

22 **Abstract**

23 Large spatial scales and long term shifts of bacterial community composition (BCC) in the
24 open ocean can often be reliably predicted based on the dynamics of physical-chemical
25 variables. The power of abiotic factors in shaping BCC at shorter time scales in shallow
26 estuarine mixing zones is less clear. We examined diurnal variation of BCC at different water
27 depths in spring and fall of 2011 at a station in the Gray's Reef National Marine Sanctuary
28 (GRNMS). This site is located in the transition zone between the estuarine plume and
29 continental shelf waters of the South Atlantic Bight. A total of 234,516 pyrotag sequences of
30 bacterial 16S rRNA genes were recovered; they were taxonomically affiliated with over
31 200 families of 23 bacterial phyla. Non-metric multidimensional scaling analysis revealed
32 significant differences of BCCs between spring and fall samples, likely due to seasonality in
33 the concentrations of dissolved organic carbon and nitrate+nitrite. Within each diurnal
34 sampling, BCC changed significantly by depth only for spring samples, and between day and
35 night only for fall samples. The former variations largely tracked changes of light
36 availability, while the latter was most correlated with concentrations of polyamines and
37 chlorophyll *a*. Our results suggest that at the GRNMS, a coastal mixing zone, diurnal
38 variation of BCCs is attributable to mixing of local and imported bacterioplankton rather than
39 to bacterial growth in response to environmental changes. Our results also indicated that like
40 roseobacters, SAR11 bacteria may play an important role in processing dissolved organic
41 material in coastal oceans.

42

43

44

45 **Keywords:** Bacterial community composition; Coastal seawater; Gray's Reef National
46 Marine Sanctuary; Pyrosequencing

47 **Introduction**

48 Spatial variability of microbes, i.e., the occurrence of distinct patterns of bacterial
49 community composition (BCC) among geographically isolated habitats or different depth
50 zones has been well established in marine environments (1–5). Annually recurring patterns
51 and pronounced seasonal variability in BCC have also been studied in a number of marine
52 environments (6–9).

53 However, BCC variations at shorter time scales, such as diurnal cycles, are relatively
54 understudied and available reports yield conflicting results. For example, BCC showed little
55 diurnal change in the Western English Channel (10) but varied significantly in coastal
56 California (11) and in the upper mixed layer of the Ligurian Sea (northwest of the
57 Mediterranean Sea) (12). Moreover, some BCC studies are based solely on molecular
58 fingerprinting methods, such as automated ribosomal intergenic spacer analysis (7) and
59 terminal restriction fragment length polymorphism (13), therefore, lacking taxonomic
60 resolution. Sequence-based studies are available but are often reported at very broad
61 taxonomic levels (such as phylum/class) (8, 11) or focused only on a few specific taxa (14).
62 Nonetheless, despite variations in sampling dimensions (spatial or vertical) and analytical
63 methods, microbial ecologists have reached a general consensus that long-term dynamics of
64 open ocean BCC are regulated by physical variables, most notably temperature (7, 8).
65 However, whether these environmental factors apply similar impacts on BCC in estuarine
66 mixing zones and at short time scales remain unclear.

67 This study investigated depth and diurnal dynamics of BCCs (i.e., relative abundance
68 of bacterioplankton taxa) at the Gray's Reef National Marine Sanctuary (GRNMS) in two
69 consecutive seasons and examined the potential correlations between BCCs and
70 environmental factors. GRNMS is located c.a. 32 km offshore of the coast of Georgia (USA)
71 in the transition zone between the nearshore estuarine plume and continental shelf waters in

72 the South Atlantic Bight. This coastal site has a long-term data set of physical and water
73 quality variables as a result of the GRNMS monitoring program and the efforts of other
74 researchers (15). Water chemistry at the reef is subjected to complex seasonal changes and
75 exhibits intense short-term dynamics due to strong tidal mixing and wind-driven advection
76 (15–17). Past GRNMS resource assessment efforts have documented community structures
77 of algae, coral, sponges and fish (18), leaving the bacterioplankton unexplored.

78 **Materials and methods**

79 *Sample collection and processing*

80 Two sets of water samples were collected on two cruises of the R/V Savannah to the
81 GRNMS (31° 24.04' N, 80° 51.51' W), one in spring (April 20-21, 2011) and the second in
82 fall (October 5-6, 2011). Water samples were collected every 3 h during a 24-hour period (8
83 or 9 casts in total) using Niskin bottles mounted on a rosette sampling system (Sea-Bird
84 Electronics, Bellevue, WA, USA). Samples taken after sunrise and before the following
85 sunset were labeled as day samples, while samples taken after sunset and before the
86 following sunrise were labeled as night samples. Depth profiles of environmental variables
87 including temperature (T), salinity (S), and photosynthetically active radiation (PAR) were
88 measured *in situ* with a conductivity-temperature-depth (CTD) water column profiler (Sea-
89 Bird Electronics, Bellevue, WA, USA) that was mounted on the rosette sampler system.
90 Water samples were collected at three different depths (~2, 4, and 17 m) in spring, when a
91 pycnocline was present, and at two depths (~2 and 17 m) in fall, when the pycnocline was
92 absent (Fig. S1). The tidal stage at time of sampling was quantified by transforming the data
93 into tidal angles (19). Briefly, a conversion factor was first calculated by dividing the time
94 interval between two successive high tides by 360°. Tidal angles were then calculated by
95 multiplying this conversion factor by the time interval between a sampling and its previous

96 high tide. As a result, tidal angles of 0° and 180° were corresponding to high and low tides,
97 respectively.

98 Immediately after collection, 1 liter of water samples were filtered sequentially
99 through $3\ \mu\text{m}$ and $0.2\ \mu\text{m}$ pore size membrane filters (Pall Life Sciences, Ann Arbor, MI,
100 USA). The filtrates were collected in 50 ml sterile conical centrifuge tubes and stored at -20
101 $^\circ\text{C}$ until analyses of organic carbon (DOC), dissolved nitrogen (DN), nitrate+nitrite (NO_x^-),
102 and soluble reactive phosphorus (SRP). Filtrates were also collected in 5 ml amber glass vials
103 and stored at -80°C for analyses of two labile dissolved organic nitrogen groups: dissolved
104 free amino acids (DFAAs) and polyamines (PAs). Cells that were collected on $0.2\ \mu\text{m}$ pore
105 size membrane filters were frozen immediately in liquid nitrogen onboard and stored at -80
106 $^\circ\text{C}$ in lab until DNA extraction. Particulate material in 500 ml whole water samples was
107 collected onto glass fiber filters (GF/F; Whatman International Ltd, Maidstone, England) for
108 chlorophyll *a* (Chl *a*) measurement, and stored in the dark at $-20\ ^\circ\text{C}$ before analysis.

109 Part of the water (2 ml) that passed through the $3\ \mu\text{m}$ pore size filters was preserved in
110 1% (final concentration) freshly prepared paraformaldehyde and stored at $4\ ^\circ\text{C}$ until cells
111 were enumerated by epifluorescence microscopy.

112 All glass ware and GF/F filters were combusted at $500\ ^\circ\text{C}$ for 4 h before use.
113 Triplicate samples were taken for analyses of nutrients and cell counts.

114 *Nutrient analysis*

115 Nutrient measurements were performed following standard procedures (20). Briefly,
116 DOC and DN concentrations were determined with a TOC/TN analyzer (TOC-V_{CPN};
117 Shimadzu Corp., Tokyo, Japan) based on combustion oxidation/infrared detection and
118 combustion/chemilluminescence detection methods, respectively. Concentrations of NO_x^- and
119 SRP were determined based on the cadmium reduction method and the molybdenum blue

120 method, respectively, using flow injection protocols (Lachat QuikChem FIA+ 8000Series,
121 Loveland, CO, USA).

122 Concentrations of DFAAs and PAs were measured using a Prominence 20A high-
123 performance liquid chromatography system (Shimadzu Corp., Tokyo, Japan) equipped with a
124 250 × 4.6 mm i.d. 5 µm particle size, Phenomenex Gemini-NX C₁₈ column (Phenomenex,
125 Torrance, CA, USA) following a protocol developed specifically for seawater samples (21).
126 Chl *a* was extracted from the GF/F filters with 90% acetone and measured by
127 spectrophotometry (22).

128 ***DNA extraction, PCR, and Pyrotag sequencing***

129 DNA was extracted from 0.2 µm pore size membrane filters using PowerSoil DNA
130 extraction kits (MoBio Laboratory Inc., Carlsbad, CA, USA). The V4-V6 regions of 16S
131 rRNA genes were PCR amplified using universal bacterial primers B530F (5'-GT GCC AGC
132 MGC NGG GG-3') (23) and B1100R (5'-GGG TTN CGN TCG TTG-3') (24). The forward
133 primers were constructed with an adaptor sequence (CCA TCT CAT CCC TGC GTG TCT
134 CCG ACT CAG) (23) and a 10 bp barcode tag, while the reverse primers were constructed
135 with an adaptor sequence (CCT ATC CCC TGT GTG CCT TGG CAG TCT CAG) only. The
136 PCR program included an initial denaturation at 95 °C for 3 min, then 30 cycles of 95 °C for
137 30 s, 58 °C for 1 min, 72 °C for 1 min, followed by a final elongation step at 72 °C for 5 min.
138 For each sample, 5 separate PCR reactions (25 µl each) were performed, the resulting 125 µl
139 of PCR amplicons were pooled and examined by gel electrophoresis (1% agarose) to verify
140 amplicon length. The amplicons were excised from the gels and purified first with the
141 QIAquick gel extraction kit (Qiagen, Chatsworth, CA, USA) and then with the Agencourt
142 AMPure XP system (Beckman Coulter, Brea, CA, USA). The quantity of purified PCR
143 amplicons was determined using the Quant-iT PicoGreen ds DNA Assay Kit (Life
144 technologies, Carlsbad, NY, USA). Equimolar amounts of PCR amplicons from 13 samples

145 (randomly assigned) were combined and sequenced in one run using a 454 GS Junior System
146 (Roche 454 Life Sciences, Branford, CT, USA) with unidirectional Lib-L chemistry. Three
147 pyrosequencing runs were performed to sequence a total of 39 samples collected in this
148 study.

149 *Taxonomic annotation of 16S rRNA gene pyrotag sequences*

150 Bacterial 16S rRNA gene pyrotag sequences were assigned to their sample of origin
151 based on barcodes. Primer and barcode sequences were then removed from each sequence.
152 Sequences that had wrong base calls in the primer regions, that were shorter than 65 bp, or
153 that contained chimeras (UCHIME) (25) were removed from further analysis. The remaining
154 sequences were clustered into operational taxonomic units (OTUs) using CD-HIT (26) at the
155 97% identity cutoff level. Singleton OTUs were excluded from the OTU list to prevent
156 potential overestimation of bacterial diversity (27). The longest sequence within each OTU
157 group was selected as a representative and blasted against the SILVA SSU database for
158 taxonomic annotation (28). Most sequences were summarized at the family level, except for
159 those that were affiliated with marine bacterial groups that have no official taxonomic
160 standings. The latter sequences were summarized at the clade level, such as SAR11 and
161 SAR116. This family/clade level organization of sequences was referred to as at the family
162 level hereafter for simplicity. Since the focus of the study was on bacterioplankton, sequences
163 annotated as chloroplasts were excluded from further analyses.

164 *Diversity and statistical analysis*

165 Diversity calculations and statistical analyses were performed using PRIMER v5
166 (Plymouth Marine Laboratory, Plymouth, UK; 29) unless mentioned otherwise. Family-level
167 diversity indices were calculated using the Shannon index. Family-level rarefaction curves
168 were constructed to infer the library coverage (30) using the vegan package in R (31).

169 The similarity of BCC, i.e., relative abundance of bacterial taxa, among samples was
170 examined using non-metric multidimensional scaling (NMDS) analysis, based on a Bray-
171 Curtis matrix that was calculated using normalized and square-root transformed relative
172 abundance of major bacterial families among samples (29). A hierarchical agglomerative
173 clustering analysis was also performed based on the same matrix using group-average
174 linkage. The robustness of NDMS grouping patterns was assessed by analysis of similarity
175 (ANOSIM), an analogue of the standard univariate analysis of variance (ANOVA). The
176 ANOSIM index r_{ANOSIM} was reported on a scale of 0 to 1. When $P < 0.05$, the sample groups
177 were reported as well-separated with $r_{ANOSIM} > 0.75$, clearly different but overlapping with 0.5
178 $< r_{ANOSIM} \leq 0.75$, separated but strongly overlapping with $0.25 < r_{ANOSIM} \leq 0.5$, or not
179 separable with $r_{ANOSIM} < 0.25$ (29). Similarity percentage (SIMPER) analysis was conducted
180 to identify the contribution of each bacterial family to the observed difference between
181 sample groups.

182 Redundancy analysis (RDA; 32) was used to explore correlations between changes in
183 BCC and environmental factors, including T, S, PAR, tidal stage (Tide), DOC, DN, NO_x^- ,
184 SRP, DFAAs, polyamines and Chl a , using the vegan package in the R software package
185 (31). The significance of the correlation values was assessed statistically by a Monte Carlo
186 analysis using 1000 permutations.

187 The significance of differences between depth and diurnal (day-night) variations in
188 bacterial diversity and composition was tested using Student's t test (for paired samples), or
189 one-way ANOVA (for multiple samples) within the R software package (33). Differences
190 were deemed significant when $P < 0.05$. Potential correlations between the relative
191 abundance of individual bacterial taxa and environmental variables were examined by
192 calculating Pearson's product momentum correlation coefficients (r) using the R software

193 package (33). Significant correlations were reported when $P < 0.05$. Bonferroni corrections of
194 P values were used for multiple tests.

195 ***Nucleotide Sequence accession number:***

196 The 16S rRNA partial sequences were deposited in the NCBI Sequence Read Archive
197 (SRA) under the project accession numbers: SRR1556928, SRR1602557 and SRR1602558.

198 **Results**

199 ***General statistics of 16S rRNA gene pyrotag sequences and taxonomic composition of***
200 ***bacterioplankton***

201 A total of 234,516 partial sequences of bacterial 16S rRNA genes with an average
202 read length of 561 bp passed quality control steps. The number of sequences per sample
203 ranged from 1,516 to 17,689 in spring and from 1,260 to 12,862 in fall, which corresponded
204 to 170 to 2,859 and 292 to 1,662 OTUs (roughly at the species level), respectively (Table S1).

205 These OTUs were widely distributed across the Domain Bacteria with over 200 unique
206 families from 23 phyla in total. Rarefaction curves of sequence libraries based on family-
207 level annotations all reached plateau (Fig. S2), despite differences in library size (either by
208 sequence or OTU number). Therefore, recovered pyrotag sequences were sufficient to
209 represent bacterioplankton diversity in our samples at the family level.

210 Over 85% of recovered bacterial sequences were affiliated with only 12 families of 5
211 phyla, namely the *Proteobacteria*, *Cyanobacteria*, *Actinobacteria*, *Deferribacteres*, and
212 *Bacteroidetes*. Each of these 12 families accounted for 2% or more of the total sequences (all
213 spring and fall libraries together) (Fig. 1; Table 1) and was designated as major taxa. Among
214 them, SAR11 (*Alphaproteobacteria*) and Family I *Cyanobacteria* (93.2% were
215 *Synechococcus*) were the most abundant, accounting for 30.7% and 13.6% of sequences on
216 average, respectively. Cluster analysis revealed that, in both seasons, patterns of SAR11
217 distribution differed significantly from Family I *Cyanobacteria* and other major taxa (Fig. 1).

218 ***Seasonal variations of BCC and influential environmental factors***

219 Family-level Shannon index (H) values showed that spring communities were more
220 diverse than those in fall (Table S1; *t* test, $P < 0.05$).

221 NMDS ordination analysis displayed a clear separation between spring and fall BCCs
222 (Fig. 2). ANOSIM ($r_{ANOSIM} = 0.99$; $P < 0.05$) analyses further confirmed the robustness of
223 this separation (Table S2). The SAR11 clade and Family I *Cyanobacteria* were the 2 most
224 abundant taxa in both seasons, with the former over-represented in spring (average 37.6% of
225 the sequences) relative to fall (21.5%) and the latter overrepresented in fall (20.1% relative
226 to spring (9.6%) (Table 1; *t* test, $P < 0.05$). SIMPER analysis further revealed that these two
227 taxa contributed the most to the observed seasonal BCC dissimilarities (24.0% and 16.5% of
228 dissimilarity, respectively), followed by *Vibrionaceae* (11.0% of dissimilarity) and
229 *Pseudoalteromonadaceae* (9.7% of dissimilarity) (Table 1).

230 RDA was used to examine the potential correlation between environmental variables
231 and the seasonality of BCC at the GRNMS (Fig. 3A). The two RDA axes, i.e., RDA1 and
232 RDA2, each captured 60.7% and 4.3% of the total variance in BCC, respectively. Spring and
233 fall samples were largely separated along RDA1, which broadly extracted the gradients of all
234 measured environmental variables ($P < 0.05$) except PAR and Tide, with NO_x^- , DOC, and T
235 contributing the most.

236 ***Depth profiles of BCC and influential environmental factors***

237 Bacterial diversity (H values) showed no significant difference between depths in fall
238 and overall samples ($P > 0.05$). However, bacterial diversity in spring was greater in surface
239 than in either mid-depth or bottom samples (Table S1; *t* test, $P < 0.05$).

240 Analysis using NMDS followed by ANOSIM revealed significant BCC differences
241 with depth (surface vs. mid-depth or bottom) only in spring ($r_{ANOSIM} = 0.52$, $r_{ANOSIM} = 0.77$,
242 respectively; $P < 0.05$), but not in fall ($r_{ANOSIM} = 0.03$, $P > 0.05$; Figs. 1 and 2; Table S2).

243 SIMPER analysis showed that nearly 60% of the difference in spring BCC was due to
244 differences in the abundances of SAR11, *Rhodobacteraceae*, Family I *Cyanobacteria* and
245 *Cytophaga* Family Incertae Sedis (Table 1). Among these taxa, SAR11 and *Cytophaga*
246 Family Incertae Sedis had higher relative abundance in mid-depth (average 42.1% and 6.5%
247 of the sequences, respectively) and bottom (39.5% and 8.4%, respectively) samples than in
248 surface samples (31.6% and 5.3%, respectively; ANOVA, $P < 0.05$). The distribution pattern
249 was opposite for *Rhodobacteraceae* and Family I *Cyanobacteria*, whose relative abundances
250 were the greatest in surface water (ANOVA, $P < 0.05$).

251 Consistently, RDA ordination also revealed a clear separation of spring BCCs based
252 on sampling depth (Fig. 3B). This separation was mostly along the RDA1_{spring} axis, which
253 explained 78.2% of the total variance and was significantly correlated with PAR ($P < 0.05$).

254 ***Diurnal dynamics of BCC and influential environmental factors***

255 Values of the Shannon index (H) revealed that, in either season, bacterioplankton
256 diversity of the day samples was not different from that of the night samples (Table S1; t test,
257 $P > 0.05$).

258 Analyses of BCC using NMDS and ANOSIM, on the other hand, identified
259 significant differences between day and night samples in fall ($r_{ANOSIM} = 0.50$, $P < 0.05$), but
260 not in spring ($r_{ANOSIM} = 0.00$, $P > 0.05$; Figures 1 and 2; Table S2). SIMPER analysis revealed
261 that over 40% of the variance between day and night BCCs in fall was contributed by shifts
262 in the relative abundances of SAR11, Family I *Cyanobacteria* (Table 1). SAR11 was more
263 abundant during the day (28.9% averaged across all fall samples) than night (15.5%) (t test, P
264 < 0.05), while, in contrast, Family I *Cyanobacteria* were more abundant at night (23.8%) than
265 during the day (15.4%) (t test, $P < 0.05$). SIMPER analysis also identified
266 *Pseudoalteromonadaceae* and *Vibrionaceae* as important contributors (each accounted for
267 ~11% of the difference) to the identified day-night BCC shifts in fall. However, this was not

268 supported by the t test, which found that the relative abundances of these taxa in day samples
269 were not significantly different from those in night samples in fall (t test; $P > 0.05$).

270 RDA analysis consistently separated the day and night BCCs in fall, mostly along the
271 RDA1_{fall} axis (47.9% of the total variance; Fig. 3C). This separation was significantly
272 correlated with variance in PAR and concentrations of polyamines and Chl a ($P < 0.05$).

273 **Correlation between major bacterial taxa and environmental factors**

274 As mentioned above, SIMPER analysis followed by ANOVA or t tests revealed that
275 the variation of BCC in time and depth was largely due to changes in the relative abundance
276 of six taxa, including SAR11, Family I *Cyanobacteria*, *Cytophaga* Family Incertae Sedis,
277 *Vibrionaceae*, *Pseudoalteromonadaceae* and *Rhodobacteraceae* (Tables 1). Pearson product-
278 moment correlation coefficients revealed that these changes were correlated significantly
279 with variations in environmental conditions between spring and fall (Table 2; $P < 0.05$ with
280 Bonferroni correction). The relative abundances of SAR11 and *Cytophaga* Family Incertae
281 Sedis were positively correlated with DOC ($r = 0.75$ and 0.76 , respectively) and NO_x^- ($r =$
282 0.52 and 0.74 , respectively), while negatively correlated with T, S, and SRP ($r \leq -0.58$). The
283 pattern was the opposite for Family I *Cyanobacteria*, *Vibrionaceae*, and
284 *Pseudoalteromonadaceae*, whose relative abundances were correlated positively with T, S,
285 and SRP ($r \geq 0.52$), but negatively with DOC and NO_x^- ($r \leq -0.50$). The relative abundance
286 of *Rhodobacteraceae* was not significantly correlated with any measured environmental
287 variable.

288 **Discussion**

289 GRNMS is a live-bottom reef ecosystem located in the transition zone between the
290 coastal/estuarine plume and continental shelf waters in the South Atlantic Bight. In
291 accordance with the complex local hydrology and chemistry, bacterioplankton communities

292 in GRNMS water were highly diverse and showed significant temporal and depth variations
293 in BCCs.

294 Despite sample variability, the two most abundant taxa were consistently SAR11
295 (*Alphaproteobacteria*) and Family I *Cyanobacteria* (93.2% were *Synechococcus*). These two
296 taxa represent numerically and ecologically significant groups of marine heterotrophic and
297 autotrophic bacterioplankton, respectively (34–38). *Rhodobacteraceae* was also identified as
298 a major taxon at the GRNMS, but their relative abundance ($7.5 \pm 0.4\%$ in average) was much
299 lower than SAR11 ($21.5 \pm 1.7\%$). The *Rhodobacteraceae* family contains several genera that
300 are collectively known as roseobacters. The relative abundance of roseobacters often exceeds
301 that of SAR11 in nutrient-rich coastal regions, likely a result of roseobacters' stronger ability
302 in processing plankton-derived DOC (39, 40). The observed predominance of SAR11 over
303 roseobacters at the GRNMS has also been found in coastal sites of the *Sorcerer II* Global
304 Ocean Sampling (41) and in Norwegian coastal waters of the North Sea (42). Moreover, we
305 found a significant correlation between nutrients (both DOC and nitrate concentrations) and
306 the relative abundance of SAR11 at the GRNMS (Table S3). These suggest that SAR11 clade
307 may play a more important role in organic substrate turnover in coastal seawaters than
308 previously thought.

309 Physical-chemical characteristics of seawater at the GRNMS, including T, S, turbidity
310 and nutrient concentrations, typically show strong temporal and depth variabilities (17, this
311 study). These factors have been repeatedly suggested as important regulators of BCC in many
312 other marine ecosystems (3, 6, 8, 9). Our data further indicates that the observed temporal
313 (both diurnal and seasonal) and depth variations of BCC at the GRNMS were correlated with
314 mixes of different environmental variables. This suggests that the impact of individual
315 environmental factors on BCCs vary among temporal and spatial scales.

316 BCC seasonality overwhelmed variations of BCC at either diurnal or depth scales
317 (Figs. 2 and 3). This pattern may be ascribed to the overriding influences of environmental
318 variables on defining the seasonal or long-term BCC dynamics at the GRNMS, as found in
319 other coastal sites (7, 9). Furthermore, we found that variables (i.e., NO_x^- , DOC, and T)
320 explaining the seasonality of BCC were not correlated with changes in BCC at shorter time
321 scales (day vs. night within 24 h) or among depths. Located in a coastal mixing zone,
322 seawater at the GRNMS actively exchanges with ambient environments through tidal
323 advection and turbulent mixing on the daily bases (15). The residence time of GRNMS water
324 has been estimated between 0.2-1.2 days (15), which may be too short to allow a detectable
325 growth of bacterial taxa. Therefore, the identified diurnal variation among our fall samples
326 (October 2011), when the water column is well mixed, may be more due to the mixing of
327 “new” bacterioplankton assemblages with local taxa than from differential growth of local
328 bacterial populations in response to changes of environmental conditions.

329 BCC variation with depth was found in spring (April 2011) (Fig. 1) when the water
330 column was stratified (Fig. S1). This indicates that the density difference between waters
331 above and below the pycnocline has created a barrier to prevent vertical mixing, which
332 allowed BCC of different depths to diverge along the PAR gradient. GRNMS is known for its
333 productive benthic community and substantial benthic nutrient fluxes (15, 16). However,
334 consistent with a previous study (15), significant vertical gradients of nutrient concentrations
335 were not observed by this study, likely resulting from the relatively large standing stock of
336 nutrients in the water column (15, Table S3). Therefore, the observed depth variations of
337 BCCs are more likely due to the difference between the content of dissolved organic and
338 inorganic nutrients across the pycnocline. Further studies on the correlation between nutrient
339 content and BCCs in GRNMS are needed to test this hypothesis.

340 **Conclusion**

341 Bacterioplankton community at the GRNMS consisted of a diverse group of taxa that
342 are typically found in coastal marine environments, including *Proteobacteria*, *Cyanobacteria*,
343 *Actinobacteria*, *Deferribacteres*, and *Bacteroidetes*. BCC at the GRNMS showed strong
344 seasonal variations and was correlated with environmental factors that regulate the growth of
345 bacterioplankton, such as T, S, and nutrient concentrations. Diurnal BCC variations at the
346 GRNMS, on the other hand, are suggested to be more impacted by water mixing events (tides
347 and vertical advection). The depth variations of BCCs are likely a result of poor vertical
348 mixing and different content of nutrients between the surface and bottom waters.

349

350 **Acknowledgments**

351 We appreciate the help from staff working on the R/V Savannah for assisting sample
352 collection and providing CTD data, C. Clevinger, M. Kelly and H. Bui for technical help. We
353 also thank two anonymous reviewers for providing comments that helped us improve this
354 paper. This study was supported by National Science Foundation Grants (OCE1029607 to
355 X.M. and OCE 1029742 to J.T. H) and Kent State University.

356 **References**

- 357 1. **Fortunato C S, Herfort L, Zuber P, Baptista AM, Crump BC.** 2012. Spatial variability
358 overwhelms seasonal patterns in bacterioplankton communities across a river to ocean
359 gradient. *ISME J.* **6**: 554–563.
- 360 2. **Friedline C, Franklin R, McCallister S, Rivera M.** 2012. Bacterial assemblages of the
361 eastern Atlantic Ocean reveal both vertical and latitudinal biogeographic signatures.
362 *Biogeosciences* **9**: 2177–2193.
- 363 3. **Logares R, Lindström ES, Langenheder S, Logue JB, Paterson H, Laybourn-Parry J,**
364 **Rengefors K, Tranvik L, Bertilsson S.** 2013. Biogeography of bacterial communities
365 exposed to progressive long-term environmental change. *ISME J.* **7**: 937–948.
- 366 4. **Qian PY, Wang Y, Lee OO, Lau SCK, Yang JK, Lafi FF, Al-Suwailem A, Wong TY.**
367 2011. Vertical stratification of microbial communities in the Red Sea revealed by 16S
368 rDNA pyrosequencing. *ISME J.* **5**: 507–518.
- 369 5. **Zinger L, Amaral-Zettler LA, Fuhrman JA, Horner-Devine MC, Huse SM, Welch**
370 **DBM, Martiny JBH, Sogin M, Boetius A, Ramette A.** 2011. Global patterns of bacterial
371 beta-diversity in seafloor and seawater ecosystems. *PLoS One* **6**: e24570.
- 372 6. **Andersson AF, Riemann L, Bertilsson S.** 2010. Pyrosequencing reveals contrasting
373 seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *ISME J.* **4**:
374 171–181.
- 375 7. **Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV, Naeem S.** 2006.
376 Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc.*
377 *Natl. Acad. Sci. U. S. A.* **103**: 13104–13109.
- 378 8. **Gilbert JA, Field D, Swift P, Newbold L, Oliver A, Smyth T, Somerfield PJ, Huse S,**
379 **Joint I.** 2009. The seasonal structure of microbial communities in the Western English
380 Channel. *Environ. Microbiol.* **11**: 3132–3139.

- 381 9. **Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B, Huse S,**
382 **McHardy AC, Knight R, Joint I, Somerfield P, Fuhrman JA, Field D.** 2012. Defining
383 seasonal marine microbial community dynamics. *ISME J.* **6:** 298–308.
- 384 10. **Gilbert JA, Field D, Swift P, Thomas S, Cummings D, Temperton B, Weynberg K,**
385 **Huse S, Hughes M, Joint I, Somerfield PJ, Mühling M.** 2010. The taxonomic and
386 functional diversity of microbes at a temperate coastal site: a ‘multi-omic’ study of seasonal
387 and diel temporal variation. *PLoS One* **5:** e15545.
- 388 11. **Olapade OA.** 2012. Diel fluctuations in the abundance and community diversity of
389 coastal bacterioplankton assemblages over a tidal cycle. *Microb. Ecol.* **63:** 96–102.
- 390 12. **Ghiglione JF, Mevel G, Pujo-Pay M, Mousseau L, Lebaron P, Goutx M.** 2007. Diel
391 and seasonal variations in abundance, activity, and community structure of particle-
392 attached and free-living bacteria in NW Mediterranean Sea. *Microb. Ecol.* **54:** 217–231.
- 393 13. **Treusch AH, Vergin KL, Finlay LA, Donatz MG, Burton RM, Carlson CA,**
394 **Giovannoni SJ.** 2009. Seasonality and vertical structure of microbial communities in an
395 ocean gyre. *ISME J.* **3:** 1148–1163.
- 396 14. **Pernthaler A, Pernthaler J.** 2005. Diurnal variation of cell proliferation in three
397 bacterial taxa from coastal North Sea waters. *Appl. Environ. Microbiol.* **71:** 4638–4644.
- 398 15. **Hopkinson CS, Fallon RD, Jasson BO, Schubauer, JP.** 1991. Community metabolism
399 and nutrient cycling at Gray's Reef, a hard bottom habitat in the Georgia Bight. *Mar. Ecol.*
400 *Prog. Ser.* **73:**105–120.
- 401 16. **National Marine Sanctuary Program (NMSP).** 2006. Gray's Reef National Marine
402 Sanctuary final management plan/final environmental impact statement. Available at
403 <http://graysreef.noaa.gov/management/pdfs/GraysReefFinalPlan.pdf>.

- 404 17. **Office of National Marine Sanctuaries (ONMS)**. 2008. Gray's Reef National Marine
405 Sanctuary condition report 2008. Available at [http://graysreef.noaa.gov/science/](http://graysreef.noaa.gov/science/publications/pdfs/grnms_condition_report08.pdf)
406 [publications/pdfs/grnms_condition_report08.pdf](http://graysreef.noaa.gov/science/publications/pdfs/grnms_condition_report08.pdf)
- 407 18. **Bauer LJ, Kendall MS, Jeffrey CF**. 2008. Incidence of marine debris and its
408 relationships with benthic features in Gray's Reef National Marine Sanctuary, Southeast
409 USA. *Mar. Pollut. Bull.* **56**: 402–413.
- 410 19. **Pinckney J L, Zingmark RG**. 1991. Effects of tidal stage and sun angles on intertidal
411 benthic microalgal productivity. *Mar. Ecol. Prog. Ser.* **76**: 81–89.
- 412 20. **Clescerl LS, Greenberg AE, Eaton AD**. 1999. Standard methods for the examination of
413 water and wastewater. American Public Health Association, Washington, D.C.
- 414 21. **Lu X, Zou L, Clevinger C, Hollibaugh JT, Liu Q, Mou X**. 2014. Temporal dynamics
415 and depth variations of dissolved free amino acids and polyamines in coastal seawater
416 determined by high-performance liquid chromatography. *Mar. Chem.* **163**: 36–44.
- 417 22. **Tett P, Kelly MG, Ijorner GM**. 1975. A method for the spectrophotometric
418 measurement of river periphyton chlorophyll *a* and pheophytin *a* using several extractions
419 with methanol. *Limnol. Oceanogr.* **20**: 887–896.
- 420 23. **Vossbrinck CR, Baker MD, Didier ES, Debrunner-Vossbrinck BA, Shaddock JA**.
421 1993. Ribosomal DNA sequences of *Encephalitozoon hellem* and *Encephalitozoon*
422 *cuniculi*: species identification and phylogenetic construction. *J. Eukaryot. Microbiol.* **40**:
423 354–362.
- 424 24. **Turner S, Pryer KM, Miao VP, Palmer JD**. 1999. Investigating deep phylogenetic
425 relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis.
426 *J. Eukaryot. Microbiol.* **46**: 327–338.
- 427 25. **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R**. 2011. UCHIME improves
428 sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200.

- 429 26. **Li W, Godzik A.** 2006. Cd-hit: a fast program for clustering and comparing large sets of
430 protein or nucleotide sequences. *Bioinformatics* **22**: 1658–1659.
- 431 27. **Kunin V, Engelbrektson A, Ochman H, Hugenholtz P.** 2010. Wrinkles in the rare
432 biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates.
433 *Environ. Microbiol.* **12**:118–123.
- 434 28. **Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO.** 2007.
435 SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA
436 sequence data compatible with ARB. *Nucleic. Acids. Res.* **35**: 7188–7196.
- 437 29. **Clarke KR, Warwick RM.** 2001. Change in marine communities: an approach to
438 statistical analysis and interpretation (2nd) Edition. PRIMER-v5, Plymouth, UK.
- 439 30. **Youssef NH, Elshahed MS.** 2008. Species richness in soil bacterial communities: a
440 proposed approach to overcome sample size bias. *J. Microbiol. Methods* **75**: 86–91.
- 441 31. **Oksanen J, Kindt R, Legendre P, O'Hara RB.** 2007. *vegan*: Community Ecology
442 Package version 1.8–5. Available at <http://r-forge.r-project.org/projects/vegan/>.
- 443 32. **Lepš J, Šmilauer P.** 2003. Multivariate analysis of ecological data using CANOCO.
444 Cambridge University Press: Oxford, UK, pp 62–63.
- 445 33. **R Core Development Team.** 2005. The R project for statistical computing.
446 <http://www.R-project.org>.
- 447 34. **Carlson CA, Morris R, Parsons R, Treusch AH, Giovannoni SJ, Vergin K.** 2009.
448 Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the
449 northwestern Sargasso Sea. *ISME J.* **3**: 283–295.
- 450 35. **Morris RM, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA,**
451 **Giovannoni SJ.** 2002. SAR11 clade dominates ocean surface bacterioplankton
452 communities. *Nature* **420**: 806–810.

- 453 36. **Scanlan DJ, West NJ.** 2002. Molecular ecology of the marine cyanobacterial genera
454 *Prochlorococcus* and *Synechococcus*. FEMS Microbiol. Ecol. **40**: 1–12.
- 455 37. **Waterbury JB, Watson SW, Guillard RR, Brand LE.** 1979. Widespread occurrence of
456 a unicellular, marine, planktonic, cyanobacterium. Nature **277**: 293–294.
- 457 38. **Waterbury JB, Watson SW, Valois FW, Franks DG.** 1986. Biological and ecological
458 characterization of the marine unicellular cyanobacterium *Synechococcus*. Can. Bull. Fish.
459 Aquat. Sci. **214**: 71–120.
- 460 39. **Hahnke S, Brock NL, Zell C, Simon M, Dickschat JS, Brinkhoff T.** 2013.
461 Physiological diversity of roseobacter clade bacteria co-occurring during a phytoplankton
462 bloom in the North Sea. Syst. Appl. Microbiol. **36**: 39–48.
- 463 40. **González JM, Simó R, Massana R, Covert JS, Casamayor EO, Pedrós-Alió C,**
464 **Moran MA.** 2000. Bacterial community structure associated with a
465 dimethylsulfoniopropionate-producing North Atlantic algal bloom. Appl. Environ.
466 Microbiol. **66**: 4237–4246.
- 467 41. **Biers EJ, Sun S, Howard EC.** 2009. Prokaryotic genomes and diversity in surface ocean
468 waters: interrogating the global ocean sampling metagenome. Appl. Environ. Microbiol.
469 **75**: 2221–2229.
- 470 42. **Giebel HA, Kalhoefer D, Lemke A, Thole S, Gahl-Janssen R, Simon M, Brinkhoff T.**
471 2011. Distribution of roseobacter RCA and SAR11 lineages in the North Sea and
472 characteristics of an abundant RCA isolate, ISME J. **5**: 8–19.
- 473
474
475
476

477 **Figure legends**

478 Fig. 1. Heatmap and accompanying cluster analysis of the relative abundance (%) of major
479 GRNMS bacterioplankton at the family level in (A) spring (April, 2011) and (B) fall
480 (October, 2011) samples. Sample identifiers are based on sampling season (sp, spring; fa,
481 fall), depth (s, surface; m, mid-depth; b, bottom) and time (in 24-hour format). Regular and
482 bold font styles are used to denote day and night samples, respectively.

483

484 Fig. 2. The non-metric multidimensional scaling ordination based on the relative abundance
485 (%) of major GRNMS bacterioplankton at the family level in spring and fall samples. Sample
486 identifiers are the same as in Fig. 1. Symbols are used to denote different sampling depths in
487 spring (black triangle: surface; gray triangle: mid-depth; white triangle: bottom) and in fall
488 (black square: surface; white square: bottom).

489

490 Fig. 3. Biplot diagrams of the RDA analysis on the relative abundance (%) of BCC by
491 environmental variables at the GRNMS for (A) overall, (B) spring and (C) fall samples.
492 Sample identifiers are the same as in Fig. 1. Environmental variables are in black boxes with
493 white font. Light and dark gray shadings are used to indicate significantly different BCC
494 sample clusters identified by ANOSIM analysis.

495

Fig. 1

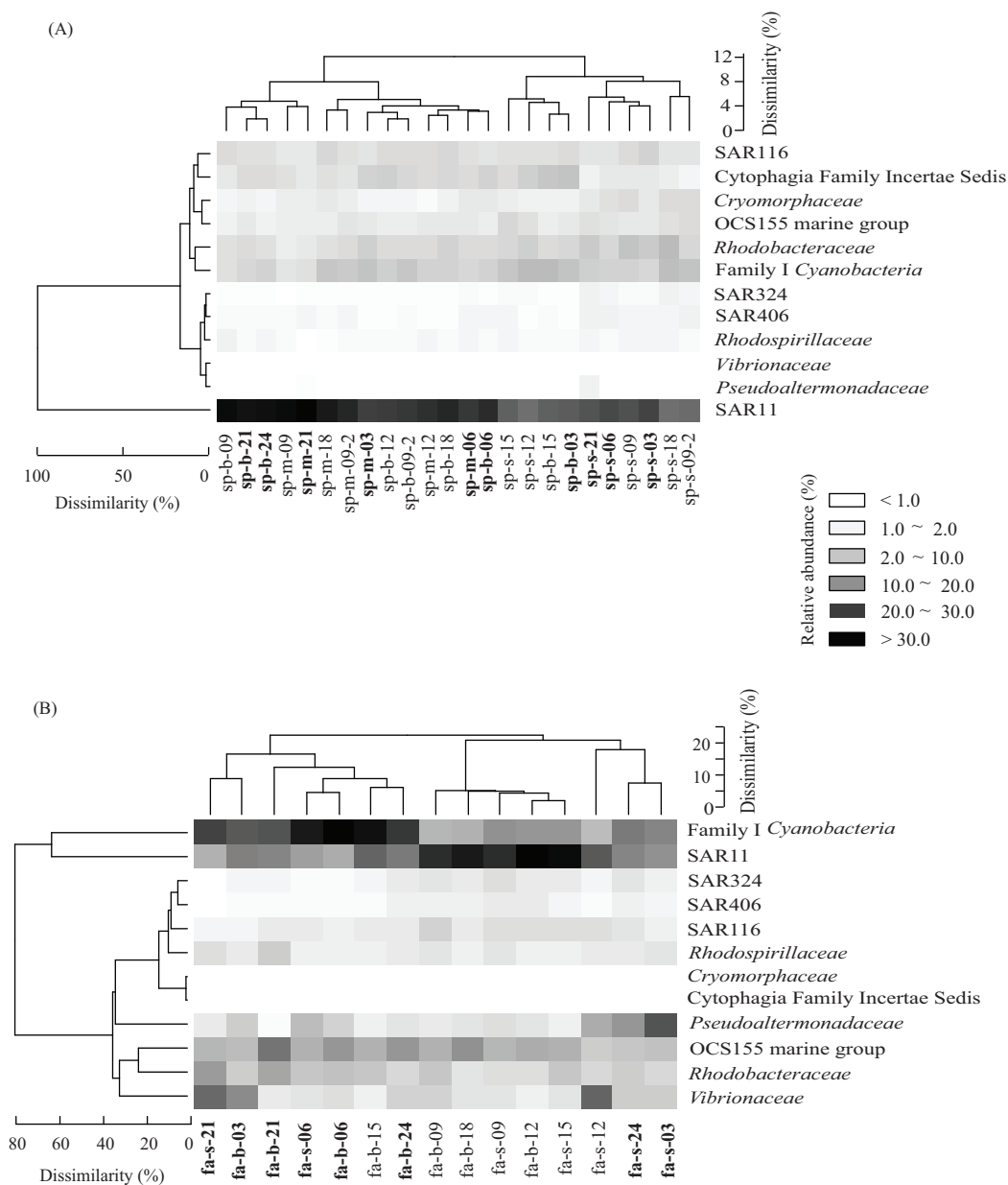


FIG 1 Heatmap and accompanying cluster analysis of the relative abundance (%) of major GRNMS bacterioplankton at the family level in (A) spring (April, 2011) and (B) fall (October, 2011) samples. Sample identifiers are based on sampling season (sp, spring; fa, fall), depth (s, surface; m, mid-depth; b, bottom), and time (in 24-hour format). Regular and bold font styles are used to denote day and night samples, respectively.

Fig. 2

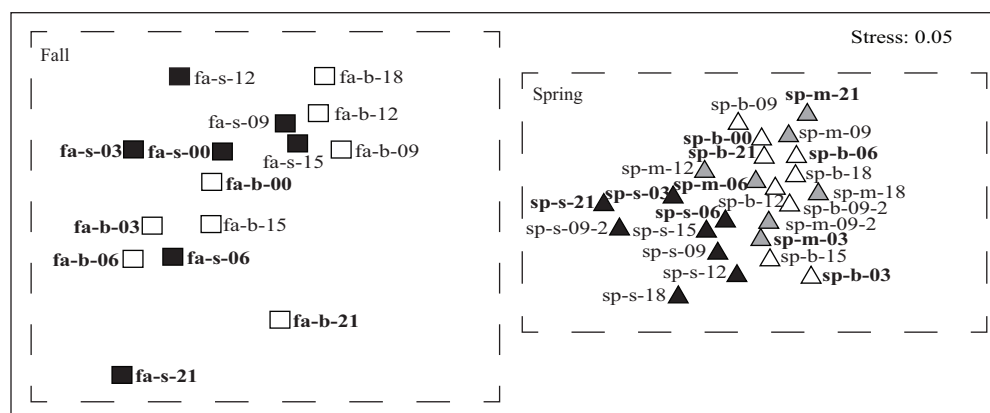


FIG 2 The non-metric multidimensional scaling ordination based on the relative abundance (%) of major GRNMS bacterioplankton at the family level in spring and fall samples. Sample identifiers are the same as in Fig. 1. Symbols are used to denote different sampling depths in spring (black triangle: surface; gray triangle: mid-depth; white triangle: bottom) and in fall (black square: surface; white square: bottom).

Fig. 3

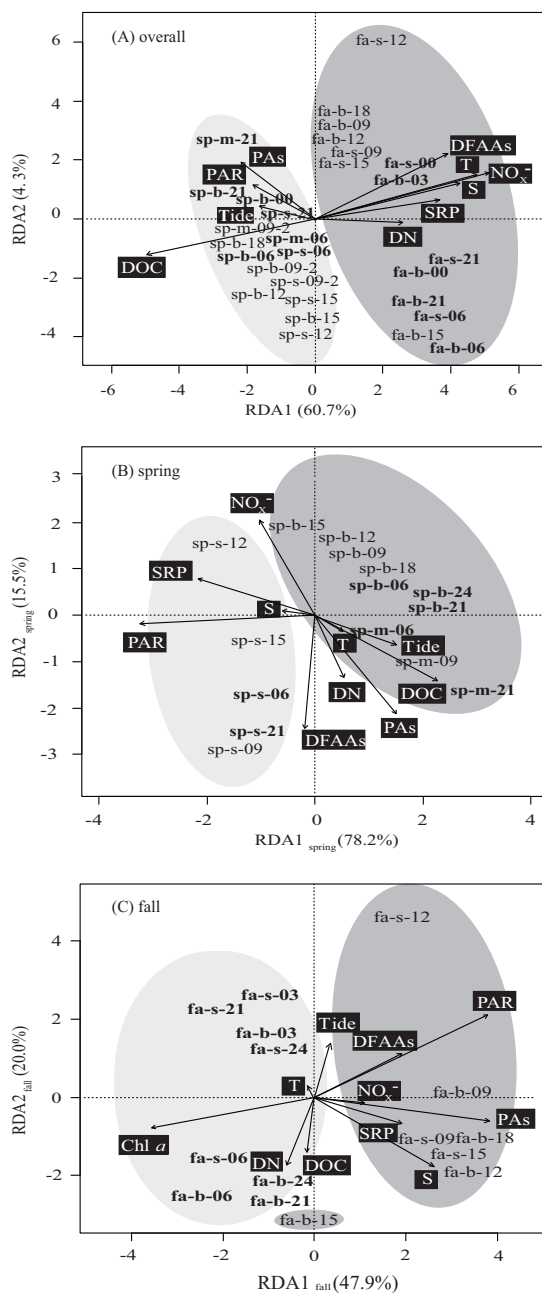


FIG 3 Biplot diagrams of the RDA analysis on the relative abundance (%) of BCC by environmental variables at the GRNMS for (A) overall, (B) spring, and (C) fall samples. Sample identifiers are the same as in Fig. 1. Environmental variables are in the black boxes with white font. Light and dark gray shadings are used to indicate significantly different BCC sample clusters identified by ANOSIM analysis.

1 Table 1. The average abundance and relative contribution of major taxa to the observed BCC
 2 differences among samples of the GRNMS.

Bacterial taxa	over-	Spring vs. Fall		Spring: Surface		Spring: Surface		Fall: day vs.	
	all	Avg sp/fa	δ_i	Avg s/m	δ_i	Avg s/b	δ_i	Avg D/N	δ_i
SAR11	30.7	37.6/21.5	24.0	31.6/42.1	29.2	31.6/39.5	29.2	28.9/15.5	22.4
Family I <i>Cyanobacteria</i>	13.6	9.6/20.1	16.5	10.2/8.8	8.6	10.2/9.6	8.7	15.4/23.8	18.1
<i>Rhodobacteraceae</i>	7.7	7.9/7.5	3.7	9.8/6.5	11.2	9.8/7.2	8.9	6.2/8.7	5.2
OCS155 marine group	7.5	4.9/11.5	9.6	6.1/4.6	5.2	6.1/4.2	6.3	11.0/11.5	5.4
SAR116	5.5	6.6/3.8	4.1	6.5/6.3	2.6	6.5/6.9	2.8	4.7/3.0	3.0
Cytophaga Family Incertae Sedis	4.3	6.8/0.3	9.6	5.3/6.5	7.4	5.3/8.4	12.0	1.7/0.9	1.8
<i>Vibrionaceae</i>	3.1	0.1/7.7	11.0	0.0/0.1	0.2	0.0/0.3	0.8	6.5/9.4	10.9
<i>Pseudoalteromonadaceae</i>	2.9	0.4/6.9	9.7	0.7/0.3	1.9	0.7/0.1	1.8	4.7/9.2	11.0
<i>Cryomorphaceae</i>	2.7	4.3/0.2	5.9	5.8/3.8	7.0	5.8/3.2	7.8	0.1/0.2	0.2
<i>Rhodospirillaceae</i>	2.7	2.1/3.6	2.2	2.3/1.8	2.1	2.2/1.9	1.4	3.2/4.0	2.2
SAR324	2.0	1.5/2.7	2.2	2.3/1.1	4.0	2.3/1.1	4.0	3.2/2.4	2.5
SAR406	2.0	2.1/1.9	1.4	2.6/2.0	2.9	2.6/1.6	3.8	2.4/1.6	2.2

3 ^a Avg, average relative abundance (in percentage); δ_i , estimated individual contributions (in
 4 percentage) of bacteria taxa to the observed between-sample differences. Shaded values are
 5 to label top four contributors for each comparison group. Sp, spring; fa, fall; s, surface water;
 6 b, bottom water; D, day; N, night.

7

8

9 Table 2. Correlations between major bacterial taxa and environmental variables based on
 10 Pearson's product momentum correlation coefficient. Variables with significant ($P < 0.05$
 11 with Bonferroni correction) correlations are in bold font.

Taxa	T	S	PAR	DOC	DN	NO _x ⁻	SRP	DFAAs	PAs
SAR11	-0.70	-0.58	-0.07	0.75	-0.27	0.52	-0.72	-0.34	0.14
Family I <i>Cyanobacteria</i>	0.64	0.56	-0.18	-0.64	0.30	-0.50	0.64	0.26	-0.15
Cytophaga Family Incertae Sedis	-0.86	-0.66	0.07	0.76	0.35	0.74	-0.79	-0.61	0.17
<i>Vibrionaceae</i>	0.64	0.55	0.09	-0.61	0.12	-0.60	0.53	0.40	-0.02
<i>Pseudoalteromonadaceae</i>	0.60	0.52	-0.02	-0.56	0.28	-0.60	0.61	0.48	0.01
<i>Rhodobacteraceae</i>	-0.02	-0.22	-0.09	-0.10	-0.24	0.10	0.01	-0.39	-0.36

12