2
malK_1
malK_2
ucpA
xylB
yajF


A.
MT/MG
Average coPTR
Salinity (PSU)
Phosphate
Silicate
Ammonium
ChlA
Bacterial Production
Nitrate
PAR
Attenuation
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
B.
PAR
Nitrate
Silicate
Bacterial Production
ChlA
Average coPTR
MT/MG
Attenuation
Salinity (PSU)
Phosphate
Ammonium
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
C.
Phosphate
Salinity (PSU)
Ammonium
Silicate
Temperature
Bacterial Production
PAR
coPTR
Nitrate
ChlA
Attenuation
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
D.
Phosphate
Salinity (PSU)
Ammonium
Nitrate
ChlA
Attenuation
coPTR
Silicate
Temperature
Bacterial Production
PAR
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
E.
Nitrate
Attenuation
Temperature
Phosphate
Silicate
Ammonium
Bacterial Production
ChlA
PAR
coPTR
Salinity (PSU)
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
F.
coPTR
Nitrate
Attenuation
Temperature
Phosphate
Silicate
Salinity (PSU)
Ammonium
Bacterial Production
ChlA
PAR
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1


A.
Growth rate (CoPTR)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
a
b
b
ab
b
a
a
b
11
9
10
10
11
5
9
8
Ia.1
Ia.3
Ia.5
Ia.6
II
IIIa.2
IIIa.4
Va


A
Ia.3_DE_Sum29DL08_25
% in Metatranscriptome
7.5
5.0
2.5
0.0
p<0.01
0
2
4
6
8
% in Metagenome
B
IIIa.4_CP_Sum27L08_15
7.5
5.0
2.5
0.0
p<0.001
0
2
5
8
10
12
% of Metagenome


All GC
CP:High:Spr:D
CP:High:Sum:D
CP:Med:Spr:D
CP:Med:Sum:D
DE:High:Spr:D
DE:High:Sum:D
DE:High:Sum:N
DE:Med:Spr:D
DE:Med:Sum:D
DE:Med:Sum:N
CP Summer
High Med
CP
DE Spring
High
Med
DE Summer
PC2: 3% variance
10
0
−10
−80
−40
0
40
PC1: 92% variance


A.
B.
Gene
Subclade
Ia.3
Ia.5
Ia.6
II
IIIa.4
V
6
3
0
−3
C
D
E
EIT
G
I
J
K
L
O
P
Q
S
U
icd
fbcH−2
fbcH−1
ctaD
ctaC
CccA
atpD
actP
yciB
ProX−3
ProX−2
ProW−3
ProW−2
ProV−2
ProV−1
pheC
metY
metE
LivM
livK
livH
livG
livF
GltB2
gcvP
gcvH
dadA1
cysE
AraD
UgpB−3
UgpB−2
msmX
melD
thlA
tgnC
PhaC
phaB
araE
rpsL
rpsK
rpsC
rpsB
rplJ
rplE
pnp
iscR
HipB
greA
trxB
LpxO2
ibpA
hslV
gsiA
groS5
groL5
dnaK
degP
PstS
PstC
pstB
PstA
ppk
phoU−2
phoB
fur
AmtB−2
maiA
FcbT3
FcbT2−2
FcbT1−6
FcbT1−5
FcbT1−4
FcbT1−3
dctM
Phasin
secY
ffh
fbcH
CYT1
ctaD
ctaC
CccA
atpE
atpD
atpB
actP
ProX−3
ProX−2
ProW−3
ProW−2
ProV−2
ProV−1
PotB
MetF
LivM
LivK
livH
LivG
livF
GltB2
GltB1
gln1−1
glnA
gcv1−2
gcvP1
gcvH
dadA1
AraD
yhdW
pheC
HisJ
UgpB−3
UgpB−2
UgpA
sdyA
MalK
adh−2
thlA
tgnC
PhaC
phaB
araE
tufA
RpsP
rpsM
rpsL
rpsK
rpsC
rpsB
rpsA
rplY
rplJ
rplE
pnp
fusA
rpoH
regA
iscR
HipB
greA
cspC−1
YhdJ
hupA
trxB
LpxO2
lon
ibpA
hslV
hslU
gsiA
groS5
groL5
dnaK
degP
TauC
tauB
PstS
PstC
pstB
PstA
ppk
phoU−2
phoB
katG
fur
AmtB−2
maiA
FcbT3
FcbT2−2
FcbT2−1
FcbT1−6
FcbT1−5
FcbT1−4
FcbT1−3
dctM
Phasin
secY
ffh