Manuscript

Click here to access/download;Manuscript;Gronniger_Revisedtext.docx

Title: A Gulf Stream frontal eddy harbors a distinct microbiome compared to adjacent waters

Running Title: Gulf Stream frontal eddy microbiome

Jessica L. Gronniger[1], Patrick C. Gray[1], Alexandria K. Niebergall[2], Zackary I. Johnson[1,3], Dana E. Hunt[1,3*]

[1]Marine Laboratory, Duke University, Beaufort NC USA

[2]Nicholas School of the Environment, Duke University, Durham NC USA

[3]Biology and Civil & Environmental Engineering, Duke University, Durham NC USA

*To whom correspondence should be directed:

Dana Hunt

135 Duke Marine Lab Rd

Beaufort NC 28516

(252) 504-7542

Abstract

Mesoscale oceanographic features, including eddies, have the potential to alter productivity and other biogeochemical rates in the ocean. Here, we examine the microbiome of a cyclonic, Gulf Stream frontal eddy, with a distinct origin and environmental parameters compared to surrounding waters, in order to better understand the processes dominating microbial community assembly in the dynamic coastal ocean. Our microbiome-based approach identified the eddy as distinct from the surround Gulf Stream waters. The eddy-associated microbial community occupied a larger area than identified by temperature and salinity alone, increasing the predicted extent of eddy-associated biogeochemical processes. While the eddy formed on the continental shelf, after two weeks both environmental parameters and microbiome composition of the eddy were most similar to the Gulf Stream, suggesting the effect of environmental filtering on community assembly or physical mixing with adjacent Gulf Stream waters. In spite of the potential for eddy-driven upwelling to introduce nutrients and stimulate primary production, eddy surface waters exhibit lower chlorophyll *a* along with a distinct and less even microbial community, compared to the Gulf Stream. At the population level, the eddy microbiome exhibited differences among the cyanobacteria (e.g. lower *Trichodesmium* and higher *Prochlorococcus*) and in the heterotrophic alpha Proteobacteria (e.g. lower relative abundances of specific SAR11 clades) versus the Gulf Stream. However, better delineation of the relative roles of processes driving eddy community assembly will likely require following the eddy and surrounding waters since inception; additionally, sampling throughout the water column could better clarify the contribution of these mesoscale features to primary production and carbon export in the oceans.

Introduction

Marine microorganisms are of the engines of biogeochemical cycles, supporting critical ecosystem functions on a global scale; yet these biogeochemical processes depend on the composition, function, and activity of microbial communities (1). Marine microbiology has emphasized environmental filtering (how environmental factors determine microbiome composition due to bottom up effects), in determining community assembly, including factors such as water temperature, salinity and nutrient availability (2, 3). However, processes affecting microbiome community assembly include biological interactions, dispersal limitation and stochasticity (4), as well as the microbiome's ecological history (5). These historical contingencies, the effects of prior conditions on current microbiome assemblages, have emerged as a significant force shaping communities (6, 7). For example, prior environment changes can result in microbiomes that are more resistant (composition remains unchanged after the disturbance) or resilient (composition is altered but eventually returns to pre-disturbance state) to future perturbations (8, 9). Similarly, microbiomes from stable environments may be more sensitive to environmental changes (9, 10). However, in aquatic ecosystems, establishing antecedent environmental conditions is particularly challenging due factors including water movement, unmeasured disturbances and environmental parameter seasonality (11-13). Further, although dispersal limitation is assumed be reduced in aquatic systems compared to other systems, such as soils, oceanographic features such as steep frontal gradients between water parcels or vertical stratification can limit dispersal (14), and thus constrain microbiome composition. Here, we examine a frontal eddy, which may be able to isolate and transport

69 microbial populations (15), to investigate the relative roles of environmental filtering, ocean
70 physics and historical contingencies in an aquatic system.
71
72 Frontal eddies form along western boundary currents when instabilities in the current, often
73 stemming from interactions with bathymetry, cause meanders to form and evolve into distinct
74 cyclonically rotating water parcels. While considered mesoscale eddies, frontal eddies are
75 smaller and less stable than typical mesoscale eddies such as Gulf Stream rings (16). Gulf Stream
76 rings form downstream from Cape Hatteras once the current is freely meandering and can "pinch
77 off", in contrast, frontal eddies do not detach from the Gulf Stream and stay trapped in the
78 meander of the current. The eddy core is shelf water that is entrained in this Gulf Stream
79 meander and a shallow warm streamer flows from the downstream meander crest around this
80 core, separating the eddy from the shelf (17). All frontal eddies are cyclonic, and thus rotation
81 typically uplifts colder, nutrient-rich waters into the euphotic zone, often stimulating
82 phytoplankton growth, though continued uplift is dependent on the eddy's age (16-18). Beyond
83 frontal eddies, there has been substantial oceanographic interest in eddies generally, particularly,
84 their roles in temporal and spatial variability in primary production, carbon export to the deep
85 sea, and rapid shifts in marine microbiomes (19-21). Ecologically, eddy formation partially
86 isolates waters, with the potential to develop microbiomes, environmental parameters, and
87 biogeochemical rates distinct from surrounding environments (22, 23). However, it is difficult to
88 predict eddy microbiomes; the specific microbial taxa stimulated vary among the phytoplankton
89 and include diatoms, prasinophytes, dinoflagellates, and *Synechoccocus* (24-26), with the
90 potential for concurrent increases in multiple phytoplankton (26), and with likely consequences
91 for heterotrophic composition and organic matter export (24, 25, 27). Differences among eddy

microbiomes have been attributed to eddy characteristics, such as the initial microbiome, environmental factors (temperature, nutrient fluxes, mixed layer depth, etc.), and the eddy stage, among others (24, 28). Thus, although mesoscale eddies are pervasive throughout the global ocean, and frontal eddies are ubiquitous along western boundary currents, we have poor predictive ability of their influence on microbiomes and associated biogeochemical processes.

This study takes place just northeast of Cape Hatteras, NC, USA (Figure 1A) which is a dynamic point where the warm, salty, rapidly-flowing Gulf Stream, the cooler and fresher Slope Sea, and the even colder and fresher mid-Atlantic Bight shelf water meet, and where the Gulf Stream separates from the continental shelf. The study area is a region of high gradients, moving from the productive continental shelf waters to the oligotrophic open ocean in a few dozen kilometers across shore. Here, we compare a Gulf Stream frontal eddy's microbiome with those of the adjacent continental shelf, continental slope, and Gulf Stream waters off the coast of Cape Hatteras. Gulf Stream frontal eddies form ~every 3-7 days in the South Atlantic Bight (18), can occupy nearly half of the front between the Gulf Stream and the continental shelf in this region (29), and have been shown to play a significant role in nutrient transport into the euphotic zone, enhancing primary and possibly secondary production both in the eddy and on the continental shelf (17, 23, 30). Using a combination of field and satellite-based (temperature and chlorophyll *a*) measurements, we take a microbiome-centric approach to identify water parcels and population-level differences between water parcels in order to better understand the forces shaping the microbiomes of these mesoscale ocean features.

Figure 1. Locations and microbiome of field samples (A) Map of sampling sites for five ~10-15 km long ship-board transects off the coast of Cape Hatteras, North Carolina, USA September

5-7, 2021. Background color represents average sea surface temperatures from August 20th to September 10th obtained from the Geostationary Operational Environmental Satellite 16 (GOES-16). Each sampling point is shown along the transect and categorized by temperature and salinity (shapes) and microbial-community based k-means clusters (colors), generated using the self-organizing map (SOM) R package 'kohonen'. Isobaths for 500m (white), 1000m (grey) and 2000m (black) are included for context. (B) Non-metric multidimensional scaling (NMDS) ordination computed based on Bray-Curtis dissimilarity for 16S rRNA gene libraries. Ellipses show the multivariate t-distribution 95% confidence interval for the mean of each SOM cluster. Environmental factors found to be statistically significant (Multiple regression permutation test, $p<0.05$) are indicated with an '*'.

Results & Discussion

Using satellite data, we followed a single Gulf Stream frontal eddy from formation on the Charleston Bump until sampling ~2 weeks later (Figure 2). This sampled Gulf Stream frontal eddy was visible via satellite as a cooler, chlorophyll *a* enriched region relative to the surrounding Gulf Stream waters (31); however, there was no evidence of a surface bloom or enhancement of chlorophyll *a* levels compared to the shelf waters. In fact, satellite- measured surface chlorophyll *a* levels decline over time, likely either through nutrient depletion or mixing with Gulf Stream waters (Figure 2)(31). Here, we *in situ* sampled the eddy (and surrounding waters) as it moved past Cape Hatteras on September 5-7th 2021. At this point the eddy was approximately 20 by 50 km (longer in the along stream dimension) and we collected five across-stream transects of 10-15 km each (49 stations with 1-2 km between stations) to obtain high-resolution measurements of surface microbial community composition and environmental conditions (Figure 1A).

Figure 2. Satellite overview of the observed frontal eddy from formation to *in situ* sampling. Black dotted lines are all 5 transects from this study. Black boxes outline the eddy location across five days from August 27th to September 6th. SST is from GOES-16 and chlorophyll *a* is from Sentinel-3's OLCI sensor. Boxes were determined manually by identifying an anomaly in SST and ocean color measurements that is consistent with a cyclonic frontal eddy and associated warm streamer.

Initially, surface temperature and salinity were used to identify water parcels (Figure S1) (31). However, these physical parameter-based categories did not align with microbiome composition, particularly among the Gulf Stream and eddy samples (Figure S2A). Thus, we used a microbiome-based, self-organizing map (SOM) approach (32) to cluster the microbiome data independently of environmental variables. We then compared these biologically-based clusters (SOM) with the temperature and salinity-based water parcel classifications. Although these two schemes did overlap, there were also a number of disagreements in the assignments (Figure 1B): SOM cluster 1 largely corresponds to the continental shelf (87.5%), cluster 2 the continental slope (67%), cluster 3 the Gulf Stream (77%) and cluster 4 the eddy (62.5%) stations. Satellite data suggests that disagreements between these simple physical parameters and microbiome-based classifications occurred at transitions between water parcels, particularly around the periphery of the eddy (Figure S3), which could be due to complex mixing between the eddy and surrounding Gulf Stream. We found that the microbiome-based (SOM) approach classified more samples as eddy-associated. Therefore, identification of the eddy based on temperature and salinity alone could underestimate the geographic extent of the eddy's biogeochemical influence (Figure S3), or fail to identify mixing of the eddy microbiome into surrounding waters. Henceforth, for clarity, these statistically-distinct SOM clusters (ANOSIM, $p<0.05$) will be referred to by their most common location (1 = continental shelf, 2 = continental slope, 3 = Gulf Stream, 4 = eddy), with the assumption that SOM clusters most accurately map potential

169 ecological differences between the samples. However, disagreements between clustering based
170 on surface properties (temperature, salinity) and microbiomes merits further investigation in both
171 the biological and physical oceanographic fields, particularly at eddy margins to better identify
172 microbiome origins and implications (e.g. changes in biogeochemical rates).
173
174 Next, we examined how environmental factors differed among these clusters; the shelf and slope
175 clusters are also clearly distinct on the NMDS, with relatively higher overlap for the Gulf Stream
176 and eddy clusters (Figure 1B). We examined environmental variables that might help explain the
177 observed microbiome patterns, the continental shelf and slope exhibit higher cell abundances
178 (total prokaryotes, *Synechococcus*) and higher PO_4 along with lower salinity and cooler water
179 temperatures compared to the eddy and Gulf Stream (Figures 3, S4; Pairwise Wilcoxon ranked
180 sum test $p < 0.05$). This is generally consistent with higher resource availability on the
181 continental shelf and slope compared to the Gulf Stream and eddy, indicating greater terrestrial
182 influence and resources in near-shore environments and oligotrophic conditions beyond the shelf
183 break (33). Ordination of the microbiome in relation to select environmental variables
184 [temperature, salinity, chlorophyll *a*, nutrients (SiO_4, NO_3, NH_4, PO_4), and flow-cytometric
185 abundances of bacteria and cyanobacteria] identified significant associations for variables other
186 than SiO_4, NO_3 and NH_4, (Figure 1B, Table S1), highlighting the factors which could explain the
187 microbiome composition. Here, we specifically focused on comparison of the Gulf Stream and
188 eddy samples as they are most similar in both environment and microbiome: the eddy is ~1 °C
189 cooler, has lower chlorophyll *a*, lower *Synechococcus* and higher non-phycoerythrin containing
190 picocyanobacteria (hereafter labeled *Prochlorococcus*) concentrations (Figures 3, S4; Pairwise
191 Wilcoxon ranked sum test $p< 0.05$). While PO_4 in the eddy is statistically higher than in the Gulf

Stream, eddy phosphate concentrations have a large range, suggesting potential variability in nutrient fluxes/ availability within this water parcel (Figure 3D). Thus, in contrast to other cyclonic eddies (23, 34), here, eddy surface waters generally exhibit oligotrophic characteristics (low nutrients, high *Prochlorococcus* concentrations, lower chlorophyll *a,* but higher PO_4) compared to the Gulf Stream, in spite of forming via entrainment of presumably higher nutrient continental shelf water and the potential for upwelling-driven nutrient additions (Figures 3, S4). Overall, these results are consistent with the satellite measurements that do not show enhanced surface photosynthetic biomass of the eddy. Possible explanations for this result include that photosynthetic biomass enhancement occurred primarily below the mixed layer and was not visible at the surface, or that high stratification in the summer along with a deep euphotic depth prevented the eddy-associated uplift from enhancing productivity. Examination of microbial community characteristics (diversity etc.) as well as specific phylotypes associated with each water parcel may offer additional clues to factors that shaped these communities in the coastal ocean.

Figure 3. Key environmental factors for Self-Organizing Map (SOM) clusters. Numbered clusters correspond to the most common environments: Continental shelf (Cluster 1), Continental slope (Cluster 2), Gulf Stream (Cluster 3), Eddy (Cluster 4). For the box and whisker plots, the center lines denote the median value while the box encompasses the 25th to 75th percentiles of the dataset. Whiskers denote the 5th and 95th percentiles and values that fall outside these ranges are shown as individual points. Brackets indicate p-values of Wilcoxon Rank Sum pairwise comparisons. Additional environmental parameter comparisons are shown in Figure S4.

Microbial Assemblages

As eddies can alter microbial community parameters, potentially increasing diversity through mixing of water parcels and influx of resources, or conversely reducing diversity and evenness

through competitive exclusion (20, 35-37), we compared diversity and evenness indexes in the four water parcels. Notably, the eddy exhibits lower richness and evenness compared to the Gulf Stream (Figure 4; Wilcoxon ranked sum test $p<0.05$), reflecting the dominance of specific taxa, particularly *Prochlorococcus* (35) (Figure 5). While the eddy initially formed from continental shelf waters (although off the coast of South Carolina) and retains comparable richness to shelf waters in these transects (Figure 4), the eddy is statistically less even than the other water parcels ~two weeks after formation, suggesting continued development of the microbiome (Wilcoxon ranked sum test $p<0.05$). While the eddy microbiome is most similar to the Gulf Stream's, if dispersal limitation significantly shapes the eddy microbiome and if we assume the shelf community where the eddy formed is similar to the shelf community off Cape Hatteras, we might predict that presence-absence data (Sorenson index) would group the eddy with the continental shelf and that relative abundances would shift to match the current environmental conditions. However, the Sorenson index also groups the eddy and the Gulf Stream samples (Figure S5), as was observed for Bray-Curtis dissimilarity (Figure 1B), suggesting that environmental filtering or physical mixing shapes frontal eddy communities.

Figure 4. Alpha diversity metrics for 16S rRNA gene community composition for each Self-Organizing Map (SOM) cluster. (A) Shannon's H index, (B) species richness, and (C) Pielou's evenness were calculated based on the absolute abundance of ASVs using 'vegan' (v2.6.2) in R (v4.0.0). SOM cluster numbers correspond to the following classifications: 1 –continental shelf, 2-continental slope, 3- Gulf Stream, 4- eddy. Brackets indicate p-values associated with Wilcoxon Rank Sum test pairwise comparisons.

Figure 5. Average family-level composition for each Self-Organizing Map (SOM) cluster. The average community was calculated by taking the mean relative abundance of each taxon across all samples within that cluster, then normalizing relative abundances to 100%. ASVs with a relative abundance of less than 1% are grouped as "Other." Amplicon Sequence Variants (ASVs were generally grouped at the family level, but when family is undefined they are labeled

at the phylum level. Additionally, the family *Synechococcaceae* was resolved at the genus level to illustrate the *Synechococcus*-*Prochlorococcus* inversion between clusters.

Across all samples, the microbiome composition reflects abundant coastal microbes including *Synechococcaceae* (~25-40%), *Pelagibacteraceae* (~28-40%), Actinobacteria family OCS155 (~7%), the Archaeal Marine Group II clade (~2-4%) and the SAR406 clade A714017 (~2%; Figure 5) (3, 33, 38, 39). Yet, we observed onshore-offshore spatial transitions (33), including the switch within *Synechococcaceae* genera, with *Synechococcus* dominating the shelf and slope, and *Prochlorococcus* the Gulf Stream and eddy (Figure 5). As specific taxa may offer additional clues about the factors which shaped these microbiomes, we examined ASV relationships with specific clusters, using 250 most abundant ASVs (representing ~73% of the microbial community) with both linear discriminant effect size analysis (LEfSe) and a Bayesian generalized joint attribute model (GJAM) (40, 41) (Tables S2, S3). Both LEfSe and GJAM identify taxa associated with specific water parcels (Figure 6); moreover, these approaches generally agree: identifying increased relative abundance of *Pelagibacter* and *Synechococcus* in continental shelf and slope waters and enrichment of *Prochlorococcus* and *OCS155* in Gulf Stream and eddy waters, which generally aligns with expectations of oligotrophic, warm water-associated taxa offshore (33). GJAM identified a single difference between the eddy and Gulf Stream associated taxa, a High Light clade I *Prochlorococcus* ASV (Tables S2,S3; Figure 6B) that was enriched in both the eddy and slope samples, potentially suggesting either a continued legacy of nearshore waters or mixing of slope waters into the eddy (Figure 6B). Although there are similarities between the eddy and Gulf Stream samples, their differences help identify the

271 current (e.g. environmental filtering) and antecedent (e.g. historical contingencies) conditions

272 that shape the microbiome.

273

274 **Figure 6. Amplicon sequence variant (ASV) associations with specific SOM groupings** (A)
275 Heatmap of ln(absolute abundance+1) of the 22 ASVs identified with at least one significant
276 SOM cluster beta coefficient obtained from generalized joint attribute modeling (GJAM)
277 analysis of the 250 most abundant ASVs across all samples. Samples are organized by SOM
278 grouping (labeled at top) (B) Heatmap of statistically significant SOM cluster-associated beta
279 coefficients (95% confidence interval does not overlap 0), non-significant beta coefficients are
280 shown as white. Colored labels at the top of each column identifies the SOM cluster (as in panel
281 A). ASVs are organized maximum likelihood phylogenetic tree constructed using Smart Model
282 Selection in PhyML 3.0 with taxonomic assignments based on the RDP naïve Bayesian classifier
283 using the Greengenes version 13.5 database.
284

285 A comparison of just the eddy and Gulf Stream using both LEfSe and DESeq2 (42) more clearly

286 reveals distinctions between their microbiomes (Table S4; Figure 7). A number of SAR11

287 lineages were enriched in the Gulf Stream compared to the eddy (Figure 7). The SAR11 clade is

288 known to contain streamlined heterotrophs (43); however, we can only speculate that ASV

289 differences could be due to resource differences or stochastic effects. Several high-light clade

290 *Prochlorococcus* ASVs were enriched in the eddy over the Gulf Stream which is consistent with

291 flow-cytometric non-phycoerythrin containing picocyanobacteria enrichment in the eddy (Figure

292 3F) and their role as efficient nutrient recyclers (44). Interestingly, although inorganic nutrients

293 are low in both the eddy and Gulf Stream (Figures 3, S4), the Gulf Stream samples are enriched

294 in a putatively nitrogen-fixing cyanobacterium *Trichodesmium* ASV (Figure 7) (45). Thus,

295 within the cyanobacteria and despite similar inorganic nitrogen concentrations, the Gulf Stream

296 and eddy microbiomes may reveal differences in nutrient fluxes or the role of historical

297 contingencies in microbiome composition (46). Moreover, we acknowledge that standing stocks

298 (e.g. nutrient concentrations) are not equivalent to nutrient turnover rates, as resources can be

299 rapidly cycled in oligotrophic conditions (47). The development of the colonial, nitrogen-fixing
300 *Trichodesmium* may require additional time or lower nitrogen fluxes. As the eddy sustains higher
301 concentrations of *Prochlorococcus* and total bacterioplankton, and statistically higher levels of
302 PO_4, this data could indicate that the eddy has experienced higher nutrient fluxes, which could
303 support higher primary and secondary production, although this is not evident in standing
304 photosynthetic biomass (as measured by chlorophyll *a*, Figure 3C). Although there was no
305 satellite-detectable eddy surface bloom, and the eddy microbiome likely shifted toward that of
306 the Gulf Stream over time, microbial communities can offer clues about the processes occurring
307 in these dynamic mesoscale features.
308
309 **Figure 7. Microbiome differences between the Gulf Stream and eddy.** Left panel: Phylotypes
310 (ASVs) that exhibited significant differential abundance in the eddy relative to the Gulf Stream
311 as identified using DESeq2, left side indicates enrichment in the Gulf Stream, right side eddy.
312 Dots indicates the Log_2 fold changes for each ASV identified as significant (Benjamin-Hochberg
313 multiple hypothesis adjusted $p<0.05$), error bars indicate the standard error of the Log_2 fold
314 changes as calculated by DESeq2. Only ASVs with a minimum average relative abundance of
315 0.05% across the samples being compared were included in the analysis. Y-axis represents
316 maximum likelihood phylogenetic tree constructed using Smart Model Selection in PhyML 3.0.
317 Taxonomic assignments were obtained from the RDP database. Right hand column indicates
318 those ASVs that were also found to have a significant association with either the Gulf Stream
319 (dark blue) or the eddy (light blue) as identified using Linear discriminant analysis Effect Size
320 (LEfSE).
321
322
323
324 **Conclusions**
325
326 Here we show that a cyclonic, frontal eddy that formed along the western boundary of the Gulf
327 Stream harbored a microbial community distinct from those of surrounding waters. This eddy
328 community appears to be shaped by either environmental selection or mixing with the Gulf

329 Stream (16, 29), rather than dispersal limitation, as it was most similar to that of the Gulf Stream
330 rather than continental slope waters where it originated. While this study was limited to surface
331 samples and encompasses a brief period in the eddy's development, new technologies are
332 enabling Lagrangian sampling of eddy parameters during development and throughout the water
333 column to better quantify upwelling, eddy coherence and mixing with the Gulf Stream, across
334 depths (48-51). High-resolution, *in situ* sampling will enable the field to develop a more
335 quantitative assessment of how source populations and environmental drivers interact to shape
336 these dynamic mesoscale features over time. Although more research is needed to understand the
337 impact of eddies globally, this study deepens our understanding of the forces shaping Gulf
338 Stream frontal eddy microbiomes off the South Atlantic Bight and provides important directions
339 for future studies. While there was not support for this eddy contributing to the biodiversity
340 hotspot located off of Cape Hatteras (36, 52), other frontal eddies may enhance biodiversity by
341 delivering distinct microbial communities, enhanced zooplankton populations, and upwelling of
342 resources (31). Furthermore, the frequency of these eddies and their predicted nutrient injections
343 into the Gulf Stream may play a significant role in carbon sequestration and productivity along
344 the South Atlantic Bight. Although we found this eddy harbored a distinct microbial community,
345 further investigations will be needed to better understand the role of historical contingencies and
346 the temporal dynamics of eddy-associated communities and their potential role in global
347 biogeochemical cycles. Moreover, this approach of microbially-defined, environmental variable
348 agnostic sample clustering can be applied more broadly in oceanography to help disentangle
349 potentially ecologically-distinct microbiomes.
350
351 **Methods**

352

353 **Sample collection and sea water parameter characterization**

354 We tracked via satellite a cyclonic, frontal eddy that formed from shelf water off of Charleston,
355 South Carolina, USA on August 25-26th 2021, was isolated by warmer, Gulf Stream-derived
356 waters, and then traveled north along the Gulf Stream's western front (31). Seawater was
357 collected on board the *R/V Shearwater* on 5 transects from September 5-7th, 2022; transects were
358 planned to cross the front between the shelf and eddy or shelf and Gulf Stream waters. Water
359 samples were collected during transit from the ship's flow-through seawater system (~1 m
360 depth). In general, the transects were 10-15 km long (over 2-3 hours) with 14-21 samples
361 collected per transect at approximately 10 min intervals for a total of 86 sampling points, 49 of
362 which included filtration for later nucleic acid analysis. Temperature and salinity were measured
363 with a SeaBird SBE38 Thermosalinograph. Chlorophyll *a* concentrations were measured by
364 filtering 100 mL of seawater onto a ~0.7 µm glass fiber filter (Whatmann APFF02500) using
365 gentle (<10mm Hg) vacuum. Chlorophyll *a* pigments were extracted in 100% methanol for 48h
366 at -20 °C and fluorescence was measured using a Turner Designs 10-AU fluorometer as
367 previously described (53). Duplicate whole seawater samples for flow-cytometric analyses were
368 collected, fixed with net 0.125% glutaraldehyde and stored at -80 °C until processing.
369 Prokaryotic phytoplankton populations were enumerated using a Becton Dickinson
370 FACSCalibur Flow Cytometer and categorized as previously described (53). Bacterioplankton
371 were quantified using SYBR Green-I on the Attune NxT acoustic flow cytometer (Life
372 Technologies) (54). Duplicate 0.22 µm Sterivex (Milipore) filtered seawater samples were
373 collected for dissolved inorganic nutrient quantification and stored at -80 °C until processing.
374 Nutrients (NO_2, NO_3, PO_4, SiO_4) were analyzed at the UC San Diego Scripps Institution of

375 Oceanography ODF chemical laboratory using a Seal Analytical continuous-flow AutoAnalyzer
376 3 (55, 56). Detection limits for the nutrients analyzed are as follows: NO_2 = 20nM, NO_3 = 20nM,
377 PO_4 =20nM, SiO_4 = 200nM. No permits were required for environmental sampling within US
378 coastal waters.
379
380 **DNA extraction and library preparation:**
381 Microbial DNA was collected by filtering ~4 liters of whole seawater through a 0.22 micron
382 Sterivex filter (Millipore). Filters were stored at -20 °C while on the research vessel and then at -
383 80 °C upon return to the lab until extraction. Genomic DNA for 16S rRNA gene libraries was
384 extracted using the Gentra Puregene Yeast/Bacteria kit (QIAGEN) supplemented with 60
385 seconds of bead beating. Extracted DNA was quantified using a Nanodrop ND-100 and 16S
386 rRNA gene libraries were generated using gene primers targeting the V4-V5 region 515F- (5′-
387 GTGYCAGCMGCCGCGGTAA) and 926R (5′-CCGYCAATTYMTTTRAGTTT) (57, 58).
388 PCR reactions were performed in triplicate with 20 µL reactions containing 20 ng template
389 DNA, 1× Taq Buffer, 0.5 µM of each primer, 200 µM of dNTPs, and 0.4 U of Q5 DNA
390 polymerase (NEB). The thermal cycling conditions were 30 sec at 98 °C, followed by 25 cycles
391 of 10 sec at 98 °C, 30 sec at 50 °C, 30 sec at 72 °C, and a final extension at 72 °C for 2 min.
392 Triplicate PCR reactions were pooled and gel purified. Libraries were pooled and sequenced at
393 the Duke Center for Genomic and Computational Biology using V2 2 × 250 bp sequencing on
394 the Illumina MiSeq.
395
396 **Sequence Processing:**

397 Barcodes were removed and sequences were demultiplexed and assigned to samples using Sabre
398 (https://github.com/najoshi/sabre). Sequences were cleaned and clustered using VSEARCH
399 v2.5.1 (59). Low quality sequence ends were trimmed at the Phred quality score (Q) of 30 using
400 a 10 bp running window. Paired-end reads with 10 bp overlap and no mismatches were merged
401 and sequences with expected errors >1 and/or a length < 360 bp were removed. Minimum
402 entropy decomposition analysis (v2.1) was used to resolve amplicon sequences variants (ASVs)
403 (60). Sequencing-error associated noise was reduced by retaining only those ASVs with unique
404 sequences with a minimum abundance of 20 occurrences. Samples contained an average of
405 39,936 representative reads. Representative ASV sequences were assigned taxonomy using the
406 RDP classifier in MacQIIME v1.9.1. Mitochondrial sequences were removed and libraries were
407 subsampled to 14,187 reads per library. AmpliCopyrighter (v.0.46) was used to correct for
408 differential gene copy numbers across taxa (61). "Absolute abundance" was calculated by
409 multiplying relative abundances of ASVs by prokaryotic cell counts obtained from flow
410 cytometry analysis. NMDS plots were generated by calculating Bray-Curtis dissimilarity scores
411 based on ASV relative abundances using the *vegdist()* function in R.
412
413 **Self-Organizing Map**
414 The Kohonen package (32) in R (v. 4.0.0) was used to cluster samples based on microbial
415 community composition. A self-organizing map was constructed using ASV relative abundances
416 and each sample was assigned to a map unit based on community similarities. Our map was
417 composed of 4x4 circular nodes in a hexagonal, nontoroidal configuration. Clusters were
418 generated using k-means clustering and a reasonable value for k was chosen by estimating the
419 point of inflection in a scree plot of within-clusters sum of squares (Figure S6).

420

421 **Discriminative taxa identification**

422 Generalized joint attribute modelling (GJAM) was applied to model the 250 most abundant
423 ASVs and select environmental factors (temperature, salinity, chlorophyll *a*, PO_4), using the
424 GJAM v. 2.6.2 package in R. Iteration was set at 20,000 and burning at 5,000. Results were
425 visualized using the built-in function 'gjamPlot'. In order to identify discriminative taxa between
426 SOM clusters, the 250 most abundant ASVs were analyzed with LEfSe (41). The threshold for
427 significance on the logarithmic LDA score for discriminative features was two. P-values were
428 obtained from the Kruskall-Wallis test for differences in ASV abundances among SOM cluster
429 assignments. These two metrics differ in approach: LEfSe identifies phylotype enrichment in a
430 specific cluster while GJAM can identify the magnitude and directionality (positive or negative)
431 of associations across multiple clusters. ASVs exhibiting differential abundances between the
432 Gulf Stream and eddy were further identified using DESeq2 (42). Only ASVs with an average
433 relative abundance >0.05% averaged across samples being compared (Gulf Stream and eddy)
434 were included. Significant differences were considered those with a Benjamin-Hochberg
435 multiple hypothesis adjusted $p<0.05$. Quantification of the processes shaping microbiome
436 community assembly using Beta Mean Nearest Taxon Distance (βMNTD) (5) was not possible
437 because our data did not exhibit the required phylogenetic signal (non-significant Mantel
438 correlogram).

439

440 **Satellite Data**

441 Satellite data from two sources is used in this work. Chlorophyll *a* products were derived from
442 ocean color imagery acquired by the European Commission Copernicus programme's Sentinel-3

Ocean and Land Color Instrument (OLCI) using the CHL_OC4ME blue to green band ratio algorithm (62). The OLCI data was used to track surface chlorophyll *a* within the eddy over its lifetime. Sea surface temperature products were from the Geostationary Operational Environmental Satellite 16 (GOES-16) which uses a retrieval from the long wave infrared bands (8µm-12µm) (63). OLCI Level 2 and obtained from https://coda.eumetsat.int and GOES-16 hourly SST was acquired from https://cwcgom.aoml.noaa.gov/erddap/griddap/goes16SSThourly.html.

Physical Water Parcel Delineations

Temperature and salinity delineations between water parcels are specific to this region and time. They are based on satellite tracking of the eddy water mass and known Temperature-Salinity properties of these waters. In general Gulf Stream waters are the hottest and saltiest, South Atlantic Bight waters are slightly fresher and cooler, Mid Atlantic Bight slope waters are cooler and fresher again, and Mid Atlantic Bight shelf water is the coolest and freshest (31, 64). To determine the exact thresholds of temperature and salinity these general patterns were informed and reinforced by observed sharp frontal gradients between water masses. This led to the following thresholds: Gulf Stream > 28.2 °C, eddy <28.2 °C and > 35.75 PSU, continental slope <35.75 PSU and > 34 PSU, and continental shelf <34 PSU (Figure S1) (31).

Data and Code Repository: Cruise field data is deposited: doi.org/10.5281/zenodo.7680135 and SSU rRNA library sequences are deposited as part of Bioproject PRJNA868376. Analysis scripts are available at https://github.com/jgronniger/Gulf-Stream-Eddy-2021.

Acknowledgements:

We acknowledge the assistance of the captain and crew of the *R/V Shearwater.*

CRediT Contributions:

Jessica Gronniger- Conceptualization; Data Curation; Formal Analysis; Visualization; Writing-original draft

Patrick C. Gray- Conceptualization; Data Curation; Formal Analysis; Visualization; Writing-review and editing

Alexandria K. Niebergall- Formal Analysis; Writing-review and editing

Zackary I. Johnson- Resources; Writing-review and editing

Dana E. Hunt- Conceptualization; Resources; Writing-original draft;

References

1. Falkowski PG, Fenchel T, Delong EF. The microbial engines that drive Earth's biogeochemical cycles. Science. 2008;320(5879):1034-9.
2. Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, et al. Structure and function of the global ocean microbiome. Science. 2015;348(6237):1261359.
3. Ward CS, Yung C-M, Davis KM, Blinebry SK, Williams TC, Johnson ZI, et al. Annual community patterns are driven by seasonal switching between closely related marine bacteria. The ISME Journal. 2017;11(6):1412-22.
4. Zhou J, Deng Y, Zhang P, Xue K, Liang Y, Van Nostrand JD, et al. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. Proceedings of the National Academy of Sciences. 2014;111(9):E836-E45.
5. Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, et al. Quantifying community assembly processes and identifying features that impose them. The ISME Journal. 2013;7(11):2069.
6. Hawkes CV, Keitt TH. Resilience vs. historical contingency in microbial responses to environmental change. Ecology Letters. 2015;18(7):612-25.
7. Stegen JC, Bottos EM, Jansson JK. A unified conceptual framework for prediction and control of microbiomes. Current Opinion in Microbiology. 2018;44:20-7.

8. Jurburg SD, Nunes I, Brejnrod A, Jacquiod S, Priemé A, Sørensen SJ, et al. Legacy effects on the recovery of soil bacterial communities from extreme temperature perturbation. Frontiers in Microbiology. 2017;8:1832.
9. Waring BG, Hawkes CV. Short-term precipitation exclusion alters microbial responses to soil moisture in a wet tropical forest. Microbial Ecology. 2015;69:843-54.
10. Wang Z, Tsementzi D, Williams T, Juarez D, Garcia N, Johnson Z, et al. Environmental stability impacts the differential sensitivity of marine microbiomes to increases in temperature and acidity. The ISME Journal. 2021;15(1):19-28.
11. Gronniger JL, Wang Z, Brandt GR, Ward CS, Tsementzi D, Mu H, et al. Rapid changes in coastal ocean microbiomes uncoupled with shifts in environmental variables. Environmental Microbiology. 2022.
12. Evans S, Martiny JB, Allison SD. Effects of dispersal and selection on stochastic assembly in microbial communities. The ISME Journal. 2017;11(1):176-85.
13. Shen D, Langenheder S, Jürgens K. Dispersal modifies the diversity and composition of active bacterial communities in response to a salinity disturbance. Frontiers in Microbiology. 2018;9:2188.
14. Bracco A, Provenzale A, Scheuring I. Mesoscale vortices and the paradox of the plankton. Proceedings of the Royal Society of London Series B: Biological Sciences. 2000;267(1454):1795-800.
15. Villar E, Farrant GK, Follows M, Garczarek L, Speich S, Audic S, et al. Environmental characteristics of Agulhas rings affect interocean plankton transport. Science. 2015;348(6237):1261447.
16. Gula J, Molemaker MJ, McWilliams JC. Submesoscale dynamics of a Gulf Stream frontal eddy in the South Atlantic Bight. Journal of Physical Oceanography. 2016;46(1):305-25.
17. Lee TN, Yoder JA, Atkinson LP. Gulf Stream frontal eddy influence on productivity of the southeast US continental shelf. Journal of Geophysical Research: Oceans. 1991;96(C12):22191-205.
18. Glenn SM, Ebbesmeyer CC. Observations of Gulf Stream frontal eddies in the vicinity of Cape Hatteras. Journal of Geophysical Research: Oceans. 1994;99(C3):5047-55.
19. Falkowski PG, Ziemann D, Kolber Z, Bienfang PK. Role of eddy pumping in enhancing primary production in the ocean. Nature. 1991;352(6330):55-8.
20. Clayton S, Dutkiewicz S, Jahn O, Follows MJ. Dispersal, eddies, and the diversity of marine phytoplankton. Limnology and Oceanography: Fluids and Environments. 2013;3(1):182-97.
21. Martin-Platero AM, Cleary B, Kauffman K, Preheim SP, McGillicuddy DJ, Alm EJ, et al. High resolution time series reveals cohesive but short-lived communities in coastal plankton. Nature Communications. 2018;9(1):266.
22. Sweeney EN, McGillicuddy Jr DJ, Buesseler KO. Biogeochemical impacts due to mesoscale eddy activity in the Sargasso Sea as measured at the Bermuda Atlantic Time-series Study (BATS). Deep Sea Research Part II: Topical Studies in Oceanography. 2003;50(22-26):3017-39.
23. Baltar F, Arístegui J, Gasol JM, Lekunberri I, Herndl GJ. Mesoscale eddies: hotspots of prokaryotic activity and differential community structure in the ocean. The ISME Journal. 2010;4(8):975-88.

24. Doblin MA, Petrou K, Sinutok S, Seymour JR, Messer LF, Brown MV, et al. Nutrient uplift in a cyclonic eddy increases diversity, primary productivity and iron demand of microbial communities relative to a western boundary current. PeerJ. 2016;4:e1973.
25. Lochte K, Pfannkuche O. Cyclonic cold-core eddy in the eastern North Atlantic. II. Nutrients, phytoplankton and bacterioplankton. Marine Ecology Progress Series. 1987:153-64.
26. Wang L, Huang B, Chiang K-P, Liu X, Chen B, Xie Y, et al. Physical-biological coupling in the western South China Sea: The response of phytoplankton community to a mesoscale cyclonic eddy. PloS One. 2016;11(4):e0153735.
27. Ahern OM, Whittaker KA, Williams TC, Hunt DE, Rynearson TA. Host genotype structures the microbiome of a globally dispersed marine phytoplankton. Proceedings of the National Academy of Sciences. 2021;118(48).
28. Harke MJ, Frischkorn KR, Hennon GM, Haley ST, Barone B, Karl DM, et al. Microbial community transcriptional patterns vary in response to mesoscale forcing in the North Pacific Subtropical Gyre. Environmental Microbiology. 2021;23(8):4807-22.
29. Gula J, Molemaker MJ, McWilliams JC. Gulf Stream dynamics along the southeastern US seaboard. Journal of Physical Oceanography. 2015;45(3):690-715.
30. Paffenhöfer G, Sherman B, Lee T. Summer upwelling on the southeastern continental shelf of the USA during 1981. Zooplankton abundance and distribution. Prog Oceanog. 1987;19:403-36.
31. Gray PC, Gronniger J, Savelyev I, Dale J, Niebergall AK, Cassar N, et al. Investigating a Gulf Stream Frontal Eddy North of Cape Hatteras: Questioning Overall Impacts on Ecology. bioRxiv. 2023:2023.02. 22.529409.
32. Wehrens R, Buydens LM. Self-and super-organizing maps in R: the Kohonen package. Journal of Statistical Software. 2007;21:1-19.
33. Wang Z, Juarez DL, Pan JF, Blinebry SK, Gronniger J, Clark JS, et al. Microbial communities across nearshore to offshore coastal transects are primarily shaped by distance and temperature. Environmental Microbiology. 2019;21(10):3862-72.
34. Benitez-Nelson CR, Bidigare RR, Dickey TD, Landry MR, Leonard CL, Brown SL, et al. Mesoscale eddies drive increased silica export in the subtropical Pacific Ocean. Science. 2007;316(5827):1017-21.
35. Lévy M, Jahn O, Dutkiewicz S, Follows MJ, d'Ovidio F. The dynamical landscape of marine phytoplankton diversity. Journal of The Royal Society Interface. 2015;12(111):20150481.
36. Barton AD, Dutkiewicz S, Flierl G, Bragg J, Follows MJ. Patterns of Diversity in Marine Phytoplankton. Science. 2010;327(5972):1509-11.
37. Pajares S. Unraveling the distribution patterns of bacterioplankton in a mesoscale cyclonic eddy confined to an oxygen-depleted basin. Aquatic Microbial Ecology. 2021;87:151-66.
38. Chow C-ET, Sachdeva R, Cram JA, Steele JA, Needham DM, Patel A, et al. Temporal variability and coherence of euphotic zone bacterial communities over a decade in the Southern California Bight. The ISME Journal. 2013;7:2259–73.
39. Yilmaz P, Yarza P, Rapp JZ, Glöckner FO. Expanding the world of marine bacterial and archaeal clades. Frontiers in Microbiology. 2016;6:1524.
40. Clark JS, Nemergut D, Seyednasrollah B, Turner PJ, Zhang S. Generalized joint attribute modeling for biodiversity analysis: Median- zero, multivariate, multifarious data. Ecological Monographs. 2017;87(1):34-56.

41. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biology. 2011;12(6):R60.
42. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014;15(12):1.
43. Giovannoni SJ. SAR11 Bacteria: The Most Abundant Plankton in the Oceans. Annual Review of Marine Science. 2017;9(1):231-55.
44. Polz MF, Hunt DE, Preheim SP, Weinreich DM. Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. Philosophical Transactions of the Royal Society B-Biological Sciences. 2006;361(1475):2009-21.
45. Palter JB, Ames EJ, Benavides M, Goncalves Neto A, Granger J, Moisander PH, et al. High N 2 fixation in and near the Gulf Stream consistent with a circulation control on diazotrophy. Geophysical Research Letters. 2020;47(16):e2020GL089103.
46. Conway TM, Palter JB, de Souza GF. Gulf Stream rings as a source of iron to the North Atlantic subtropical gyre. Nature Geoscience. 2018;11(8):594-8.
47. Benitez-Nelson CR, Buesseler KO. Variability of inorganic and organic phosphorus turnover rates in the coastal ocean. Nature. 1999;398(6727):502-5.
48. Ntaganou N, Kourafalou V, Beron-Vera FJ, Olascoaga MJ, Le Hénaff M, Androulidakis Y. Influence of Caribbean eddies on the Loop current system evolution. Frontiers in Marine Science. 2023;10:961058.
49. Calil P, Richards K. Transient upwelling hot spots in the oligotrophic North Pacific. Journal of Geophysical Research: Oceans. 2010;115(C2).
50. Tarshish N, Abernathey R, Zhang C, Dufour CO, Frenger I, Griffies SM. Identifying Lagrangian coherent vortices in a mesoscale ocean model. Ocean Modelling. 2018;130:15-28.
51. Zhang Y, Ryan JP, Hobson BW, Kieft B, Romano A, Barone B, et al. A system of coordinated autonomous robots for Lagrangian studies of microbes in the oceanic deep chlorophyll maximum. Science Robotics. 2021;6(50):eabb9138.
52. Byrd BL, Hohn AA, Lovewell GN, Altman KM, Barco SG, Friedlaender A, et al. Strandings as indicators of marine mammal biodiversity and human interactions off the coast of North Carolina. 2014.
53. Johnson ZI, Shyam R, Ritchie AE, Mioni C, Lance VP, Murray JW, et al. The effect of iron-and light-limitation on phytoplankton communities of deep chlorophyll maxima of the western Pacific Ocean. Journal of Marine Research. 2010;68(2):283-308.
54. Marie D, Partensky F, Jacquet S, Vaulot D. Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. Applied and Environmental Microbiology. 1997;63(1):186-93.
55. Atlas EL, Hager SW, Gordon LI, Park PK. A Practical Manual for Use of the Technicon (Trade Name) Autoanalyzer (Trade Name) in Seawater Nutrient Analyses; Revised. OREGON STATE UNIV CORVALLIS DEPT OF OCEANOGRAPHY; 1971.
56. Gordon LI, Jennings Jr JC, Ross AA, Krest JM. A suggested protocol for continuous flow automated analysis of seawater nutrients (phosphate, nitrate, nitrite and silicic acid) in the WOCE Hydrographic Program and the Joint Global Ocean Fluxes Study. WOCE hydrographic program office, methods manual WHPO. 1993(68/91):1-52.
57. Needham D, Fichot E, Parada A, Yeh Y-C, Fuhrman J. Fuhrman Lab 515F-926R 16S and 18S rRNA Gene Sequencing Protocol V.2. protocolsio. 2019.

58. Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environmental Microbiology. 2016;18(5):1403-14.
59. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4:e2584.
60. Eren AM, Morrison HG, Lescault PJ, Reveillaud J, Vineis JH, Sogin ML. Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. The ISME Journal. 2015;9(4):968-79.
61. Angly FE, Dennis PG, Skarshewski A, Vanwonterghem I, Hugenholtz P, Tyson GW. CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. Microbiome. 2014;2(1):11.
62. Morel A, Huot Y, Gentili B, Werdell PJ, Hooker SB, Franz BA. Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. Remote Sensing of Environment. 2007;111(1):69-88.
63. Applications NCFS, Research. GHRSST NOAA/STAR GOES-16 ABI L2P Western Atlantic SSTskin v2.70 dataset in GDS2. NASA Physical Oceanography Distributed Active Archive Center; 2019.
64. Seim HE, Savidge D, Andres M, Bane J, Edwards C, Gawarkiewicz G, et al. Overview of the processes driving exchange at Cape Hatteras program. Oceanography. 2022;35(2).

Supplemental materials:

Table S1. Multiple regression analysis of select environmental variables on Bray-Curtis dissimilarity for 16S rRNA gene libraries using 'envfit' in R (v 4.0.0) with 9999 permutations.

Table S2. Linear discriminant analysis effect size (LEfSE) results for the 250 most abundant ASVs across all samples.

Table S3. GJAM beta-coefficients of ASVs found to have a significant association with at least one SOM cluster.

Table S4. Linear discriminant analysis effect size (LEfSE) results for the 250 most abundant ASVs in Gulf Stream and eddy comparison.

Figure S1. Environmental variables for physical parameter defined water parcels (A) Temperature and (B) salinity for each of the temperature and salinity-defined water parcels

(Continental Shelf, Continental Slope, Gulf Stream, Eddy) as determined using the flow-through seawater system at ~1m depth. Center lines denotes the median value while the box encompasses the 25th to 75th percentiles of the dataset. Whiskers denote the 5th and 95th percentiles and values that fall outside these ranges are shown as individual points. Brackets indicate p-values associated with Wilcoxon Rank Sum test pairwise comparisons.

Figure S2. Non-metric multidimensional scaling (NMDS) ordination comparisons of bacterial communities grouped by physical and biological parameters. NMDS computed based on Bray-Curtis dissimilarity for 16S rRNA gene libraries and colored by (A) oceanographic water parcels (temperature and salinity) and (B) Self-Organizing Map (SOM) clusters based on the microbiome. Ellipses show the multivariate t-distribution 95% confidence interval for the mean microbial community of each water parcel or SOM cluster.

Figure S3. Maps of sampling sites for five transects off the coast of Cape Hatteras, North Carolina, USA from September 5-7, 2021. Background color represents sea surface temperatures at time of sampling obtained from the Geostationary Operational Environmental Satellite 16 (GOES-16). Each sampling point is shown along the transects and categorized by oceanographic water parcel (shape), defined by temperature and salinity, and microbial-community based k-means clusters (color), generated using the self-organizing map R package 'kohonen'. Black dotted lines indicate the approximate outline of the frontal eddy during the time of sampling based on a qualitative analysis of the SST imagery.

Figure S4. Environmental conditions for Self-Organizing Map (SOM) clusters. Numbered clusters correspond to the most common environments: Continental shelf (Cluster 1), Continental slope (Cluster 2), Gulf Stream (Cluster 3), Eddy (Cluster 4). For the box and whisker plots, the center lines denote the median value while the box encompasses the 25th to 75th percentiles of the dataset. Whiskers denote the 5th and 95th percentiles and values that fall outside these ranges are shown as individual points. Brackets indicate p-values of Wilcoxon Rank Sum pairwise comparisons.

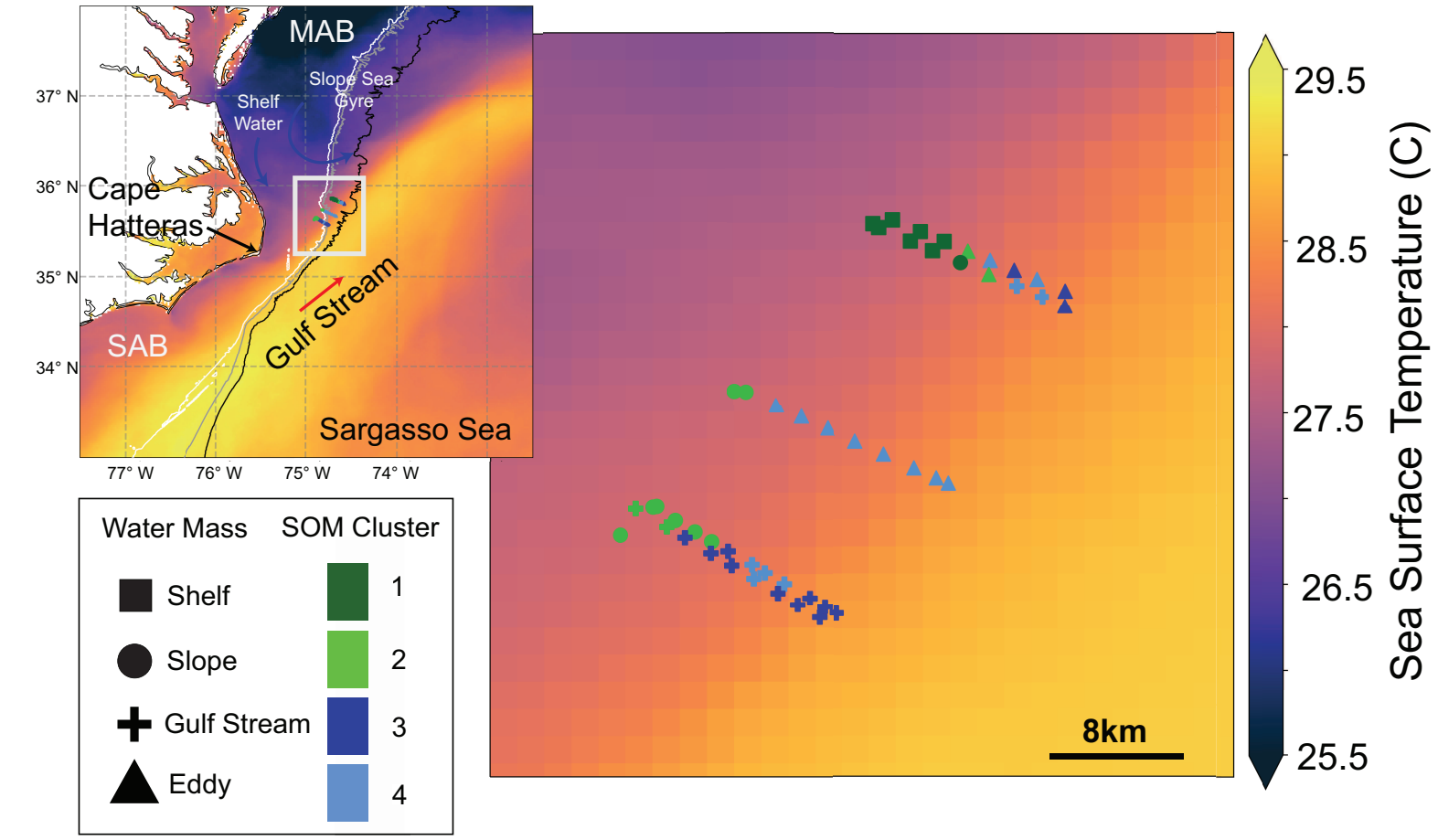
Figure S5. Non-metric multidimensional scaling (NMDS) ordination computed based on Sørensen similarity for 16S rRNA gene libraries. Each sampling point is categorized by oceanographic water parcel (shapes), defined by temperature and salinity, and microbial-community based k-means clusters (colors), generated using the self-organizing map R package 'kohonen'. Ellipses show the multivariate t-distribution 95% confidence interval for the mean of each Self-Organizing Map (SOM cluster).

Figure S6. **Self-Organizing Map within-clusters sum of squares plot for cluster number selection**. Calculated using a 4x4 grid with hexagonal mapping units in the 'kohonen' (v.3.0.11) in R (v4.0.0).

Figure 1

Click here to access/download;Figure;SW_figure1.eps

A



B

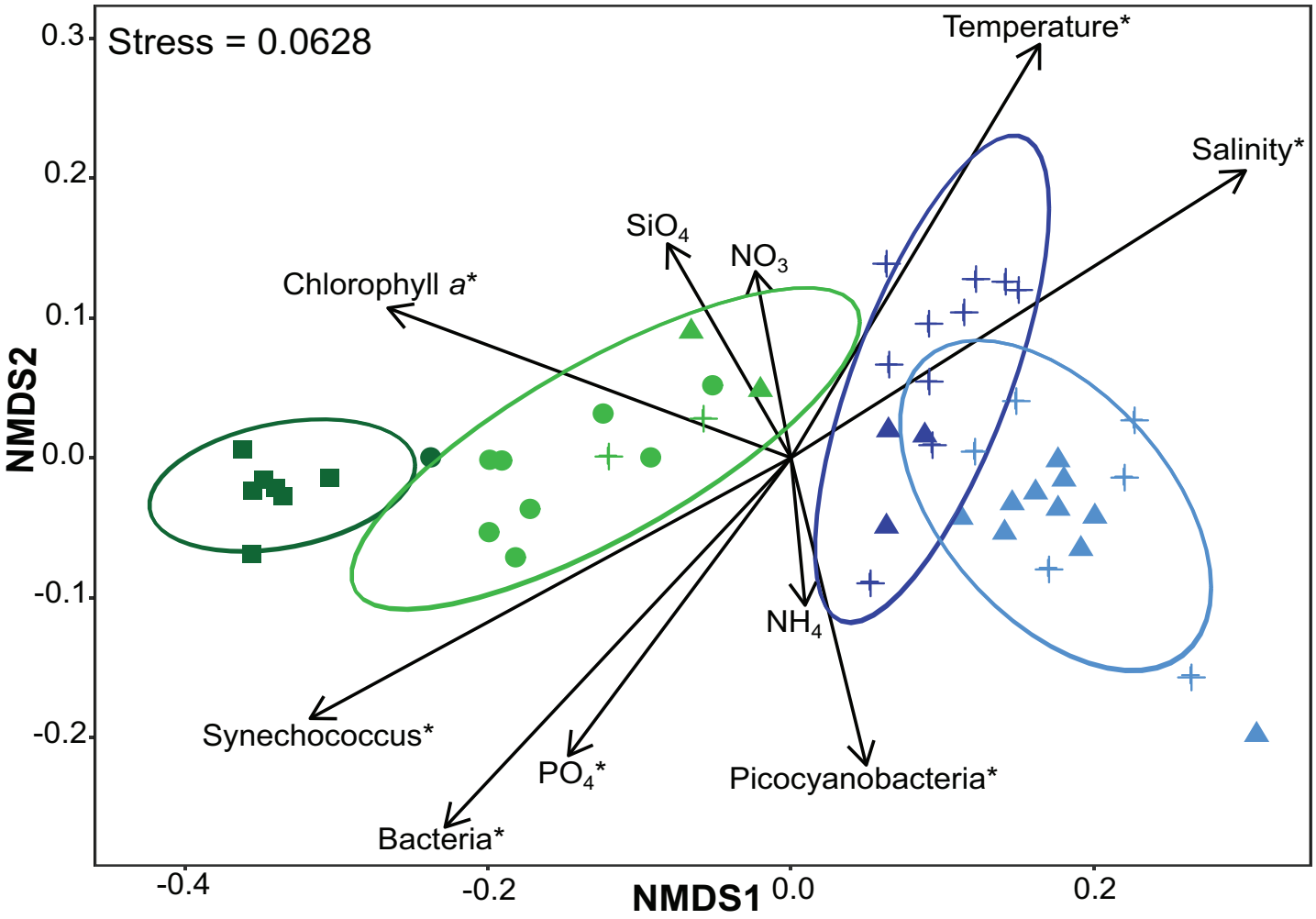
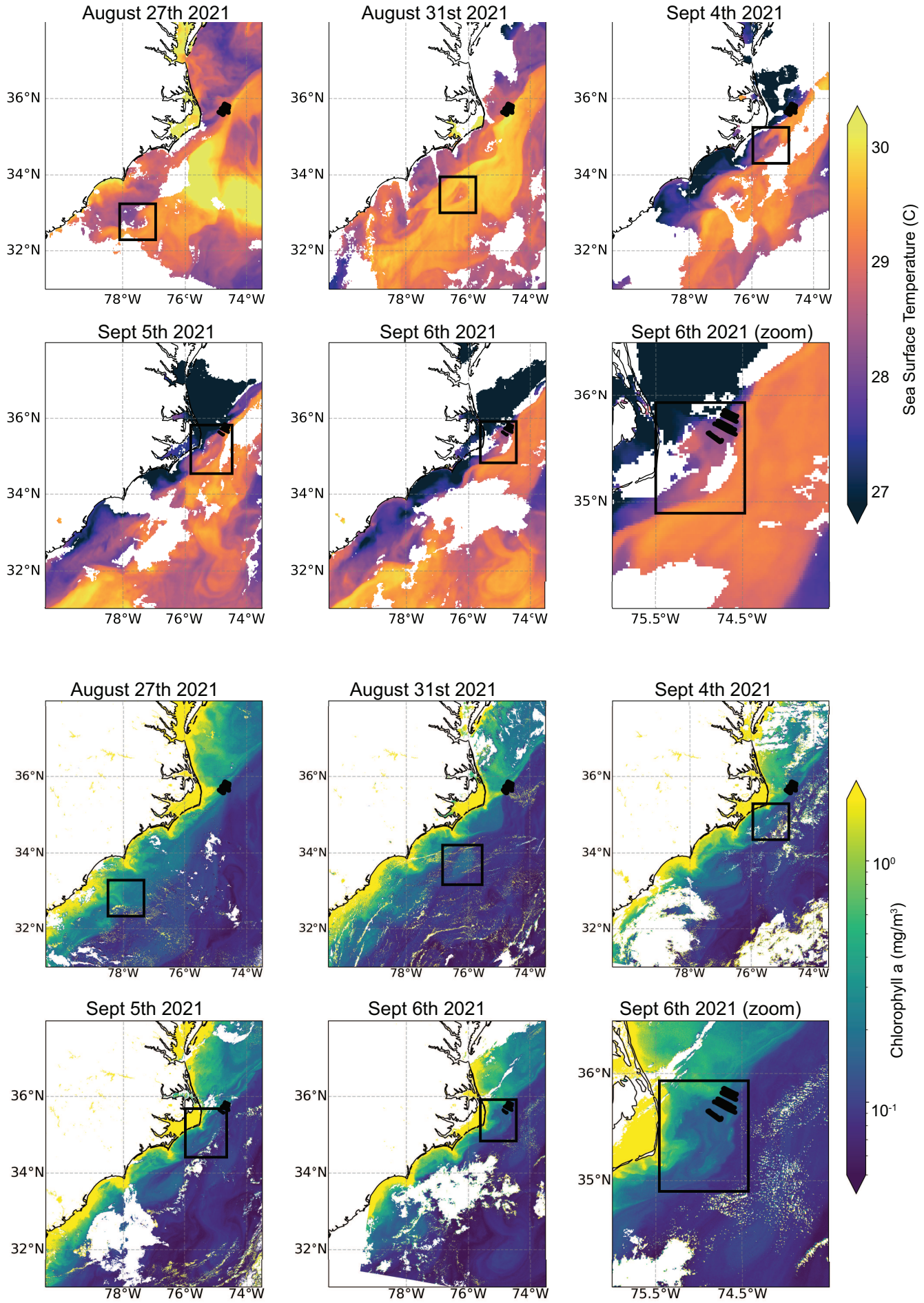


Figure 2

Click here to access/download;Figure;SW_Figure2.eps



Click here to access/download;Figure;SW_Figure3.eps

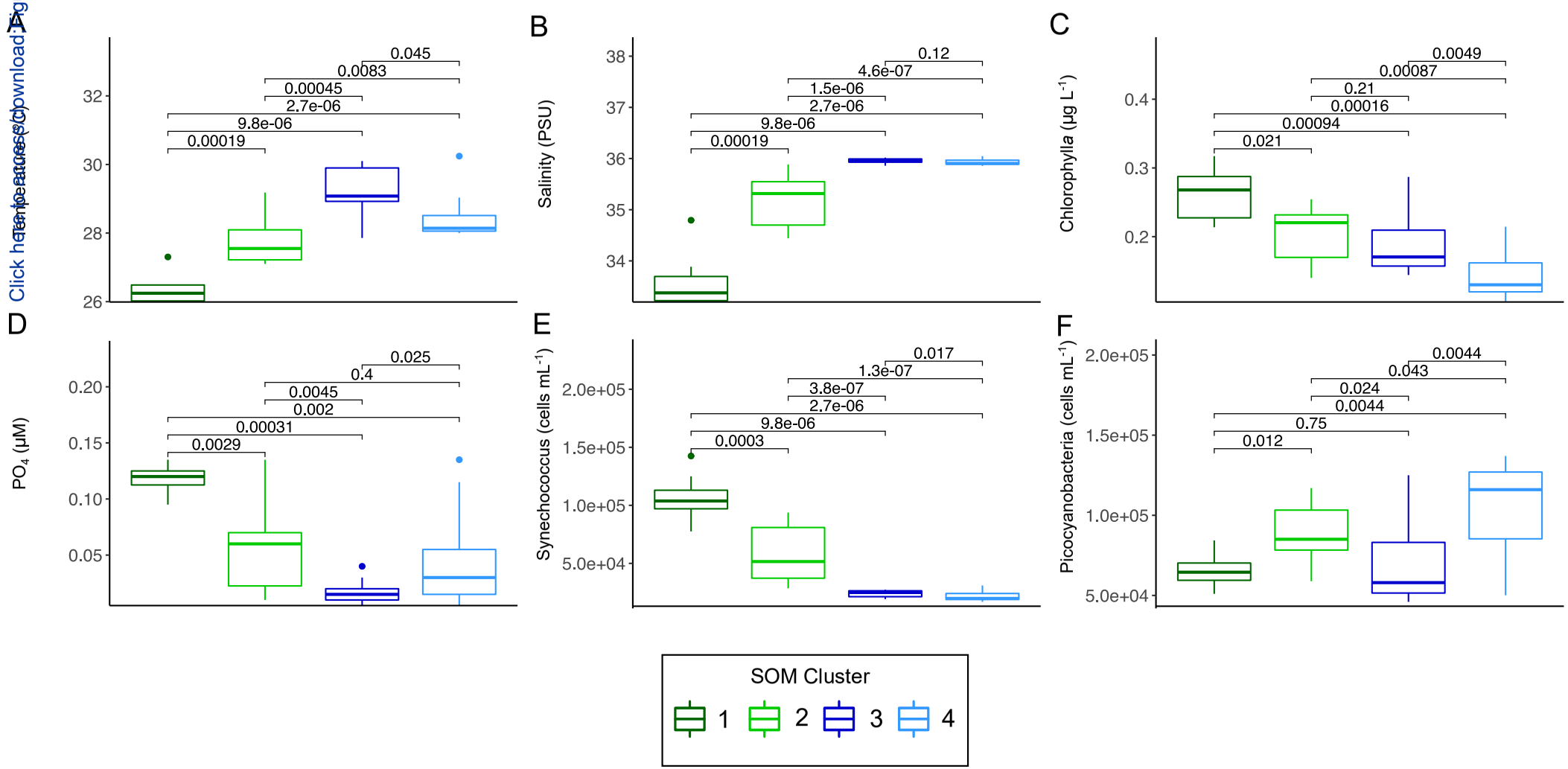


Figure 3

Figure 4

Click here to access/download;Figure;SW_Figure4.eps

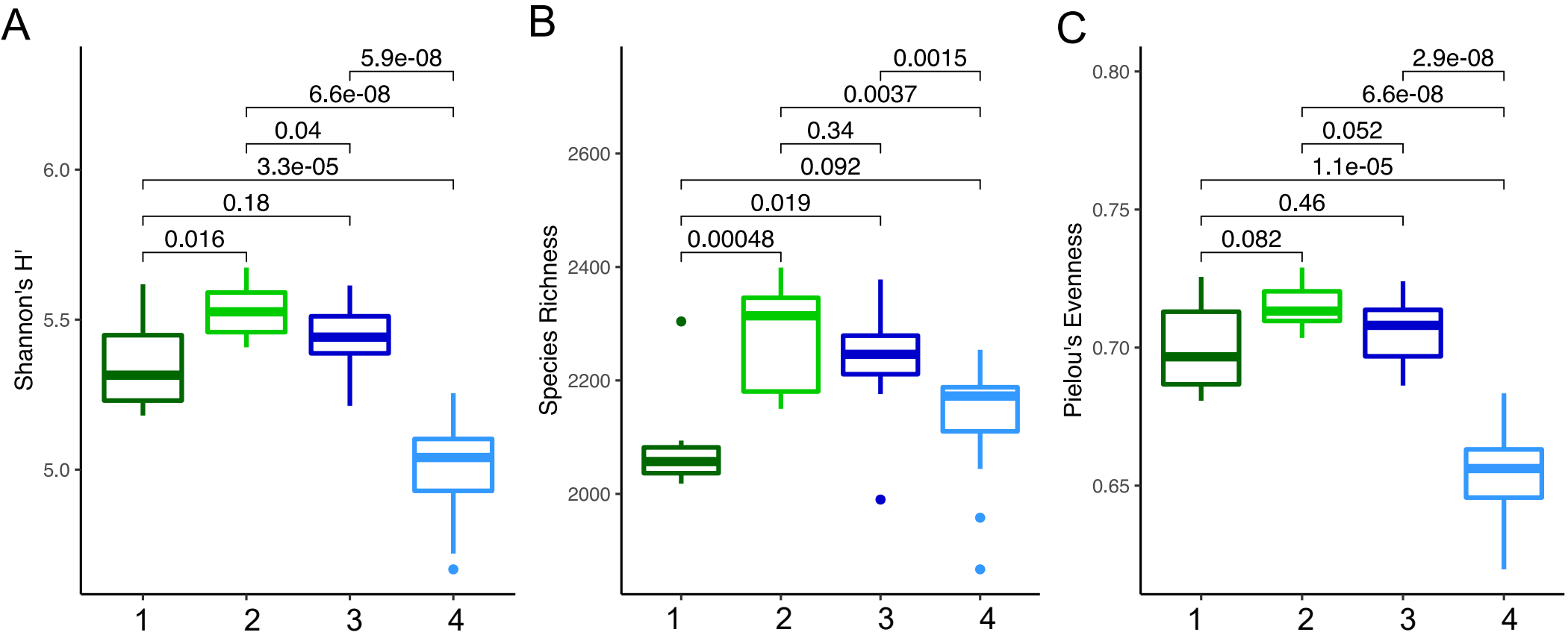


Figure 5

Click here to access/download;Figure;SW_Figure5.eps

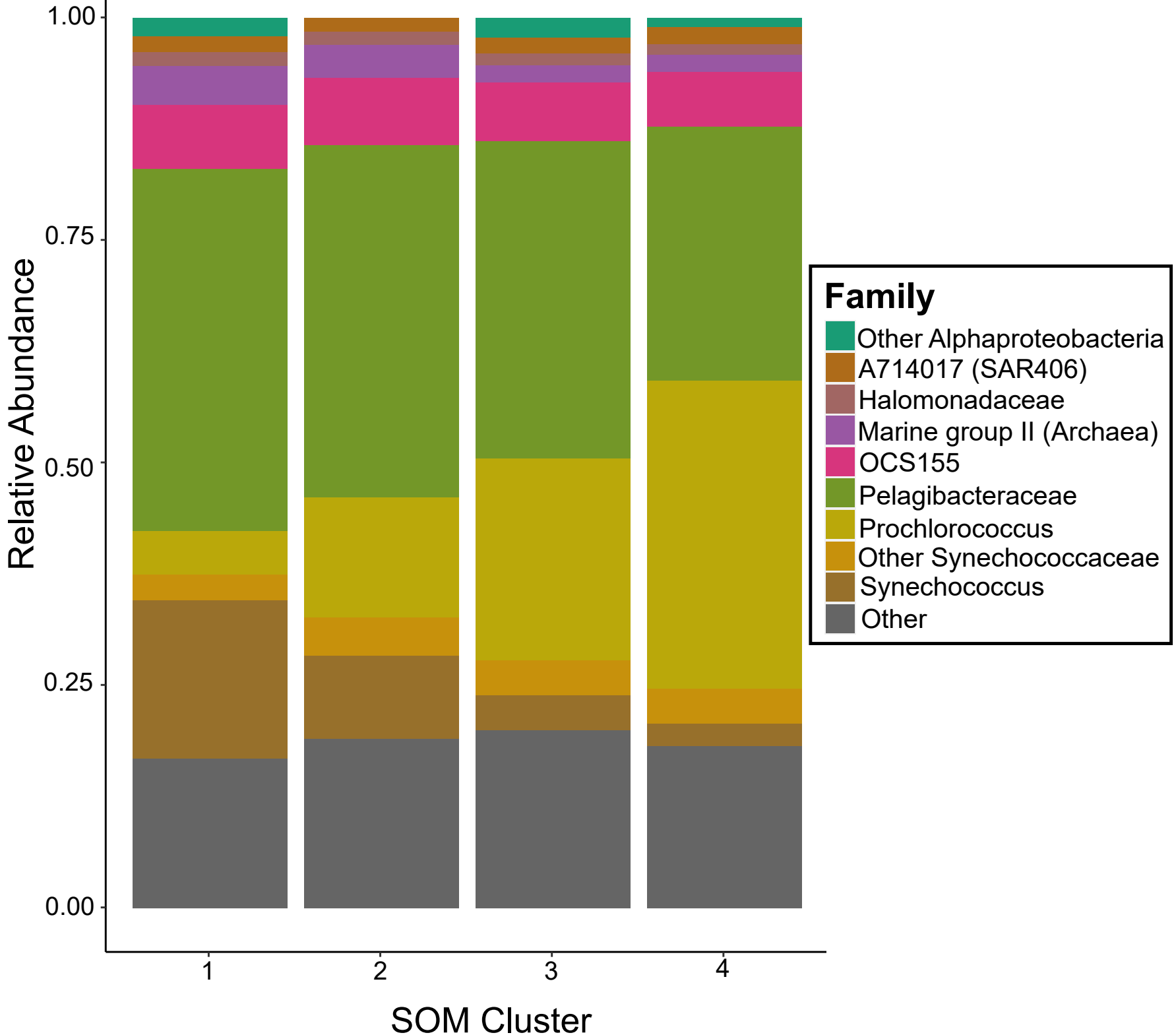


Figure 6

Click here to access/download;Figure;SW_Figure6.eps

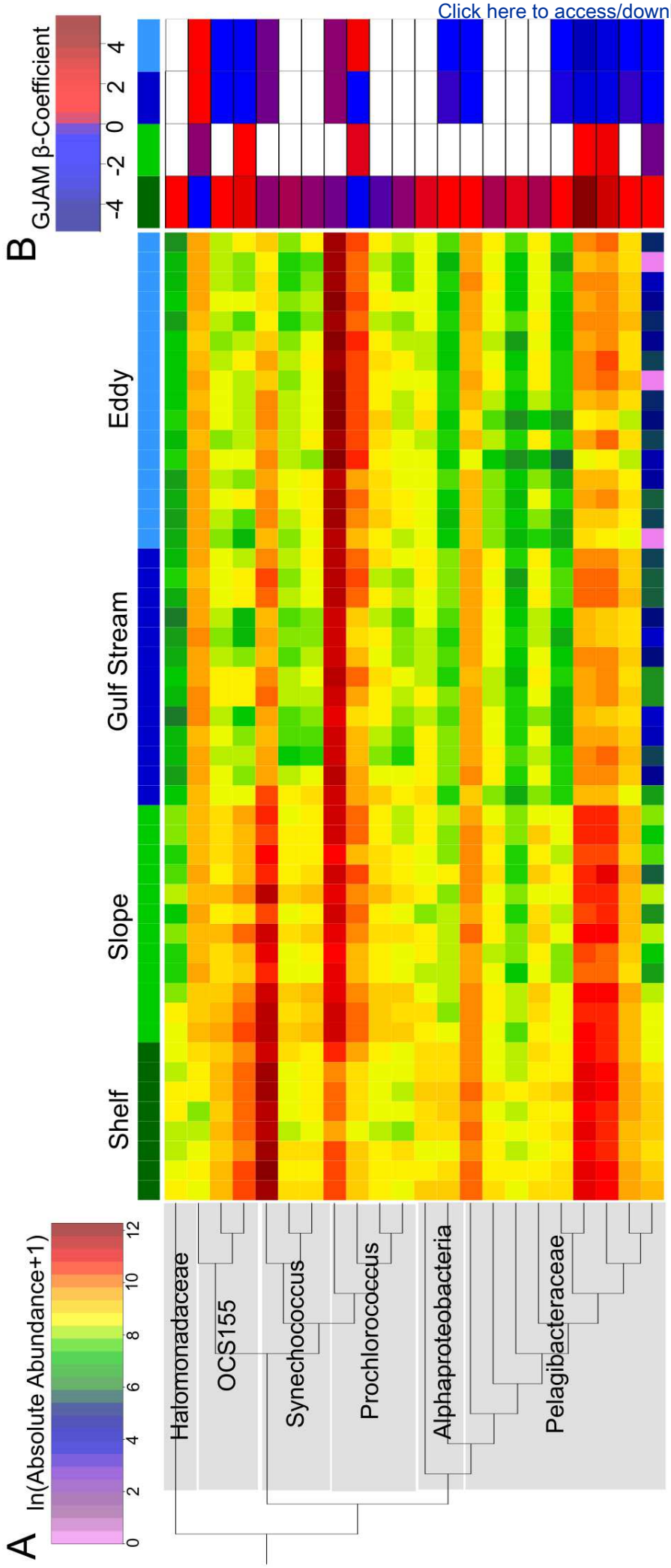
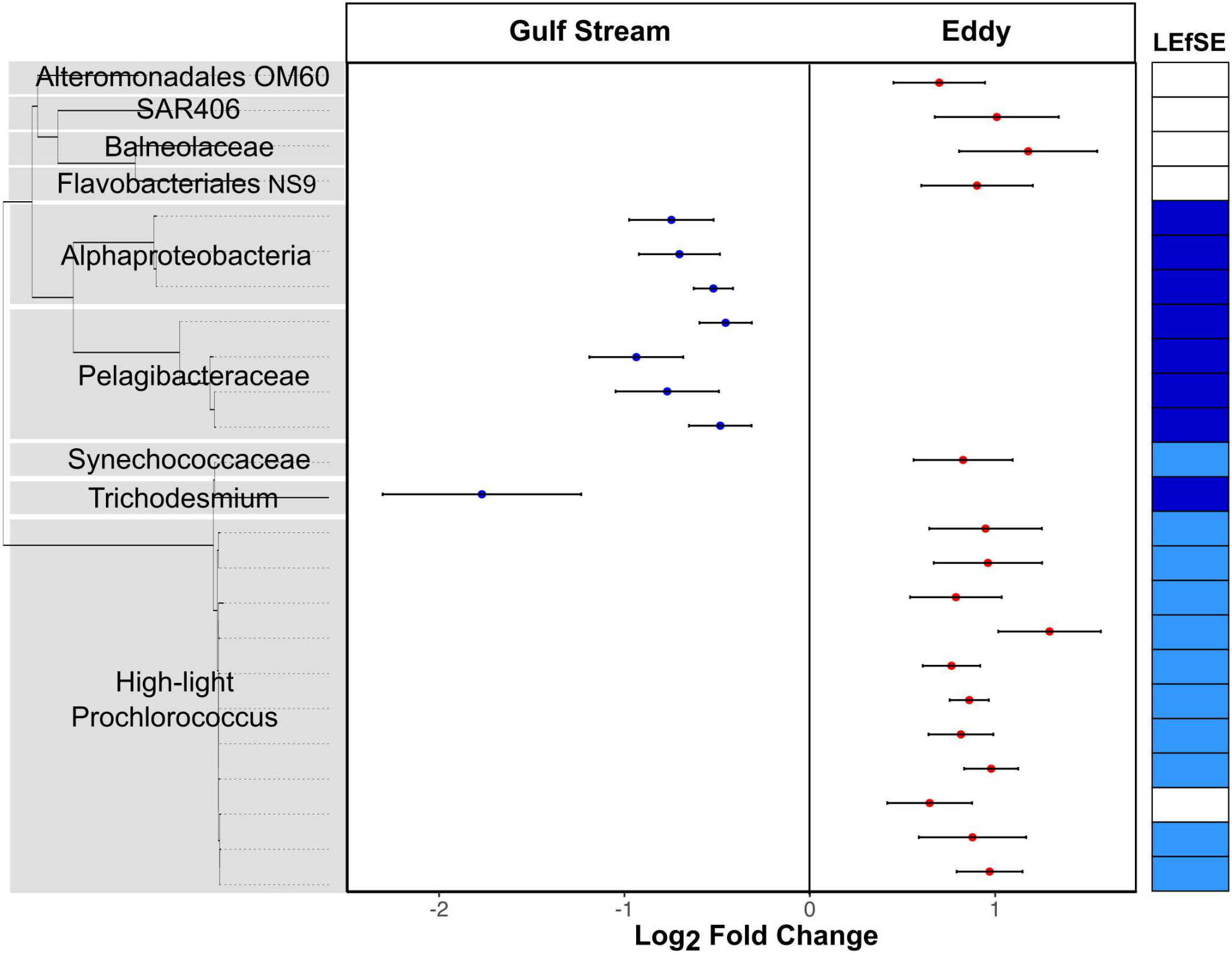


Figure 7

Click here to access/download;Figure;SW_Figure7.eps



Click here to access/download
Supporting Information
Gronniger_Supplementary_Materials.docx