

Temporal dynamics of Bacteria, Archaea and protists in equatorial coastal waters

**Caroline Chénard¹, Winona Wijaya¹, Daniel Vaulot^{2, 1}, Adriana Lopes dos Santos^{1, 3},
Patrick Martin¹, Avneet Kaur¹, and Federico M Lauro^{1, 4 *}**

¹Asian School of the Environment, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798

²Sorbonne Université, CNRS, UMR7144, Ecology of Marine Plankton team, Station Biologique de Roscoff, 29680 Roscoff, France

³GEMA Center for Genomics, Ecology & Environment, Universidad Mayor, Camino La Pirámide, 5750, Huechuraba, Santiago, Chile

⁴Singapore Centre for Environmental Life Sciences Engineering (SCELSE), Nanyang Technological University, 60 Nanyang Dr, Singapore 637551

*Corresponding author: Federico M. Lauro, 50 Nanyang Avenue, Singapore 639798, Tel: (65) 6790-9845, Fax: (65) 6515-6751, Email: flauro@ntu.edu.sg

ABSTRACT

Singapore, an equatorial island in South East Asia, is influenced by a bi-annual reversal of wind directions which defines two monsoon seasons. We characterized the dynamics of the microbial communities of Singapore coastal waters by collecting monthly samples between February 2017 and July 2018 at four sites located across two straits with different trophic status, and sequencing the V6-V8 region of the small sub-unit ribosomal RNA gene (rRNA gene) of Bacteria, Archaea, and Eukaryota. Johor Strait, which is subjected to wider environmental fluctuations from anthropogenic activities, presented a higher abundance of copiotrophic microbes, including Cellvibrionales and Rhodobacterales. The mesotrophic Singapore Strait, where the seasonal variability is caused by changes in the oceanographic conditions, harboured a higher proportion of typically marine microbe groups such as Synechococcales, Nitrosupumilales, SAR11, SAR86, Marine Group II Archaea and Radiolaria. In addition, we observed seasonal variability of the microbial communities in the Singapore Strait, which was possibly influenced by the alternating monsoon regime, while no seasonal pattern was detected in the Johor Strait.

Submitted to

Date: May 31, 2019

1 Introduction

2 Marine planktonic communities harbour representatives from all three domains of life (Bacteria,
3 Archaea and Eukaryota). Together, these organisms perform a range of global biological and
4 geochemical processes. In contrast to the open ocean, coastal environments are influenced by
5 local disturbances such as freshwater land run-off, land to sea transfer of nutrients and organic
6 matter through river discharge and mixing in shallow areas induced by tidal currents, in addition
7 to seasonal cycles. These variable conditions may lead to complex dynamics where patterns of
8 reoccurring microbial communities are harder to access and predict.

9 In temperate waters, annual cycles of plankton communities have been studied for a long
10 time^{1,2}. The seasonal changes in microbial community composition and biomass are often
11 attributed to responses to changes in environmental conditions driven by seasonal climate
12 cycles^{3–6}. For example, Lambert *et al.*⁴ demonstrated a strong rhythmicity in Bacteria, Archaea
13 and protist communities despite sporadic meteorological events and irregular nutrient availability
14 in a coastal site situated in the North Western Mediterranean Sea. In contrast, few studies have
15 investigated the seasonal dynamics of coastal microbial communities in equatorial waters.

16 Equatorial waters are subjected to monsoons, periods when the prevailing winds over land
17 and adjacent ocean areas reverse directions on a seasonal basis in response to differences
18 of heating pattern between land and ocean. These differences alter the patterns of rising and
19 sinking air near the equator, resulting in the seasonal migration of the intertropical convergence
20 zone (ITCZ), an area of low atmospheric pressure where Northeast and Southeast Trade Winds
21 meet⁷. By influencing vertical mixing⁸, upwelling⁹ and advective transport¹⁰, monsoon systems
22 have previously been shown to influence phytoplankton dynamics in tropical waters^{8,11,12}. Miki
23 *et al.*⁸ demonstrated a higher concentration of chlorophyll during the NE monsoon than the SW
24 monsoon in the Sulu Sea off the southwest side of the Philippines. A similar trend in chlorophyll
25 concentration was also observed in the Strait of Malacca¹¹.

26 In Singapore, a highly urbanized island city state of South East Asia, located just one degree
27 North of the equator, a bi-annual reversal of wind directions and two monsoon seasons named

28 after the prevailing wind direction⁷ characterize the annual meteorological conditions in the island.
29 The Northeast Monsoon (NE Monsoon), generally occurring from December to early March,
30 brings primarily north-easterly winds and long periods of heavy rain across the region, while the
31 Southwest Monsoon (SW Monsoon), from June to September, exhibits south-westerly winds with
32 higher temperatures, scattered showers and thunderstorms mostly occurring in the afternoon.
33 Inter-monsoon periods are generally less windy and experience lower precipitation. Singapore is
34 also surrounded by two main bodies of water with different trophic status: the Singapore Strait in
35 the South and the Johor Strait in the North. The Singapore Strait exhibits oceanic conditions
36 and is influenced by strong tidal currents up to 2 m s^{-1} and alternate oceanic influxes from
37 the South China Sea and the Java Sea¹³. The Johor Strait, less than 1 km wide, is subjected
38 to greater environmental fluctuations resulting from a combination of anthropogenic activities,
39 sporadic riverine inputs and reduced tidal mixing¹⁴. Recurrent spring/summer blooms have
40 been reported during the inter-monsoon months April/May in the years of 1935, 1948, 1949 and
41 1968¹⁵. More recent studies either applied morphological identification methods to a restricted
42 number of taxa¹⁴ or used techniques such as flow cytometry and pigments analysis that have
43 limited taxonomic resolution^{16,17} to investigate the dynamics of the plankton communities in
44 Singapore coastal waters.

45 Here, we present an 18-month study that uses SSU rRNA metabarcoding to characterize
46 the temporal variation of the marine microbial community in Singapore waters. We observed
47 taxonomic compositions reflecting differences in trophic status between the two Straits and
48 identified a seasonal signature in the microbial community composition of the Singapore Strait,
49 probably linked to the monsoonal current reversal. In contrast, no seasonal patterns, but large
50 month-to-month changes, in beta-diversity were observed in the Johor Strait, suggesting that
51 localized pulse disturbances might be responsible for shaping community composition.

52 Materials and Methods

53 Sampling sites and water collection

54 Surface water (~ 1 m depth) was collected with a submersible pump monthly from February
55 2017 to July 2018 at four different stations around Singapore (Figure S1, Supplementary Data
56 S1). Sampling was performed within two days of neap tide, except for samples # 21 collected
57 during the full moon (28 June 2018). Two stations were in the Singapore Strait, St. John (1.2383°
58 N, 103.8536° E) and East Coast (1.2972° N, 103.9207° E), and two in the Johor Strait, Pasir Ris
59 (1.3886° N, 103.9515° E) and Sembawang (1.4716° N, 103.8051° E). Four monsoon periods
60 were identified using the prevailing wind direction from the annual climatological report for 2017
61 and 2018 from the Meteorological Service of Singapore¹⁸: Northeast (NE), inter-monsoon 1
62 (IM-1), Southwest (SW), and inter-monsoon 2 (IM-2). A month was assigned to the NE or SW
63 Monsoon if the wind recorded at Changi airport blew for more than 50% of the month from either
64 the first quadrant (NE) or third quadrant (SW) direction, respectively, and assigned to the IM
65 periods in all other cases.

66 Temperature and salinity were measured at ~ 1m using a Eureka Water Probes Manta+
67 Trimeter Multiprobe. Chlorophyll was measured using an AquaFlash™ Handheld Active Fluoro-
68 meter (Turner Designs). For dissolved inorganic nutrients, water was syringe-filtered (0.22
69 μm Acrodisc, PALL) into acid-washed, polypropylene centrifuge tubes, immediately flash-frozen
70 in liquid nitrogen, and stored at -20° C until analysis. For microbial community samples, 1L
71 of seawater was first filtered through a 150 μm net mesh filter to remove large detritus and
72 zooplankton, and then the biomass was collected on 0.22 μm Sterivex filters (Millipore). The
73 Sterivex filters were then immediately stored at -80° C until further analysis. Samples for flow
74 cytometry (980 μL) were fixed for 15 min at 4° C in the dark with 25% electron microscopy (EM)
75 grade glutaraldehyde (0.5% final concentration ; Sigma-Aldrich), flash-frozen in liquid nitrogen
76 and stored at -80° C until analysis.

77 **Nutrient analysis**

78 Samples for dissolved inorganic macronutrients, $\text{NO}_3 + \text{NO}_2$ (referred to as NO_x hereafter), NO_2 ,
79 NH_4 , PO_4 , and $\text{Si}(\text{OH})_4$ were thawed at room temperature and immediately measured on a
80 SEAL AA3 segmented-flow autoanalyser according to SEAL methods for seawater analysis.
81 NH_4 was measured fluorometrically¹⁹. For samples collected on 3 April 2017, 02 May 2017
82 and 31 May 2017, the dissolved inorganic macronutrients were measured using the APHA²⁰
83 method and Si (4500-SiO₂ Flow injection analysis for Molybdate-Reactive Silicate) as a service
84 provided by DHI-Singapore Seawater. Total dissolved inorganic nitrogen (DIN) was calculated
85 as $\text{NO}_x + \text{NH}_4$.

86 **Enumeration of bacteria**

87 Bacteria were enumerated in duplicate water samples by flow cytometry using a CytoFLEX
88 (Beckman Coulter, Singapore) equipped with blue (488 nm) and violet (405 nm) lasers. Prior to
89 analysis, samples were thawed and diluted (between 5-10x) with 0.2 μm filtered sterile 10 mM
90 Tris, 1 mM ethylene-diamine-tetra-acetic acid buffer (pH 8.0) and stained with SYBR Gold (Life
91 Technologies, Singapore). Samples were incubated in the dark for 15 minutes prior to analysis.
92 The analysis was performed for 1 minute at medium flow rate ($\sim 30 \mu\text{L min}^{-1}$) with an event
93 rate of ~ 200 -1200 particles per second with a threshold set on the green fluorescence channel
94 (525/40 BP). Bacteria were discriminated based on their signature on a scatter plot of green
95 fluorescence against violet side-scatter using CytExpert software provided with the CytoFLEX.

96 **DNA extraction and 16S/18S PCR amplification and sequencing**

97 DNA was extracted from the Sterivex filters using a modified protocol for MoBio PowerSoil kit
98 (MoBio Laboratories, Carlsbad, CA), as previously published by Jacobs *et al.* 2009²¹. Briefly, the
99 Sterivex filter casing was broken open and the filter was cut with a sterile razor blade into small
100 lengthwise strips and placed into the PowerBead tube. 60 μL of the manufacturer solution C1
101 were added to the tube which was incubated for 10 min at 70° C with constant shaking (500 rpm)
102 twice and vortexed at maximum speed between each incubation. Following the last incubation,

103 700 μ L of phenol-chloroform-isoamyl alcohol was added into the PowerBead tube and vortexed
104 at maximum speed for 10 min. The PowerBead tube was centrifuged at 10,000 g for 30 seconds
105 and 800 μ L of supernatant transferred to a new collection tube to perform the DNA extraction
106 as per the manufacturer's instructions. DNA was eluted into 100 μ L of nuclease-free water and
107 quantified using a Qubit® 2.0 fluorometer with the dsDNA Broad Range assay kit (Invitrogen,
108 Singapore). Amplicon libraries were then generated using a modified version of the Illumina
109 16S Metagenomic Sequencing Library Preparation Protocol²². Universal primer pairs (926wF:
110 5'-AAACTYAAAKGAATTGRCGG-3' and 1392R: 5'-ACGGGCGGTGTGTRC-3') targeting the
111 V6-V8 hyper-variable region of the 16S/18S Ribosomal RNA gene with an Illumina-specific
112 overhang were used²³. For each sample, triplicate PCR reactions obtained with 22 cycles were
113 pooled and purified using Agencourt AMPure XP beads (Beckman Coulter, Singapore). Products
114 from this first step were then sent to the sequencing facility at the Singapore Centre for Life
115 Science and Engineering (SCELSE) where a second round of PCR was performed to add dual
116 barcodes to each amplicon library. Afterwards, PCR products were purified with Agencourt
117 AMPure XP beads, pooled at equal volume and sequenced with an Illumina MiSeq sequencing
118 run (2x300 bp).

119 Sequence analysis

120 The raw sequencing data was initially processed by removing primers with cutadapt²⁴. Paired-
121 end joining, denoising and taxonomic assignment of Amplicon Sequence Variants (ASV) were
122 performed using Quantitative Insights Into Microbial Ecology (QIIME) release 2018.6²⁵. Briefly,
123 after importing the sequence data into the QIIME2 environment, the denoising and pair-end
124 joining were performed using DADA2²⁶. After discarding all the ASVs with less than 10 se-
125 quence representatives, phylogenetic relationships were inferred by aligning representative
126 sequences with MAFFT²⁷, filtering the alignment and constructing phylogenetic trees using
127 fasttree²⁸ with a midpoint root. Taxonomic classification was performed using the classify-sklearn
128 method²⁹ using the SILVA version 132 database as a reference. All ASVs identified as eu-
129 karyotic (nuclear, plastid) were further assigned against the PR² database³⁰ version 4.12.0

130 (<https://github.com/pr2database/pr2database>) using the RDP naive Bayesian classifier⁹ as implemented in the R *dada2* package. ASVs with a bootstrap value lower than 90% at the superegroup level were discarded. For statistical analyses, Metazoa (mostly copepods) were removed from the final ASV table because they are likely to represent eggs or organismal fragments that were not captured by the pre-filtering step. Similarly, eukaryotic plastid sequences were removed to avoid counting photosynthetic organisms twice. The final ASV table contained a total of 2,571 ASVs.

137 **Statistical analyses**

138 Pairwise comparisons between the different locations (Johor Strait, Singapore Strait) and monsoon periods were performed using ANOSIM with the software R³¹. Pairwise community dissimilarity was calculated using Bray-Curtis distance as implemented in the *vegan* package. The Shannon index (H') and richness calculated with the *vegan* package, nMDS plots were generated with the metaMDS function also in the *vegan* package. The Envfit analysis generated using the *vegan* package used the following environmental parameters: salinity, temperature, chlorophyll, bacterial abundance, PO₄, DIN, Si(OH)₄, NH₄ and amount of rainfall seven days (rain-7) prior to the sampling day. The differential abundance of relevant ASVs was extrapolated using Gneiss³², with Ward's correlation clustering as a function of monsoon period for the Singapore Strait. The ASVs contributing significantly to the difference between the two monsoon regimes were computed using the DESeq2 package³³ using a threshold for the corrected p-value of 0.01. The magnitude of changes occurring in the community during the time series was inferred with the first-distances method³⁴ applied to the unweighted Unifrac distance matrix³⁵.

151 **Results**

152 **Physico-chemical parameters in Singapore coastal waters**

153 Despite their proximity and physical connection, the Singapore and Johor Straits exhibit different environmental conditions (Figure 1, Figure S2, Supplementary Data S1). Temperature showed little variation, but was slightly higher in the Johor (29.0–32.8°C) than the Singapore (27.1–30.6°C)

156 Strait (Figure S2A). The lowest temperatures were consistently found during the NE monsoon.
157 Much larger variability was observed in salinity: the two sites in the Singapore Strait were very
158 similar to each other, with salinity ranging from 29.8 to 32.6 and lowest values during the SW
159 monsoon (Figure 1). In contrast, salinity was lower and more variable in the Johor Strait, ranging
160 from 18.1 to 30.3, with lowest values mostly during the NE monsoon. Sembawang, the sampling
161 site closest to the causeway between Singapore and Malaysia, typically had lower salinity than
162 Pasir Ris (Figure 1).

163 Dissolved nutrient concentrations also differed strongly between the Singapore and Johor
164 Straits: NO_x , NO_2 , NH_4 and PO_4 were typically 2–10-fold higher in the Johor compared to the
165 Singapore Strait, while Si(OH)_4 was on average more similar, but with higher variability in the
166 Johor Strait (Figure 1, Supplementary Data S1, Figure S2). The Singapore Strait showed more
167 oligo- to mesotrophic characteristics, with NO_x ranging from $< 1\text{--}5 \mu\text{mol L}^{-1}$, NO_2 and NH_4
168 almost always $< 0.5 \mu\text{mol L}^{-1}$ and often close to or below detection limits, and PO_4 frequently
169 $< 0.2 \mu\text{mol L}^{-1}$. Si(OH)_4 in the Singapore Strait showed a seasonal pattern inverse to that of
170 salinity. In contrast, the Johor Strait mostly had $\text{NO}_x > 5 \mu\text{mol L}^{-1}$ and often $> 10 \mu\text{mol L}^{-1}$, with
171 NO_2 regularly contributing more than 50% of NO_x . NH_4 mostly ranged at Pasir Ris from 1 to
172 $8 \mu\text{mol L}^{-1}$, and at Sembawang mostly around $10\text{--}40 \mu\text{mol L}^{-1}$, reaching as high as $75 \mu\text{mol L}^{-1}$
173. PO_4 in the Johor Strait was generally in the range of $0.5\text{--}3.0 \mu\text{mol L}^{-1}$, but occasionally
174 dropped to $< 0.1 \mu\text{mol L}^{-1}$. Higher values were usually found during the SW monsoon. In the
175 Singapore Strait, NO_x and PO_4 were generally highest just before or during the SW monsoon,
176 and lowest during the NE monsoon. The Johor Strait displayed less coherent seasonal patterns,
177 with NO_x and PO_4 showing opposite seasonal trends in Sembawang, and no clear pattern for
178 Si(OH)_4 . There were no differences in total precipitation for the 7 days prior to sampling between
179 the Johor and Singapore Straits throughout the year, with values ranging from 0 to 128 mm
180 (Supplementary Data S1).

181 Bacterial abundance was different between the Singapore and Johor Straits (Supplemen-
182 tary Data S1, Figure S2), with abundances ranging from 0.4×10^6 to $1.6 \times 10^6 \text{ cells}\cdot\text{mL}^{-1}$ and
183 1.1×10^6 to $9.4 \times 10^6 \text{ cells}\cdot\text{mL}^{-1}$, respectively. In general, higher bacterial counts were recorded at

184 Sembawang.

185 Chlorophyll concentration varied substantially between sites, broadly reflecting the concen-
186 tration of nutrients: low values were found in the Singapore Strait (mostly $<2 \mu\text{g L}^{-1}$, but up to
187 $13 \mu\text{g L}^{-1}$), while the Johor Strait rarely had less than $10 \mu\text{g L}^{-1}$, and up to $94 \mu\text{g L}^{-1}$ (Figure 1,
188 Supplementary Data S1). In the Singapore Strait, chlorophyll was lowest during the SW monsoon
189 and highest during inter-monsoon 1 or the NE monsoon, while in the Johor Strait, concentrations
190 were highest during the NE monsoon or inter-monsoon 2, and lower during the SW monsoon.

191 Community composition of Singapore coastal waters

192 A total of 70 samples, 30 from the Singapore Strait and 40 from the Johor Strait from February
193 2017 to July 2018 were analyzed by rRNA metabarcoding using primers that amplify both
194 prokaryotic 16S and eukaryotic 18S rRNA²³. After quality control filtering, end-pair joining and
195 chimera filtering, a total of 5,036,783 sequences were retained with an average of 71,954
196 sequences per sample which were assigned to 2,556 ASVs (Supplementary Data S2).

197 Based on the Shannon index (H'), Singapore Strait communities generally had a higher
198 diversity than those from the Johor Strait for Archaea and Eukaryota, but not for the Bacteria
199 (Figure 2). The richness, inferred from the number of ASVs, was also generally higher in the
200 Singapore Strait for Archaea and Eukaryota (Figure 2).

201 The taxonomic composition of Bacteria, Archaea and Eukaryota communities varied between
202 straits (Figures 3, S3 and S4). At the domain level, Bacteria dominated both straits but Archaea
203 were more prevalent than Eukaryota in the Singapore Strait than the Johor Strait (Figure 3).

204 Among prokaryotes, Cyanobacteria, specific proteobacterial groups and the archaeal phyla
205 Euryarchaeota and Thaumarchaeota were over-represented in the Singapore Strait. Within the
206 Archaea, the phylum Thaumarchaeota was mainly represented by the order Nitrosopumilales,
207 which was present in both the Singapore and Johor Straits, but had a higher relative proportion in
208 the Singapore Strait (Figure S4). In the Singapore Strait, both *Candidatus Nitrosopelagicus* and
209 *Candidatus Nitrosopumilus* were abundant while *Candidatus Nitrosopumilus* was abundant in
210 the Johor Strait (Figure S5. Euryarchaeota were mostly found in the Singapore Strait, with many

211 ASVs assigned to marine group II Euryarchaeota (MGII). Marine group III Euryarchaeota (MG-III)
212 were also present in the Singapore Strait, but in lower proportion, and absent in most samples
213 from the Johor Strait. Other prevalent members of the microbial communities of the Singapore
214 Strait included the alphaproteobacterial group SAR11, the gammaproteobacterial group SAR86,
215 the deltaproteobacterial group SAR324 and the candidate actinobacterial genus *Actinomar-*
216 *rina*. In the Johor Strait, Proteobacteria and Bacteroidetes were the most abundant groups
217 (Figure S4). The most abundant family within the Bacteroidetes in Johor included the Flavobac-
218 tericeae and the NS5 marine group. Within the Proteobacteria there was a high abundance
219 of reads assigned to *Roseobacter* strain HIMB11, to members of the gammaproteobacterial
220 OM60/NOR5 clade, Cellvibrionales (in particular members of the genus *Luminiphilus*) and the
221 family Burkholderiaceae, Figure S4, Figure S3, Figure S6).

222 The eukaryotic planktonic community in both straits was dominated by Dinoflagellata (Alve-
223 olate), Ochrophyta (Stramenopiles, mostly diatoms) and Chlorophyta (Figure S4-C). Among
224 Dinoflagellata, Dinophyceae and Syndiniales were the main groups in the Singapore Strait, while
225 the Johor Strait was dominated by Dinophyceae (Figure S4-C). Most of Dinophyceae reads from
226 the Johor Strait were affiliated with the order Gymnodiniales (e.g. *Gyrodinium* and *Woloszynskia*)
227 and Gonyaulacales, while the Singapore Strait reads were affiliated with uncultured dinoflagellate
228 sequences (Figure S5). For example, the sequence of the most abundant Dinophyceae ASV
229 (0115) in the Singapore Strait is 100% similar to that of a marine eukaryote clone o10.5-16
230 (KX532243) from the South China Sea (Supplementary Data S2). The Dinophyceae community
231 in the Singapore Strait was more diverse than in the Johor Strait, where two ASVs (0011, *Gyro-*
232 *dinium* and 0043, *Gonyaulax*) practically dominated the whole community (Figure S5). These
233 ASVs were also within the top 50 most abundant ASV when considering the whole dataset
234 (Bacteria, Eukaryota and Archaea).

235 Reads assigned to the marine parasites Syndiniales were nearly absent in the Johor Strait.
236 In fact, the most abundant Syndiniales ASV found in the Singapore strait were absent in
237 Johor (Figure S5 Supplementary Data S2). Syndiniales are divided into 5 main groups³⁶.
238 Syndiniales sequences from the Singapore Strait were mainly affiliated to the highly diverse

group II (Figure S7), which contains a single genus formally described, *Amoebophrya*, which infects a wide range of dinoflagellate hosts. Most Syndiniales reads could only be assigned to uncultured clades known from environmental sequences (Supplementary Data S2). For example, the sequence of ASV-0285, the main Syndiniales ASV from the Singapore Strait, was 100% similar to the uncultured eukaryote clone ST5900.074 (KF130025) obtained from the South China Sea.

Among the other Alveolata groups, very few reads were assigned to Apicomplexa (e.g Gregarines and Perskinsea, Supplementary Data S2). Only Ciliophora were found in both straits, but more abundant in the Singapore Strait (Figure S4-C). Ciliophora reads were mainly assigned to the class Spirotrichea, with the genera *Parastrombidinopsis* and *Pelagostrobilidium* present in the Johor Strait vs. *Strombidiida* and the *incertae sedis* genus *Mesodinium* in the Singapore Strait (Figure S5).

Bacillariophyta (Ochrophyta) was the second and third most abundant group in the Johor and the Singapore Straits, respectively (Figure S4). Mediophyceae (polar centrics) were well represented in the two straits while Coscinodiscophyceae (radial centrics) were mainly found in the Singapore Strait. *Cyclotella*, *Cerataulina* and *Thalassiosira* (all Mediophyceae) were the main genera found in the Johor Strait, while *Leptocylindrus* (Coscinodiscophyceae), *Guinardia* (Coscinodiscophyceae) and *Skeletonema* (Mediophyceae) dominated the Singapore Strait (Figure S5). The highly diverse genus *Chaetoceros* (Mediophyceae) was the only dominant genus common to both straits. Only two ASVs with a low number of sequences were assigned to Bacillariophyceae, which contains pennate diatom genera such as the toxic blooming *Pseudo-nitzschia* (Supplementary Data S2). Other classes of Ochrophyta (except for ASV-2888 belonging to Raphidophyceae) were not detected based on nuclear 18S rRNA. However Pelagophycae, Dictyochophyceae and Chrysophyceae, all Orchophyta classes, were detected based on plastid 16S rRNA (Supplementary Data S2). Stramenopile groups such as the diverse group of heterotrophic flagellates MAST (Marine stramenopiles) and fungus-like members of Oomycota and Labyrinthulea were also found in both straits, although in low abundance (Figure S4-C and Supplementary Data S2).

267 Mamiellophyceae and Trebouxiophyceae were the two main classes of Chlorophyta (green
268 algae) found in the Singapore and Johor Straits, respectively. Trebouxiophyceae ASVs were as-
269 signed to the highly diversified marine and brackish water coccoid genus *Picochlorum* (Figure S5),
270 known for its broad halotolerance^{37,38}. Members of the two widespread Mamiellophyceae gen-
271 era, *Micromonas* and *Ostreococcus*, were found in both straits (Figure S7 and Supplementary
272 Data S2). *Micromonas* was the main photosynthetic genus in the Singapore Strait (Figure S5)
273 with two clades, one corresponding to the ubiquitous species *M. commoda*, and the other, to
274 the undescribed clade B5,³⁹ found at the two stations in the Singapore Strait all year around
275 (Figure S7). Two clades of *Ostreococcus* were found mainly at Pasir Ris station in the Johor
276 Strait during the SW monsoon (Figure S7): Clade B, which seems to be the dominant clade in
277 tropical waters³⁹, and another clade that could be the new Clade E, which was found in Singa-
278 pore during the Ocean Sampling Day survey³⁹, and for which no culture is available yet. The
279 ubiquitous genus *Bathycoccus* was less abundant in Singapore waters, appearing sporadically
280 (Supplementary Data S2).

281 Within the supergroup Hacrobia, reads belonging to Haptophyta and Cryptophyta divisions
282 were the most abundant. Among haptophytes, all ASVs belonged to the class Prymnesiophyceae
283 (Figure S5 and Supplementary Data S2) and were closely related to the widespread and non-
284 calcifying genera *Chrysochromulina* and *Phaeocystis*, mostly in Singapore Strait. Only ASV
285 1427, with low abundance, was assigned to the calcifying coccolithophorids order Isochrysidales
286 (*Gephyrocapsa*) (Supplementary Data S2). Cryptophytes were mainly detected in the Singapore
287 Strait and assigned to the genus *Geminigera* (Figure S5).

288 Reads belonging to the supergroup Rhizaria were found in both straits. Radiolaria were
289 dominant in the Singapore Strait but nearly absent in the Johor Strait. The most abundant
290 Radiolaria ASV was assigned to the order Taxopodia clade RAD-B (Figure S5). Rhizaria
291 reads from the Johor Strait belonged to the division Cercozoa and were assigned to the Class
292 Filosa-Imbricatea (Figure S5)⁴⁰.

293 **Factors structuring microbial communities off Singapore**

294 Non-metric multidimensional scaling analysis (nMDS) of microbial community composition
295 revealed that samples clustered based on strait (Figure 4). While there was little dissimilarity
296 between the St. John and East Coast sites of the Singapore Strait cluster, there was a clear
297 separation between the Johor Strait sampling sites (Sembawang and Pasir Ris stations). An
298 analysis of similarity (ANOSIM) based on different grouping factors (strait, site, monsoon)
299 confirmed significant differences based on strait and site (Table 1). Envfit analysis revealed
300 that most environmental factors (salinity, temperature, chlorophyll a, DIN, NH₄ and bacterial
301 abundance were significant drivers ($p < 0.05$) for the difference between straits (Figure 4).

302 An nMDS analysis taking only into account the Singapore Strait displayed little overlap
303 between the SW monsoon and the others monsoon periods (Figure 4). ANOSIM confirmed that
304 there was a significant difference based on monsoon (Table 1) and this was confirmed by gneiss
305 balance analysis (Figure S9-top). PO₄, DIN and Si(OH)₄ were the most significant environmental
306 variables influencing separation of SW-Monsoon cluster from other monsoon periods ($p < 0.1$,
307 Figure 4). In contrast, the nMDS restricted to Johor Strait samples did not show clustering based
308 on monsoon (Figure S8) and ANOSIM demonstrated a significant difference only based on site
309 (Table 1).

310 **Seasonal variability in Singapore coastal water communities**

311 Seasonal variability associated with the monsoon was observed in the Singapore Strait. Linear
312 regression models implemented in Gneiss analysis indicated a significant difference between
313 SW and NE monsoon regime (Figure S9top), which was driven by an increase in abundance of
314 members of the order Nitrosopumilaceae (Figure S6).

315 A DESeq2 analysis revealed that the abundance of 71 specific taxa was influenced by the
316 NE or SW monsoon. Among those ASVs related to Alteromonadales (0292, 0512, 0747), Pirellu-
317 lales(0514, 0490), SAR11(0217, 0407, 0448, 0130, 0181, 0280, 0074, 0054) SAR116(0371)
318 and *Synechococcus* clade CRD1 (0277) were associated with the SW monsoon, while the ASVs
319 related to Desulforculaceae (0772, 0921), Opitutales MB11C04 marine group (0819, 0274) and

320 *Cyanobium* (0424, 0470) were associated with the NE monsoon. For the archaea, different
321 ASVs assigned to the Marine Group II and Marine Group III showed a change in their abundance
322 by monsoon based on the DESeq2 analysis. ASVs related to Marine Group II (0398, 0185, 0358,
323 0101) and Marine Group III (0297) were more abundant during the SW monsoon while ASVs
324 related to Marine Group II (0403,441) and Marine Group III (0423) were more abundant during
325 the NE monsoon. An abundant archaeal ASV (0005) assigned to *Candidatus Nitrosopumilus*
326 also showed an increase during the SW monsoon. Among eukaryotes, thirteen ASVs had their
327 abundance influenced by the monsoon according to DESeq2 analysis (Figure S9). ASVs as-
328 signed to Syndiniales group I (clades 1 and 2) and group II (clades 7 and 10-11), Mediophyceae
329 and Ciliophora genus *Mesodinium* had their abundance increased or detected during the SW
330 monsoon (Figure S9). During the NE monsoon an increase in abundance of ASVs related to
331 the Syndiniales group I (clades 1 and 4) and II (clade 16), Dynophyceae order Gymnodiniales,
332 Coscinodiscophyceae genus *Guinardia* and Ciliophora Class Spirotrichea was observed. Within
333 these 13 ASVs, only 3 (0532, 0380 and 0324) were among the top 30 most abundant ASVs in
334 the Singapore Strait (Figure S5).

335 Volatility values for the first distance analysis of the unweighted Unifrac matrix (Figure S10),
336 which can be interpreted as the month-to-month change in the composition of microbial com-
337 munities, were generally higher for samples from the Johor Strait (Sembawang and Pasir Ris)
338 and lower for the Singapore Strait sites (St. John and East Coast). The comparison of Figure 1
339 and Figure S10 revealed that, in the Johor Strait, peaks in volatility corresponded to large
340 month-to-month variations in salinity.

341 Discussion

342 Johor Strait microbial community is influenced by short term events

343 The lower values and the absence of a clearly defined seasonal pattern in salinity compared to
344 the Singapore Strait (Figure 1) suggest that Johor Strait is influenced by river and land runoff.
345 The lower salinity at Sembawang compared to Pasir Ris also suggests that the central portion
346 of the Johor Strait is affected by land-runoff from storm drains, reservoirs and smaller rivers

such as the Sungai Tebrau, located just north of Sembawang station. The oscillation in salinity was also reflected in the eukaryotic community composition, which harbored brackish water (e.g *Cyclotella*) and marine euryhaline groups (e.g *Picochlorum*). Groups known to have an exclusive marine lifestyle such as Syndiniales and *Micromonas* were absent at Sembawang samples and only sporadically found in the samples from Pasir Ris (Figure S7). On the other hand, the cosmopolitan genus *Gyrodinium*, known to inhabit marine, brackish and freshwaters, was found at both Johor stations (Figure S7).

The Johor Strait is clearly eutrophic compared to the Singapore Strait, with consistently higher nutrient and chlorophyll concentrations. Higher concentrations of NO₂ and NH₃, and the higher bacterial counts, indicate that heterotrophic recycling is likely more pronounced in the Johor Strait, consistent with eutrophication. It is unclear to what extent the nutrient concentrations in the Johor Strait are controlled by direct inputs via run-off, or whether sedimentary recycling processes are significant. The fact that the concentration of DIN is inversely correlated to salinity in the Johor Strait (Figure S11) but not to precipitation (data not shown) suggests that reservoir run-off processes may play an important role. Moreover, the Johor Strait exhibited a higher abundance of copiotrophic families of microbes such as the Flavobactericeae, the Burkholderiaceae, taxa affiliated with the *Roseobacter* and the OM60/NOR5 clades. All these groups have been previously implicated in nutrient remineralization and rapid responses to pulses of organic carbon such as those resulting from phytoplankton blooms and coastal runoff^{41,42}. The most abundant prokaryotic ASV in the Johor Strait *Roseobacter* strain HIMB11 (Figure S5 was also previously shown to have the genomic potential for degradation of algal-derived compound such as DMSP⁴³. The Burkholderiaceae have also been shown to be one of the dominant groups in the network of urban waterways of Singapore⁴⁴ which might represent a reservoir and source for their presence in Johor Strait.

A lower alpha diversity and a larger beta diversity was observed in the Johor compared to the Singapore Strait. The overall lower alpha diversity in the Johor Strait was mostly driven by Eukaryota and Archaea. The frequent and short pulse disturbances (short-term events with release of nutrients) experienced by the Johor Strait planktonic community seemed to affect

375 its stability and ultimately the diversity of the system⁴⁵. The wider Johor sample dispersion
376 in the nMDS plots suggested little mixing between the planktonic communities from the Jo-
377 hor and the Singapore straits as well as within the Johor Strait, as supported by the cluster
378 separation between Pasir Ris and Sembawang samples (Figure 4). The seasonal change in
379 regional seawater circulation due to monsoon wind reversal did not seem to affect the microbial
380 community composition of the Johor Strait directly. The only seasonal trend we observed in the
381 community, linked to the SW monsoon, was the presence of the two pico-phytoplanktonic genera
382 *Ostreococcus* and *Micromonas* (clade B5) in Pasir Ris samples during July of 2018 and 2019
383 (Figure S7).

384 Singapore Strait community and seasonality

385 The Singapore Strait samples contained a higher proportion of autotrophic and oligotrophic
386 microbes such as the Synechococcales, Nitrosupumilales, marine group II Euryarchaeota (MGII)
387 and SAR11. The most prevalent cyanobacteria found in the Singapore Strait belonged to the
388 order Synechococcales and were most closely related to *Synechococcus* representative of clade
389 II, which has been identified as the dominant type in warm and oligotrophic waters⁴⁶. The phylum
390 Euryoarchaeota found in high abundance in the Singapore Strait is related to Marine Group
391 II, which is widely distributed in the oceanic euphotic zone^{47,48} and is the dominant planktonic
392 Archaeon at the surface water of the South China Sea⁴⁹. Marine group III Euryarchaeota (MG-III)
393 was also present in the Singapore Strait, but in a lower proportion. This group is known to
394 be prevalent in deep-sea waters, but few studies have reported their presence in the photic
395 zone^{50–52}. Other prevalent members of the microbial communities of the Singapore Strait
396 included the Alphaproteobacteria clade SAR11, which has previously been found to dominate
397 heterotrophic bacterial communities in coastal and open ocean environments^{53–55}.

398 The microbial community shift in the Singapore Strait is synchronous with the seasonal
399 reversal of ocean currents between the Java Sea and the South China Sea¹³ suggesting
400 that the advection of microbial communities from different water basins might be causing this
401 shift. During the SW monsoon, currents flow northward from the Java Sea¹³, bringing less

402 saline water into the Singapore Strait due to the high precipitation over Sumatra, Borneo,
403 and Java. The high freshwater input from these islands into the Java Sea is likely also the
404 source of nutrients transported to Singapore at the start of the SW monsoon period (Figure 1).
405 The net northerly water transport during the SW monsoon transports a community with a
406 higher proportion of Archaea to the Singapore Strait, especially the ammonia-oxidizing archaea
407 *Candidatus Nitrosopelagicus* and *Candidatus Nitrosopumilus*, presumably reflecting higher
408 nutrient concentrations and greater dominance of heterotrophic recycling of organic matter.
409 During the NE monsoon, the current direction reverses and water from the South China Sea
410 flows into the Singapore Strait¹³, carrying more saline and more oligotrophic water with less
411 terrestrial input into the Singapore Strait. Consequently, all nutrients decrease in concentration
412 except for a small peak around the beginning of the NE monsoon (Figure 1) that is probably
413 driven by local run-off (since this is typically the period with the highest rainfall in Singapore).

414 While our study showed an increasing trend in nutrient concentration during the SW mon-
415 soon, the chlorophyll did not show any trend. Other parts of South East Asia have also been
416 shown to experience seasonal dynamics in the chlorophyll concentration as a result of monsoon
417 systems^{8,10,11}. The analysis of the ASVs that differ most significantly between the two mon-
418 soons suggested that ecotype replacement is likely occurring during the seasonal cycle. This
419 replacement involves ASVs belonging to SAR11 clades, SAR406 and the NS4 and NS5 groups,
420 which have been shown to display seasonal dynamics in temperate waters^{56–60}.

421 Although few eukaryotic ASVs had their abundance increased or were present during either
422 the SW or NE monsoon period, no clear seasonal pattern was observed among eukaryotes.
423 *Micromonas*, the main photosynthetic picoeukaryotes in our dataset, have been reported as
424 the dominant group in the sub-tropical waters of the South China Sea⁶¹ and off Taiwan⁶². In
425 Taiwan, high abundances of clade B5 were detected during the summer and autumn, with a peak
426 in July when the local hydrographic characteristics were high temperature and irradiance, and
427 oligotrophy⁶². In our study, *Micromonas* sp. clade B5 seem to have their abundance increased
428 during the NE and inter monsoon months (Figure S7), a period when the water of the Singapore
429 Strait tends to be more oligotrophic.

430 Members of the *Mesodinium* species complex are known to form periodic or recurrent non-
431 toxic blooms (red tides)^{63,64}. These ciliates can photosynthesize by acquiring and maintaining
432 organelles from cryptophyte prey^{65,66}. Cryptophytes are an important component of phyto-
433 plankton communities in coastal ecosystems, especially estuaries environments^{67,68}. High
434 abundances of cryptophytes have been associated with either preceding⁶⁹ or co-occurring
435 peaks⁷⁰ of *Mesodinium* in coastal ecosystems. Although *Mesodinium* and the cryptophyte
436 species complex *Geminigera* were among the abundant groups in Singapore strait, we did not
437 observe a *Geminigera* - *Mesodinium* dynamic in our dataset. Also, to our knowledge, red tides
438 have not been reported in the Singapore Strait.

439 Blooms of *Coscinodiscus* and *Chaetoceros*, during the NE and inter-monsoon periods, were
440 reported in the Singapore Strait 70 years ago¹⁵. Although these taxa were among the dominant
441 phytoplankton group that we found in the strait, neither a seasonal trend nor a bloom were
442 observed during the present study. During the 1968 SW monsoon, a bloom of pennate diatoms
443 (Bacillariophyceae) was also reported¹⁵, and this group is known to be part of the phytoplankton
444 community in the Singapore Strait¹⁴. Surprisingly, very few reads in our dataset were assigned to
445 Bacillariophyceae, probably due to the fact that the forward primer used displays one mismatch
446 to all Ochrophyta (data not shown), which is confirmed by the fact that this group was detected
447 through its plastid 16S rRNA sequences (Supplementary Data S2). As this is the first long-term
448 study of the protist communities in the waters of Singapore since the early 50's¹⁵, it is impossible
449 to known whether either the absence of blooms today or the appearance of blooms in the past
450 are anomalies of the ecosystem.

451 Conclusion

452 In this 18-month long study we successfully captured the environmental characteristics of
453 two Straits with different trophic status. In the more eutrophic Strait, the Johor Strait, large
454 environmental fluctuations were observed throughout the year, yet no recurring seasonal pattern
455 could detected in the microbial community composition. Conversely, the Singapore Strait showed
456 a seasonal variability which might be a result of the different monsoon regimes. Our study

457 suggests that even in the vicinity of the Equator, where irradiance and temperature show little
458 variation, seasonal trends are reflected in the microbial community composition in relation to
459 monsoon alternation. Based on these findings, we argue that longer sustained observations
460 over a broader spatial scale are needed, in equatorial waters, to identify the patterns of microbial
461 community assembly and the longterm trajectories in the seasonal succession of specific
462 taxonomic groups of Bacteria, Archaea and Eukaryota.

463 References

- 464 1. Colebrook, J. M. Continuous Plankton Records: Seasonal cycles of phytoplankton and
465 copepods in the North Atlantic ocean and the North Sea. *Mar. Biol.* **51**, 23–32, DOI:
466 10.1007/BF00389027 (1979).
- 467 2. Winder, M. & Cloern, J. E. The annual cycles of phytoplankton biomass. *Philos. Transactions
468 Royal Soc. B: Biol. Sci.* **365**, 3215–3226, DOI: 10.1098/rstb.2010.0125 (2010).
- 469 3. Piredda, R. *et al.* Diversity and temporal patterns of planktonic protist assemblages at
470 a Mediterranean Long Term Ecological Research site. *FEMS Microbiol. Ecol.* **93**, DOI:
471 10.1093/femsec/fiw200 (2017).
- 472 4. Lambert, S. *et al.* Rhythmicity of coastal marine picoeukaryotes, bacteria and archaea
473 despite irregular environmental perturbations. *ISME J.* **13**, 388–401, DOI: 10.1038/
474 s41396-018-0281-z (2019).
- 475 5. Giner, C. R. *et al.* Quantifying long-term recurrence in planktonic microbial eukaryotes. *Mol.
476 Ecol.* **28**, 923–935, DOI: 10.1111/mec.14929 (2019).
- 477 6. Steele, J. A. *et al.* Marine bacterial, archaeal and protistan association networks reveal
478 ecological linkages. *ISME J.* **5**, 1414–1425, DOI: 10.1038/ismej.2011.24 (2011).
- 479 7. Zhisheng, A. *et al.* Global Monsoon Dynamics and Climate Change. *Annu. Rev. Earth
480 Planet. Sci.* **43**, 29–77, DOI: 10.1146/annurev-earth-060313-054623 (2015).

- 481 **8.** Miki, M., Ramaiah, N., Takeda, S. & Furuaya, K. Phytoplankton dynamics associated with
482 the monsoon in the Sulu Sea as revealed by pigment signature. *J. Oceanogr.* **64**, 663–673
483 (2008).
- 484 **9.** Wang, J., Qi, Y. & Jones, I. S. An analysis of the characteristics of chlorophyll in the Sulu
485 Sea. *J. Mar. Syst.* **59**, 111–119, DOI: 10.1016/j.jmarsys.2005.09.004 (2006).
- 486 **10.** Sriwoon, R., Pholpunthin, P., Lirdwitayaprasit, T., Kishino, M. & Furuya, K. Population
487 dynamics of green *Noctiluca scintillans* (Dinophyceae) associated with the monsoon cycle
488 in the upper Gulf of Thailand. *J. Phycol.* **44**, 605–615, DOI: 10.1111/j.1529-8817.2008.00516.x
489 (2008).
- 490 **11.** Siswanto, E. & Tanaka, K. Phytoplankton Biomass Dynamics in the Strait of Malacca
491 within the Period of the SeaWiFS Full Mission: Seasonal Cycles, Interannual Variations and
492 Decadal-Scale Trends. *Remote. Sens.* **6**, 2718–2742, DOI: 10.3390/rs6042718 (2014).
- 493 **12.** Abdul-Hadi, A., Mansor, S., Pradhan, B. & Tan, C. K. Seasonal variability of chlorophyll-a
494 and oceanographic conditions in Sabah waters in relation to Asian monsoon—a remote
495 sensing study. *Environ. Monit. Assess.* **185**, 3977–3991, DOI: 10.1007/s10661-012-2843-2
496 (2013).
- 497 **13.** Xu, M. & Chua, V. P. A numerical study on circulation and volume transport in Singapore
498 coastal waters. *J. Hydro-Environment Res.* **12**, 90, DOI: 10.1016/j.jher.2015.11.005 (2016).
- 499 **14.** Tan, K. S., Acerbi, E. & Lauro, F. M. Marine habitats and biodiversity of Singapore's coastal
500 waters: A review. *Reg. Stud. Mar. Sci.* **8**, 340–352, DOI: 10.1016/j.rsma.2016.01.008 (2016).
- 501 **15.** Tham, A. K. Seasonal distribution of the plankton in Singapore Straits. *Special Publ. Dedic.*
502 to Dr N K Panikkar
- 503 **16.** Gin, K. Y.-H. Dynamics and size structure of phytoplankton in the coastal waters of Singapore.
504 *J. Plankton Res.* **22**, 1465–1484, DOI: 10.1093/plankt/22.8.1465 (2000).

- 505 17. Gin, K. Y. H. Phytoplankton community structure in Singapore's coastal waters using HPLC
506 pigment analysis and flow cytometry. *J. Plankton Res.* **25**, 1507–1519, DOI: 10.1093/plankt/
507 fbg112 (2003).
- 508 18. Meteorological Service Singapore. Annual Climate Reports.
- 509 19. Kérouel, R. & Aminot, A. Fluorometric determination of ammonia in sea and estuarine waters
510 by direct segmented flow analysis. *Mar. Chem.* **57**, 265–275, DOI: 10.1016/S0304-4203(97)
511 00040-6 (1997).
- 512 20. Clesceri, L., Greenberg, A. & Eaton, A. *Standard Methods for the Examination of Water and
513 Wastewater, 22nd Edition* (APHA American Public Health Association, 2012).
- 514 21. Jacobs, J., Rhodes, M., Sturgis, B. & Wood, B. Influence of environmental gradients on
515 the abundance and distribution of *Mycobacterium* spp. in a coastal lagoon estuary. *Appl.
516 Environ. Microbiol.* **75**, 7378–7384, DOI: 10.1128/AEM.01900-09 (2009).
- 517 22. Illumina, I. Illumina 16S Metagenomic Sequencing Workflow (2017).
- 518 23. Wilkins, D., van Sebille, E., Rintoul, S. R., Lauro, F. M. & Cavicchioli, R. Advection shapes
519 Southern Ocean microbial assemblages independent of distance and environment effects.
520 *Nat. Commun.* **4**, 2457, DOI: 10.1038/ncomms3457 (2013).
- 521 24. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
522 *EMBnet.journal* **17**, 10, DOI: 10.14806/ej.17.1.200 (2011). ISSN2226-6089.
- 523 25. Bolyen, E. *et al.* QIIME 2: Reproducible, interactive, scalable, and extensible microbiome
524 data science. *PeerJ Prepr.* DOI: 10.7287/peerj.preprints.27295v2 (2018).
- 525 26. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data.
526 *Nat. Methods* **13**, 581–583, DOI: 10.1038/nmeth.3869 (2016). 15334406.
- 527 27. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:
528 Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780, DOI: 10.1093/
529 molbev/mst010 (2013).

- 530 **28.** Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood
531 trees for large alignments. *PLoS ONE* **5**, e9490, DOI: 10.1371/journal.pone.0009490 (2010).
532 Price,MorganN.,2010,FastTree2.
- 533 **29.** Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. Tech. Rep. (2011).
- 534 **30.** Guillou, L. *et al.* The Protist Ribosomal Reference database (PR²): a catalog of unicellular
535 eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**,
536 D597–D604, DOI: 10.1093/nar/gks1160 (2013).
- 537 **31.** R Development Core Team. R: A Language and Environment for Statistical Computing, DOI:
538 10.1007/978-3-540-74686-7 (2013).
- 539 **32.** Morton, J. T. *et al.* Balance trees reveal microbial niche differentiation. *mSystems* **2**,
540 e00162–16 (2017).
- 541 **33.** Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
542 RNA-seq data with DESeq2. *Genome Biol.* **15**, 550, DOI: 10.1186/s13059-014-0550-8 (2014).
- 543 **34.** Bokulich, N. A. *et al.* q2-longitudinal: Longitudinal and Paired-Sample Analyses of Micro-
544 biome Data. *mSystems* **3**, e00219–18, DOI: 10.1128/msystems.00219-18 (2018).
- 545 **35.** Lozupone, C. & Knight, R. UniFrac: A new phylogenetic method for comparing microbial
546 communities. *Appl. Environ. Microbiol.* DOI: 10.1128/AEM.71.12.8228-8235.2005 (2005).
547 NIHMS150003.
- 548 **36.** Guillou, L. *et al.* Widespread occurrence and genetic diversity of marine parasitoids belong-
549 ing to Syndiniales (Alveolata). *Environ. Microbiol.* **10**, 3349–3365, DOI: 10.1111/j.1462-2920.
550 2008.01731.x (2008).
- 551 **37.** Wang, S., Lambert, W., Giang, S., Goericke, R. & Palenik, B. Microalgal assemblages in a
552 poikilohaline pond. *J. Phycol.* **50**, 303–309, DOI: 10.1111/jpy.12158 (2014).
- 553 **38.** Foflonker, F. *et al.* Genome of the halotolerant green alga *Picochlorum* sp. reveals strategies
554 for thriving under fluctuating environmental conditions. *Environ. Microbiol.* **17**, 412–426,
555 DOI: 10.1111/1462-2920.12541 (2015).

- 556 39. Tragin, M. & Vaultot, D. Novel diversity within marine Mamiellophyceae (Chlorophyta) unveiled
557 by metabarcoding. *Sci. Reports* **9**, 5190, DOI: 10.1038/s41598-019-41680-6 (2019). 449298.
- 558 40. Bass, D. & Cavalier-Smith, T. Phylum-specific environmental DNA analysis reveals re-
559 markably high global biodiversity of Cercozoa (Protozoa). *Int. J. Syst. Evol. Microbiol.* **54**,
560 2393–2404, DOI: 10.1099/ijsm.0.63229-0 (2004).
- 561 41. Yan, S. *et al.* Biogeography and phylogeny of the NOR5/OM60 clade of Gammaproteobacte-
562 ria. *Syst. Appl. Microbiol.* **32**, 124–139, DOI: 10.1016/j.syapm.2008.12.001 (2009).
- 563 42. Spring, S. *et al.* Taxonomy and evolution of bacteriochlorophyll *a*-containing members
564 of the OM60/NOR5 clade of marine gammaproteobacteria: Description of *Luminiphilus*
565 *syltensis* gen. nov., sp. nov., reclassification of *Haliea rubra* as *Pseudohaliea rubra* gen.
566 nov., comb. nov., and emendation of *Chromatocurvus halotolerans*. *BMC Microbiol.* **13**, DOI:
567 10.1186/1471-2180-13-118 (2013).
- 568 43. Durham, B. P. *et al.* Draft genome sequence of marine alphaproteobacterial strain
569 HIMB11, the first cultivated representative of a unique lineage within the Roseobacter
570 clade possessing an unusually small genome. *Standards Genomic Sci.* **9**, 632–645, DOI:
571 10.4056/sigs.4998989 (2014).
- 572 44. Saxena, G. *et al.* Metagenomics Reveals the Influence of Land Use and Rain on the
573 Benthic Microbial Communities in a Tropical Urban Waterway. *mSystems* **3**, 1–14, DOI:
574 10.1128/msystems.00136-17 (2018).
- 575 45. Shade, A. *et al.* Fundamentals of Microbial Community Resistance and Resilience. *Front.*
576 *Microbiol.* **3**, DOI: 10.3389/fmicb.2012.00417 (2012). NIHMS150003.
- 577 46. Farrant, G. K. *et al.* Delineating ecologically significant taxonomic units from global patterns
578 of marine picocyanobacteria. *Proc. Natl. Acad. Sci.* **113**, E3365–E3374, DOI: 10.1073/pnas.
579 1524865113 (2016).
- 580 47. Massana, R., DeLong, E. F. & Pedros-Alio, C. A Few Cosmopolitan Phylotypes Dominate
581 Planktonic Archaeal Assemblages in Widely Different Oceanic Provinces. *Appl. Environ.*

- 582 *Microbiol.* **66**, 1777–1787, DOI: 10.1128/AEM.66.5.1777-1787.2000 (2000).
- 583 **48.** Zhang, C. L., Xie, W., Martin-Cuadrado, A. B. & Rodriguez-Valera, F. Marine Group II
584 Archaea, potentially important players in the global ocean carbon cycle. *Front. Microbiol.* **6**,
585 1–9, DOI: 10.3389/fmicb.2015.01108 (2015).
- 586 **49.** Liu, H. *et al.* Marine Group II dominates planktonic archaea in water column of the North-
587 eastern South China Sea. *Front. Microbiol.* **8**, 1–11, DOI: 10.3389/fmicb.2017.01098 (2017).
- 588 **50.** Galand, P. E., Casamayor, E. O., Kirchman, D. L., Potvin, M. & Lovejoy, C. Unique archaeal
589 assemblages in the arctic ocean unveiled by massively parallel tag sequencing. *ISME J.* **3**,
590 860–869, DOI: 10.1038/ismej.2009.23 (2009).
- 591 **51.** Galand, P. E., Gutiérrez-Provecho, C., Massana, R., Gasol, J. M. & Casamayor, E. O. Inter-
592 annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea (Blanes
593 Bay Microbial Observatory). *Limnol. Oceanogr.* **55**, 2117–2125, DOI: 10.4319/lo.2010.55.5.
594 2117 (2010).
- 595 **52.** Haro-Moreno, J. M., Rodriguez-Valera, F., López-García, P., Moreira, D. & Martin-Cuadrado,
596 A. B. New insights into marine group III Euryarchaeota, from dark to light. *ISME J.* **11**,
597 1102–1117, DOI: 10.1038/ismej.2016.188 (2017).
- 598 **53.** Eiler, A., Hayakawa, D. H., Church, M. J., Karl, D. M. & Rappé, M. S. Dynamics of the
599 SAR11 bacterioplankton lineage in relation to environmental conditions in the oligotrophic
600 North Pacific subtropical gyre. *Environ. Microbiol.* **11**, 2291–2300, DOI: 10.1111/j.1462-2920.
601 2009.01954.x (2009).
- 602 **54.** Carlson, C. A. *et al.* Seasonal dynamics of SAR11 populations in the euphotic and
603 mesopelagic zones of the northwestern Sargasso Sea. *ISME J.* **3**, 283–295, DOI:
604 10.1038/ismej.2008.117 (2009).
- 605 **55.** Morris, R. M., Frazar, C. D. & Carlson, C. A. Basin-scale patterns in the abundance of SAR11
606 subclades, marine Actinobacteria (OM1), members of the Roseobacter clade and OCS116 in

- 607 the South Atlantic. *Environ. Microbiol.* **14**, 1133–1144, DOI: 10.1111/j.1462-2920.2011.02694.x
608 (2012).
- 609 **56.** Alonso-Sáez, L., Díaz-Pérez, L. & Morán, X. A. G. The hidden seasonality of the rare
610 biosphere in coastal marine bacterioplankton. *Environ. microbiology* **17**, 3766–3780, DOI:
611 10.1111/1462-2920.12801 (2015).
- 612 **57.** Chow, C. E. T. *et al.* Temporal variability and coherence of euphotic zone bacterial communi-
613 ties over a decade in the Southern California Bight. *ISME J.* DOI: 10.1038/ismej.2013.122
614 (2013).
- 615 **58.** Cram, J. A. *et al.* Seasonal and interannual variability of the marine bacterioplankton
616 community throughout the water column over ten years. *ISME J.* **9**, 563–580, DOI: 10.1038/
617 ismej.2014.153 (2015).
- 618 **59.** Gilbert, J. A. *et al.* Defining seasonal marine microbial community dynamics. *ISME J.* **6**,
619 298–308, DOI: 10.1038/ismej.2011.107 (2012).
- 620 **60.** Salter, I. *et al.* Seasonal dynamics of active SAR11 ecotypes in the oligotrophic Northwest
621 Mediterranean Sea. *ISME J.* DOI: 10.1038/ismej.2014.129 (2015).
- 622 **61.** Wu, W., Huang, B., Liao, Y. & Sun, P. Picoeukaryotic diversity and distribution in the
623 subtropical-tropical South China Sea. *FEMS Microbiol. Ecol.* **89**, 563–579, DOI: 10.1111/
624 1574-6941.12357 (2014).
- 625 **62.** Lin, Y. C. *et al.* Community Composition of Photosynthetic Picoeukaryotes in a Subtropical
626 Coastal Ecosystem, with Particular Emphasis on Micromonas. *J. Eukaryot. Microbiol.* DOI:
627 10.1111/jeu.12370 (2017).
- 628 **63.** Herfort, L., Peterson, T. D., Campbell, V., Futrell, S. & Zuber, P. *Myrionecta rubra* (*Meso-*
629 *dinium rubrum*) bloom initiation in the Columbia River estuary. *Estuarine, Coast. Shelf Sci.*
630 **95**, 440–446, DOI: 10.1016/j.ecss.2011.10.015 (2011).

- 631 **64.** Crawford, D. W., Purdie, D. A., Lockwood, A. P. & Weissman, P. Recurrent red-tides in
632 the Southampton Water estuary caused by the phototrophic ciliate *Mesodinium rubrum*.
633 *Estuarine, Coast. Shelf Sci.* **45**, 799–812, DOI: 10.1006/ecss.1997.0242 (1997).
- 634 **65.** Gustafson, D. E., Stoecker, D. K., Johnson, M. D., Van Heukelem, W. F. & Sneider, K.
635 Cryptophyte algae are robbed of their organelles by the marine ciliate *Mesodinium rubrum*.
636 *Nature* **405**, 1049–1052, DOI: 10.1038/35016570 (2000).
- 637 **66.** Qiu, D., Huang, L. & Lin, S. Cryptophyte farming by symbiotic ciliate host detected in situ.
638 *Proc. Natl. Acad. Sci.* **113**, 12208–12213, DOI: 10.1073/pnas.1612483113 (2016).
- 639 **67.** Mallin, M. A. Phytoplankton ecology of North Carolina estuaries. *Estuaries* **17**, 561–574,
640 DOI: 10.1016/j.crvi.2003.09.002 (1994).
- 641 **68.** Adolf, J. E., Yeager, C. L., Miller, W. D., Mallonee, M. E. & Harding, L. W. Environmental
642 forcing of phytoplankton floral composition, biomass, and primary productivity in Chesapeake
643 Bay, USA. *Estuarine, Coast. Shelf Sci.* **67**, 108–122, DOI: 10.1016/j.ecss.2005.11.030 (2006).
- 644 **69.** Johnson, M. D., Stoecker, D. K. & Marshall, H. G. Seasonal dynamics of *Mesodinium rubrum*
645 in Chesapeake Bay. *J. Plankton Res.* **35**, 877–893, DOI: 10.1093/plankt/fbt028 (2013).
- 646 **70.** Kim, S., Park, M. G., Moon, C., Shin, K. & Chang, M. Seasonal variations in phytoplankton
647 growth and microzooplankton grazing in a temperate coastal embayment, Korea. *Estuarine,
648 Coast. Shelf Sci.* **71**, 159–169, DOI: 10.1016/j.ecss.2006.07.011 (2007).

649 **Acknowledgments**

650 We would like to thank the members of the Singapore Laboratory for Integrated Microbial
651 Ecology and DHI explorer crews for help with sampling. We thank *Indigo V Expeditions* for
652 providing sampling equipment, Christaline George, Sandra Kolundžija, Rosalie Chai and Halimah
653 Razali for help with sampling, Daniela Drautz-Moses for technical assistance with Illumina
654 sequencing, and Chen Shuang for nutrient analysis. This study was supported by the National
655 Research Foundation, Prime's Minister's Office, Singapore under its Marine Science Research
656 and Development Programme (Awards No. MSRDP-P13).

657 **Author contributions statement**

658 CC and FML conceived the study. CC, WW and AK collected and processed the samples. CC,
659 FML, WW, DV and AL analyzed the data. CC and FML drafted the manuscript. CC, DV, AL, PM,
660 FML edited the final version of the paper.

661 **Additional information**

662 ***Data availability***

663 Raw sequencing data have been deposited to GenBank under Bioproject number PRJNA497851.
664 Processed data and scripts are available from <https://github.com/slimelab/Singapore-metabarcodes>.

665 ***Competing interests***

666 The authors declare no competing interests.

Table 1. ANOSIM analysis.

Sample set	Grouping	R	Significance
All sites	Strait	0.802	0.001
	Site	0.652	0.001
	Monsoon	0.005	0.34
Singapore Strait	Site	-0.072	0.498
	Monsoon	0.270	0.002
Johor Strait	Site	0.421	0.001
	Monsoon	-0.026	0.675

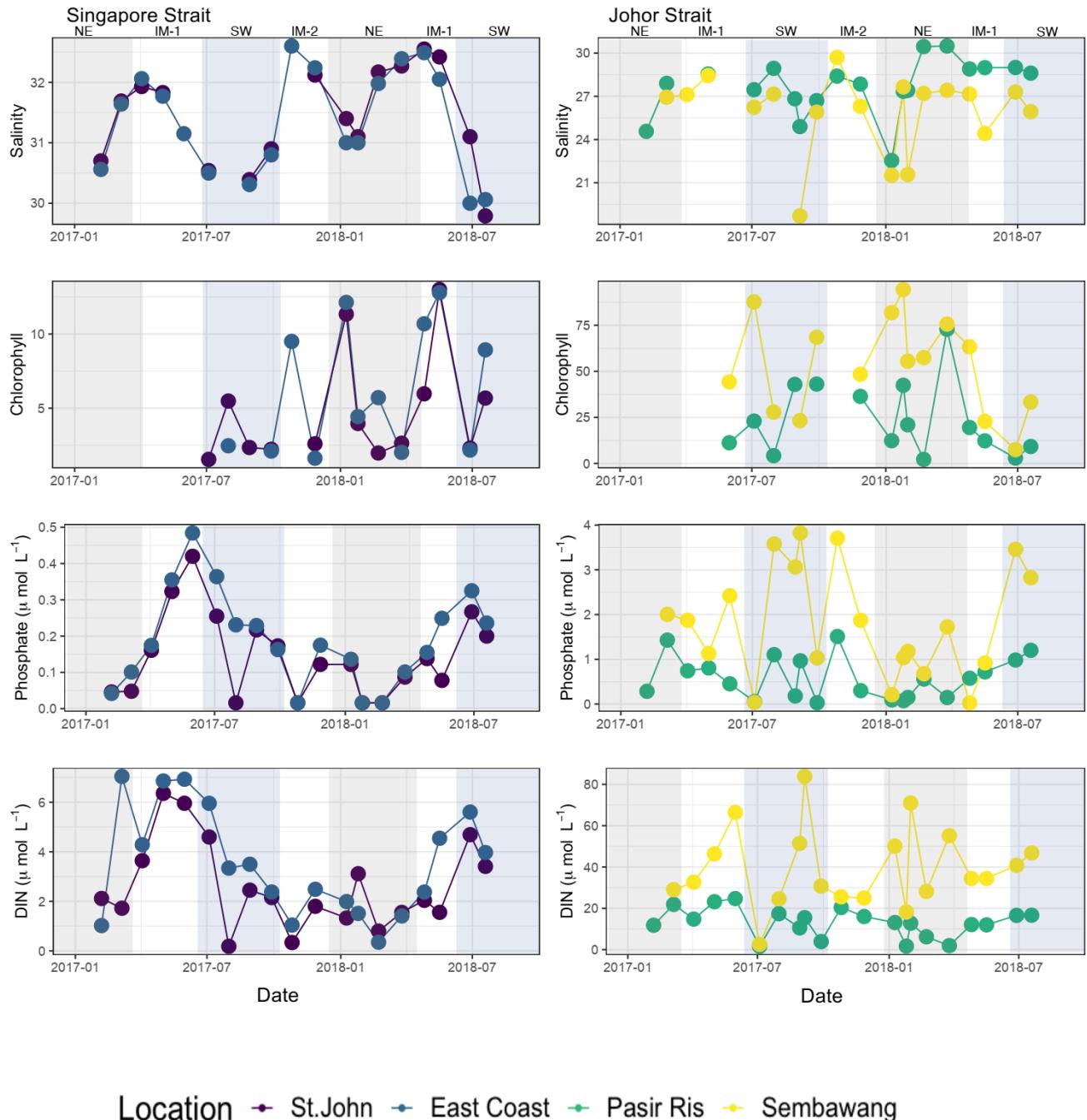


Figure 1. Salinity, chlorophyll *a*, phosphates and DIN (dissolved inorganic nitrogen) during the 18-month time series in Singapore coastal waters. Highlights in grey and blue represent NE and SW monsoon, respectively.

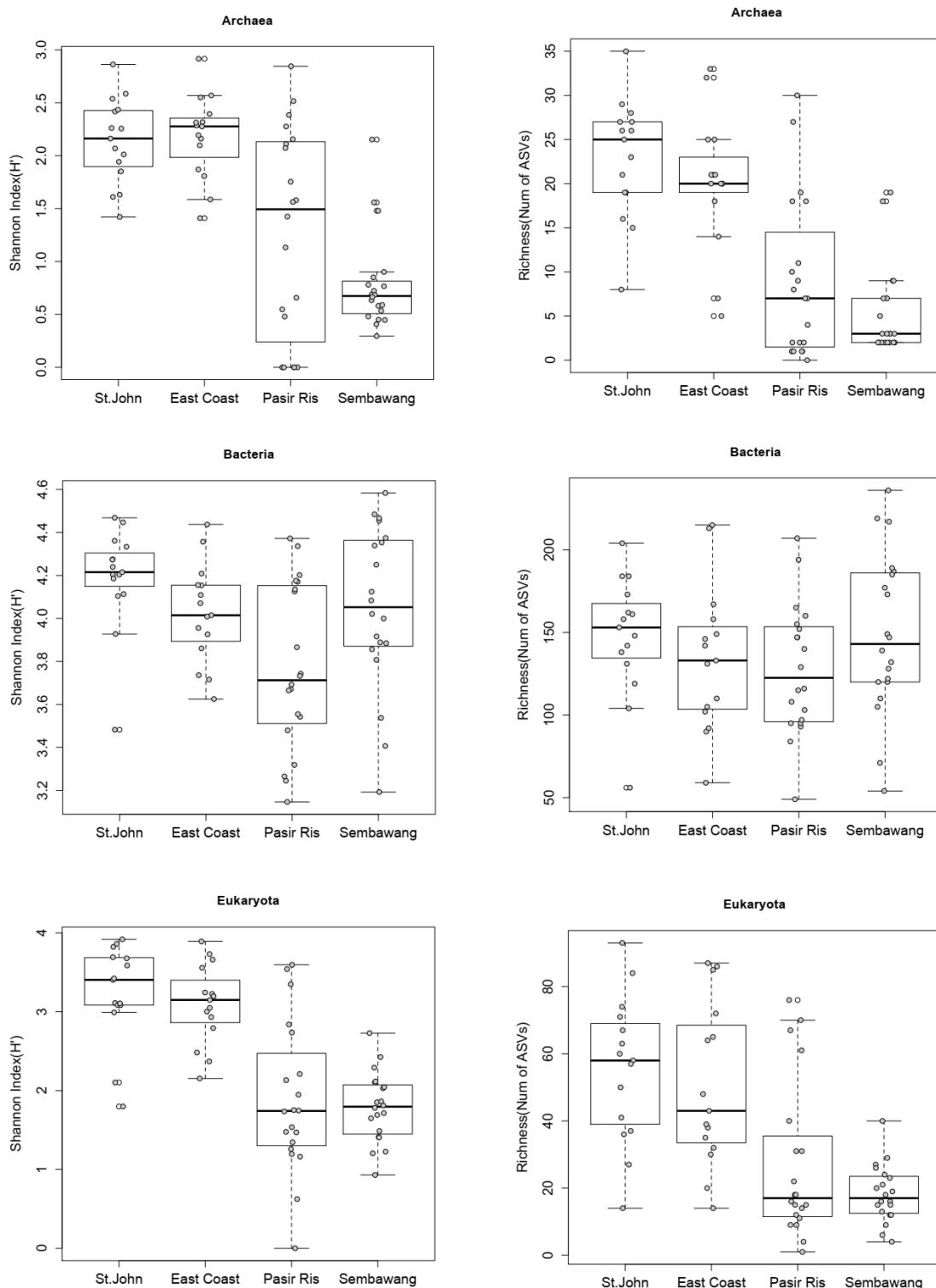


Figure 2. Shannon Index (H' , left) and Richness (Number of ASVs, right) of microbial communities at 4 stations in Singapore coastal waters for Archaea (top,A-B), Bacteria (middle,C-D) and Eukaryota (bottom, E-F). Empty dots represent outliers.

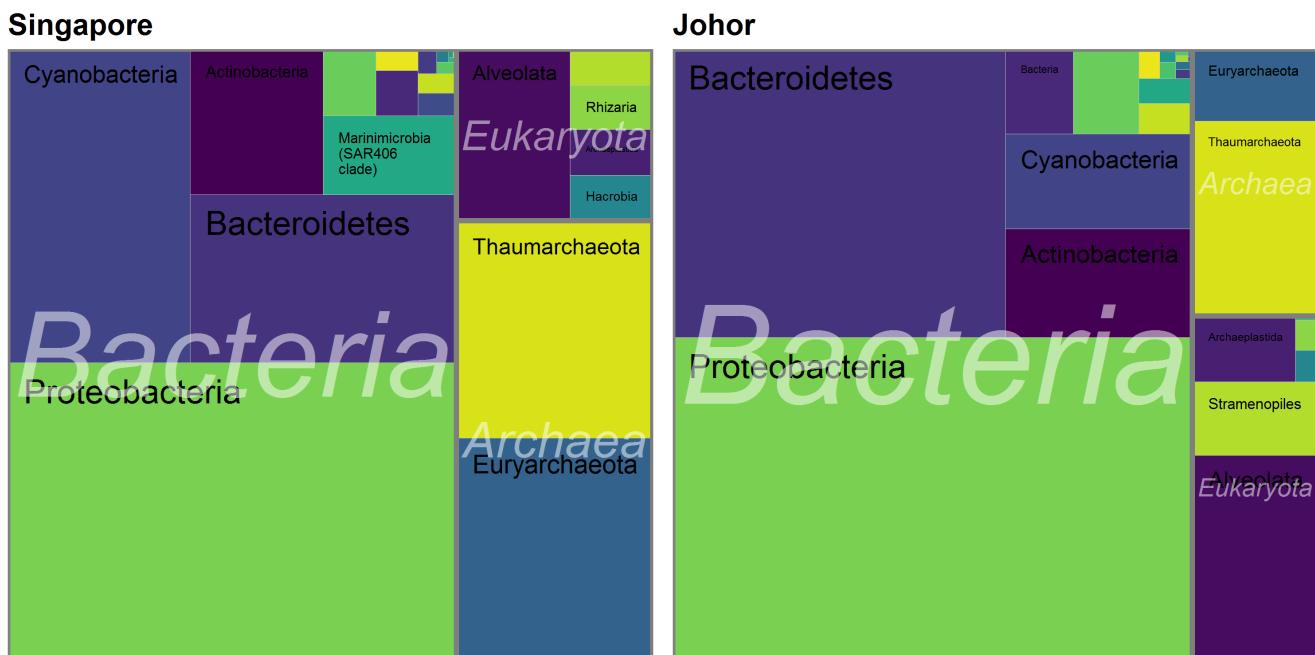


Figure 3. Tree map with the major taxonomic groups in Singapore and Johor Straits.

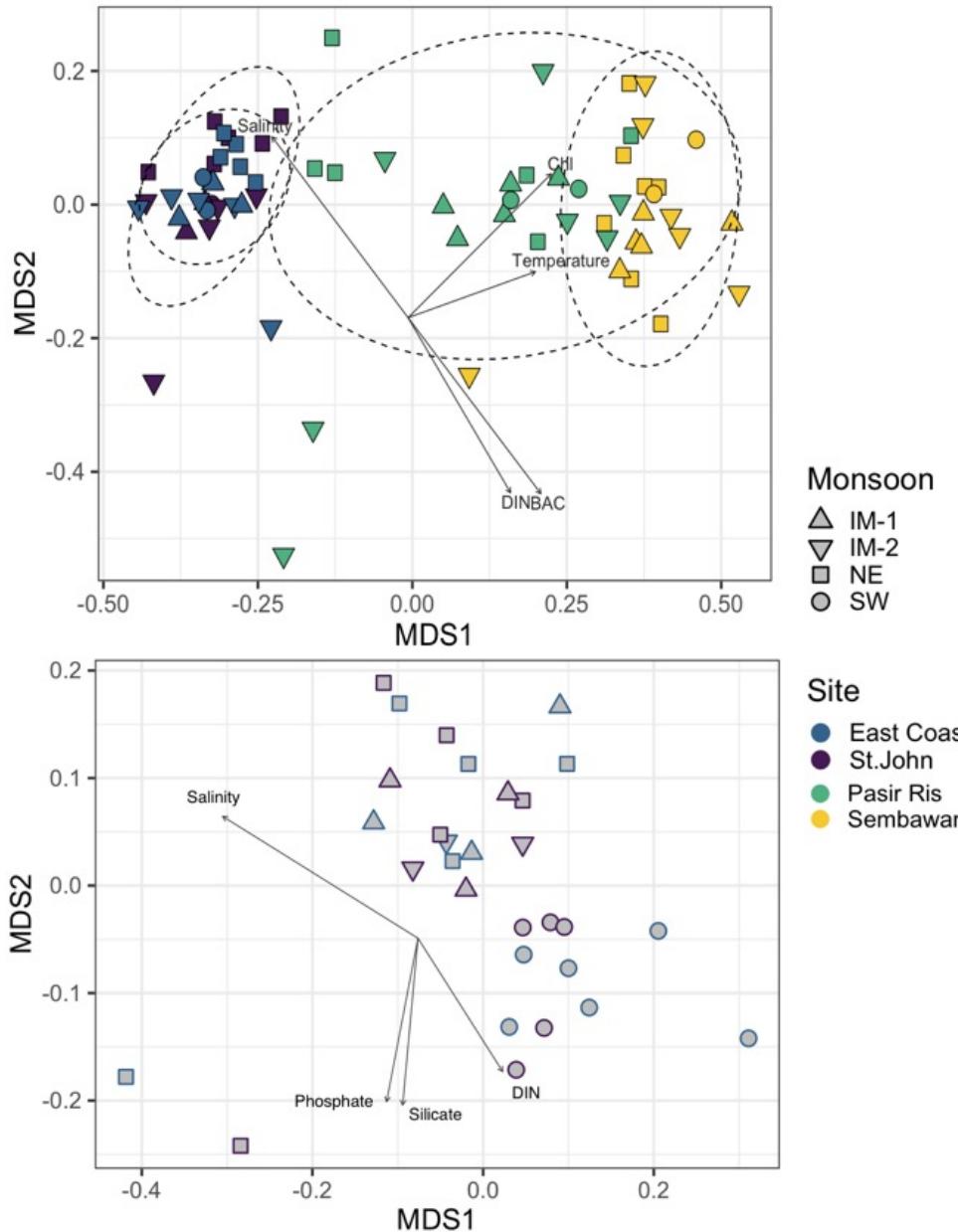


Figure 4. Non-metric multidimensional scaling (nMDS) analysis of Bray-Curtis similarity index. Each sample is labelled by location and monsoon period. Arrows represent environmental parameters with $p < 0.001$ when performing an *envfit* analysis. The ellipses represent 95% confidence intervals for samples collected at the same station. Top. All stations. Bottom. Only Singapore Strait samples (STJ and EC stations).

667 **1 Supplementary material**

668 All supplementary material is available at <https://github.com/slimelab/Singapore-metabarcodes>

669 **1.1 Supplementary Data**

670 Supplementary Data S1: (monsoonpaper_env_data.csv) List of samples collected with environ-
671 mental parameters.

672 Supplementary Data S2 (Singapore ASV_table.xlsx). Sheet ASV: List of ASVs with taxonomic
673 affiliation, assignment bootstrap values for eukaryotes, sequence and number of reads in each
674 sample (only samples from four stations, STJ, EC, SBW and PR are used in this paper). Sheet
675 Blast eukaryotes: Summary of BLAST assignments against PR²³⁰ and GenBank for Eukaryota
676 ASVs (nuclear 18S rRNA).

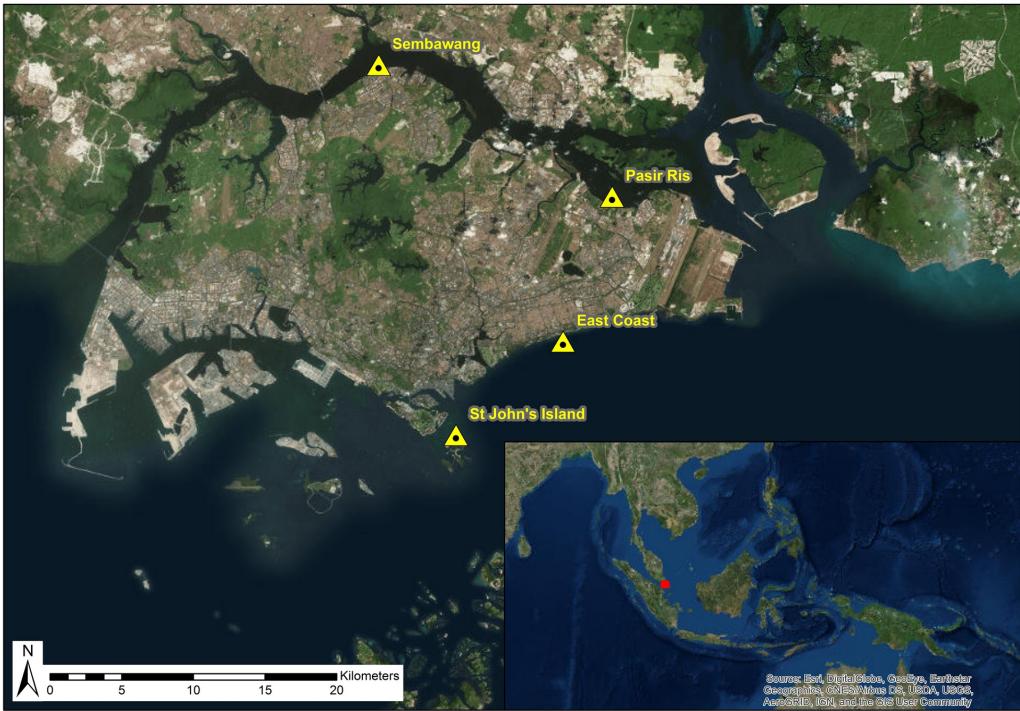


Figure S1. Map of Singapore showing the four stations sampled in this study.

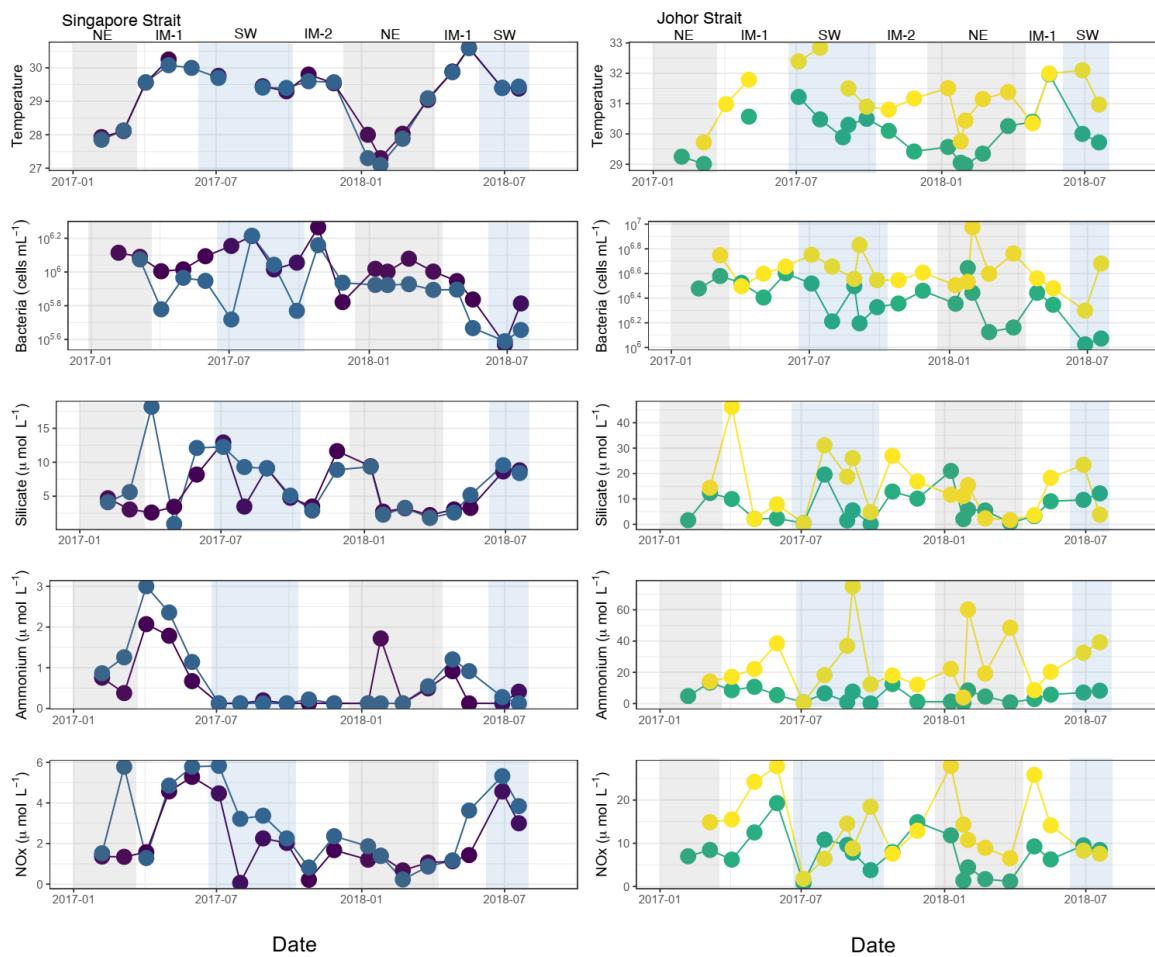


Figure S2. Temperature, bacteria abundance, silicates, ammonium and NO_x during the 18-month time series in Singapore coastal waters. Highlights in grey and blue represent NE period and SW periods respectively.

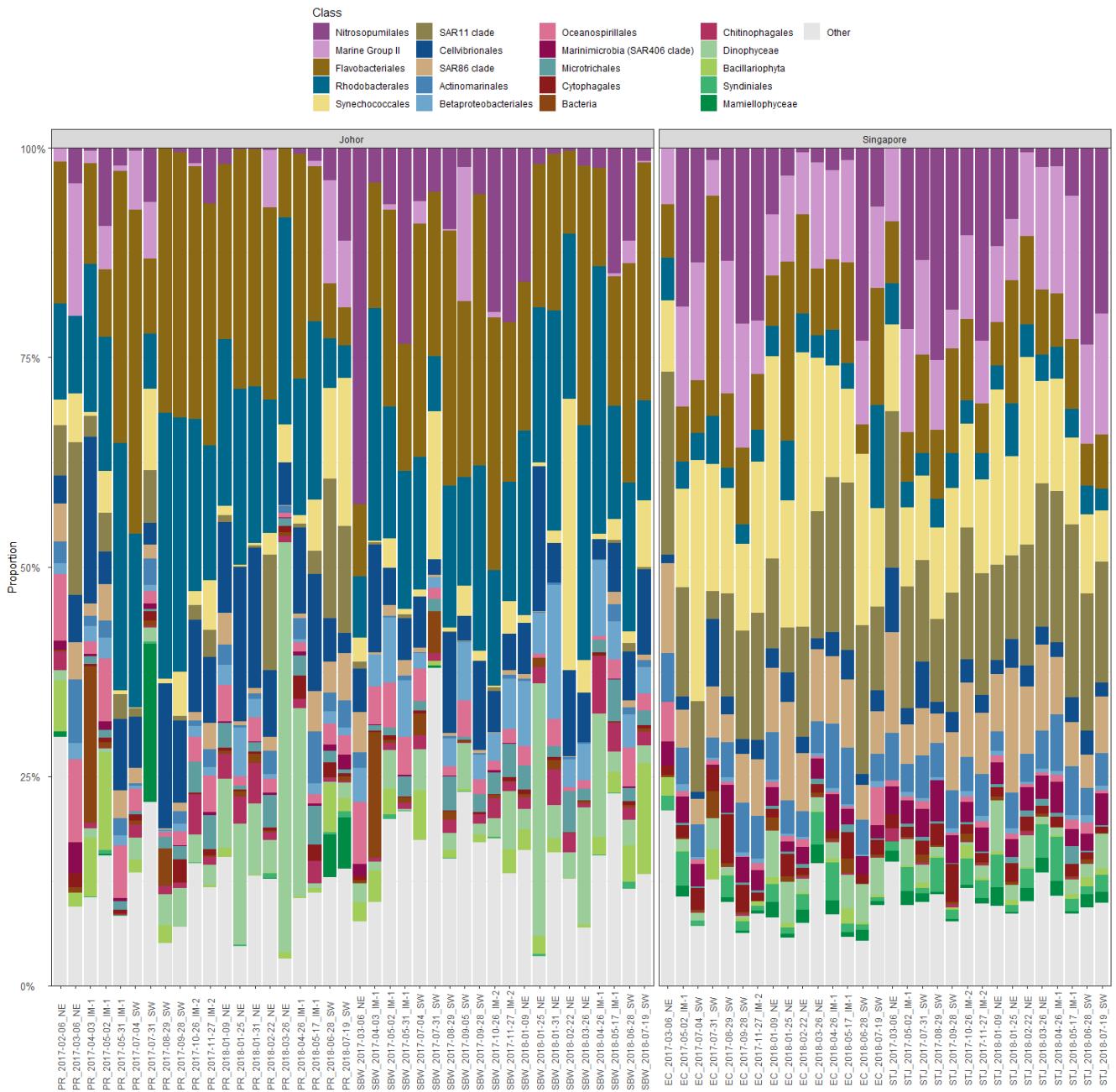


Figure S3. Taxonomic composition of the microbial communities in Singapore coastal waters focusing on the top 20 most abundant groups (class level) for the three domains (Archaea, Bacteria, Eukaryota).

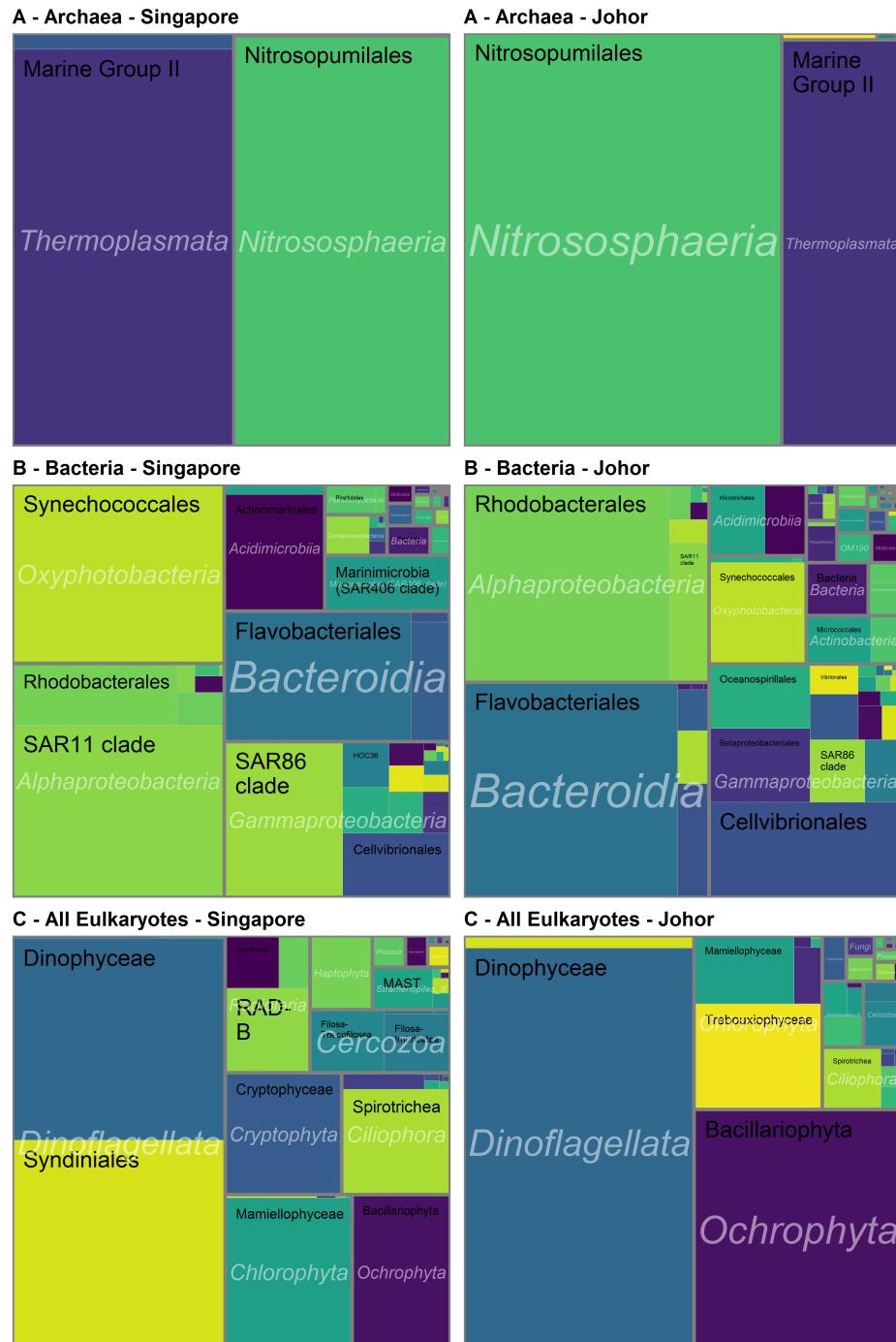


Figure S4. Tree map with the major taxonomic groups in Singapore and Johor Straits for the 3 domains: Archaea, Bacteria and Eukaryota.



Figure S5. Thirty most abundant Archaea and Bacteria (top) and Eukaryota (bottom) ASVs for Singapore (left) and Johor (right) straits. ASVs that are unique to one strait are labelled with an asterisk.

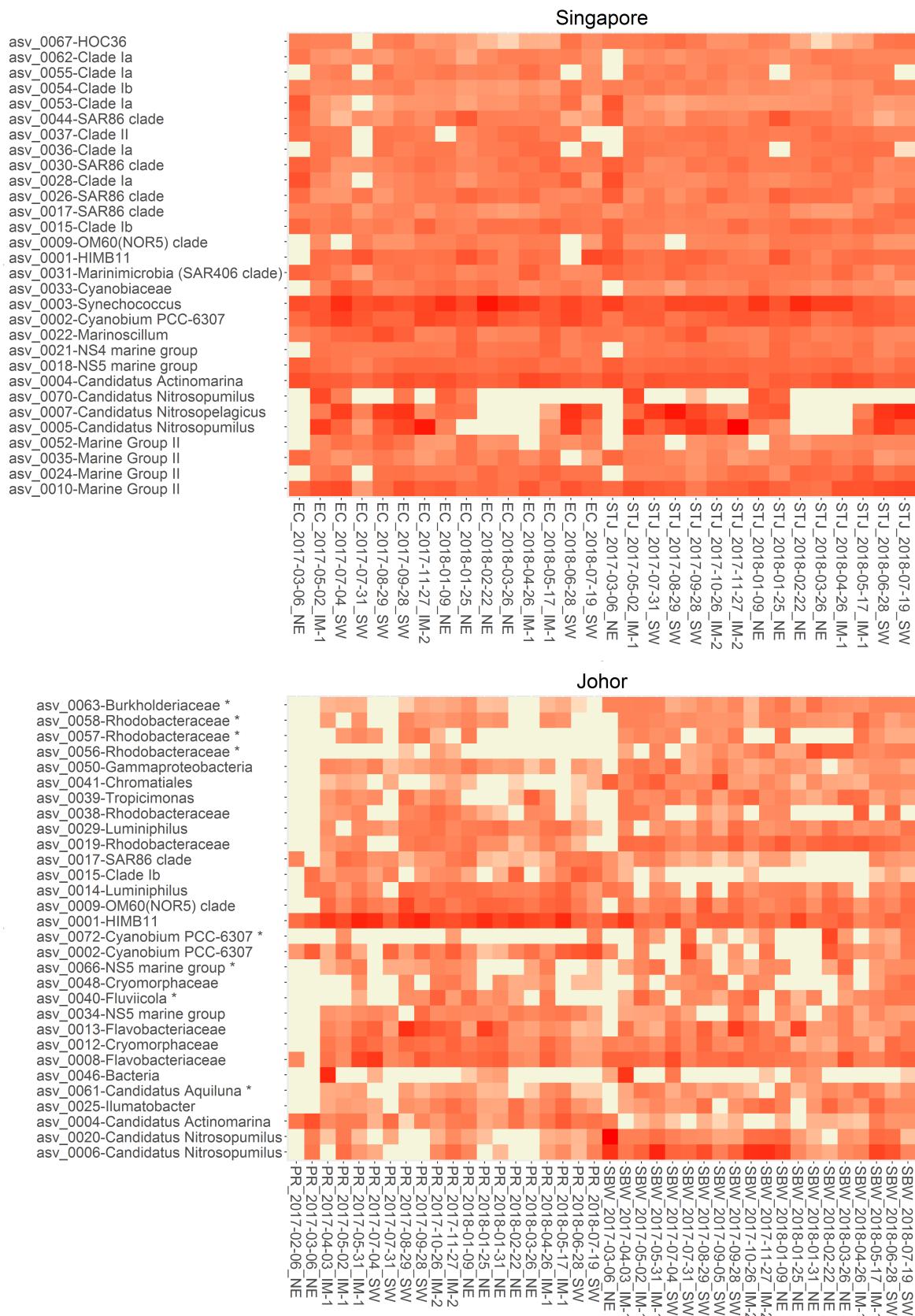


Figure S6. Sample heatmap of the thirty most abundant Archaea and Bacteria ASVs for Singapore (top) and Johor (bottom) straits. ASVs that are unique to one strait are labelled with an asterisk.

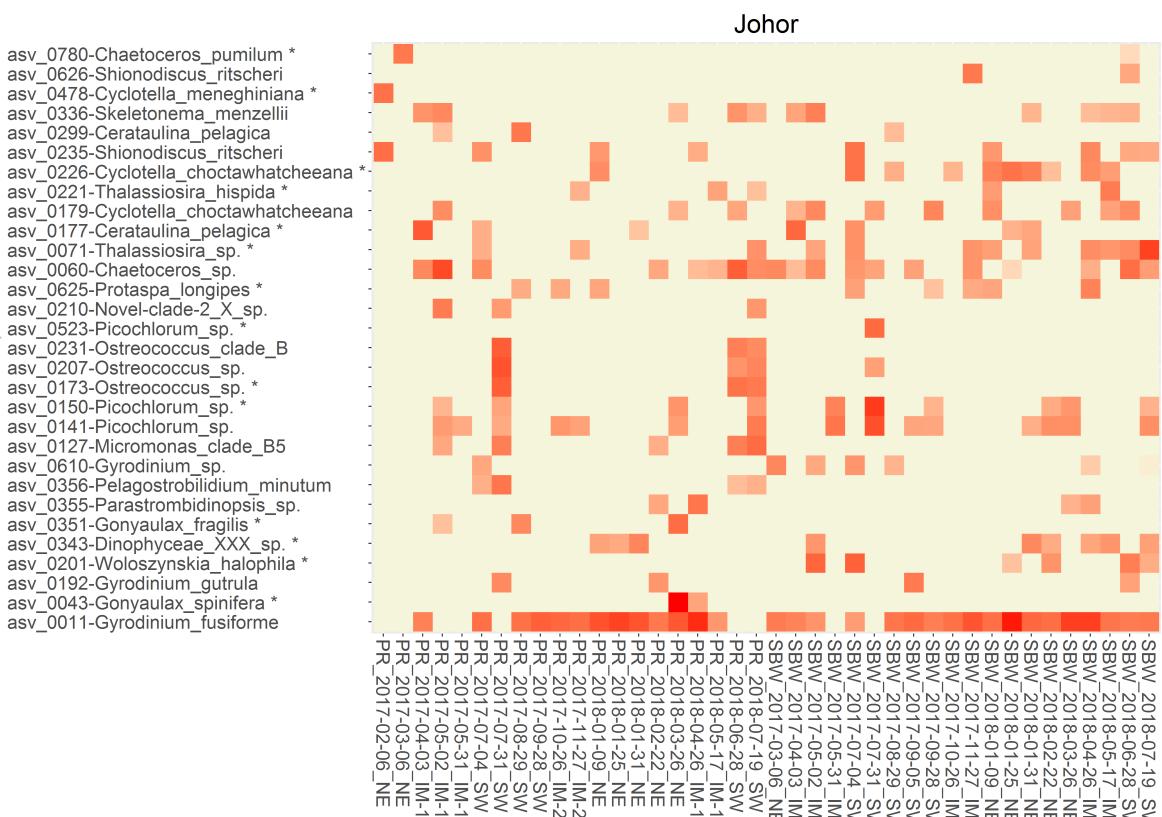
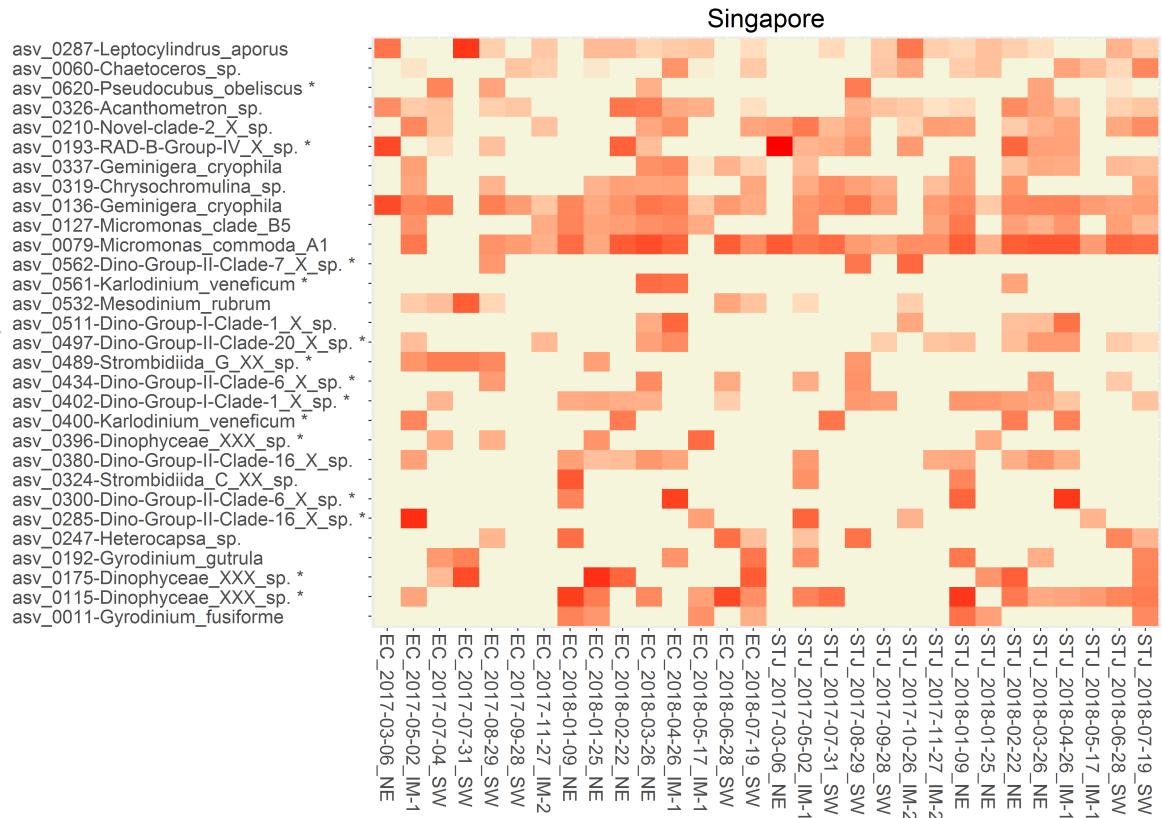


Figure S7. Sample heatmap of the thirty most abundant Eukaryota ASVs for Singapore (top) and Johor (bottom) straits. ASVs that are unique to one strait are labelled with an asterisk.

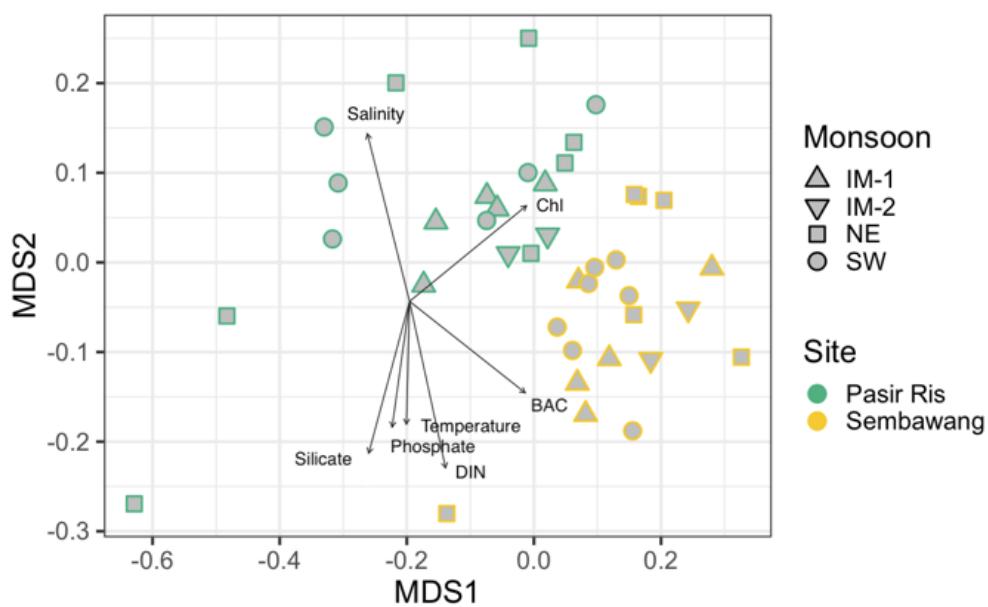


Figure S8. Non-metric multidimensional scaling (nMDS) of Bray-Curtis similarity index for Johor Strait (PR and SBW stations). Each sample is labelled based on location and monsoon period. The arrows represent environmental parameters with $p < 0.05$ when performing an *envfit* analysis.



Figure S9. Top. Gneiss analysis for Singapore strait based on monsoon period. Bottom. Members of the communities that drive the difference between the NE and SW monsoons based on DESeq2 analysis³³ using a threshold p-value < 0.01. Symbol transparency is inversely proportional to ASV rank.

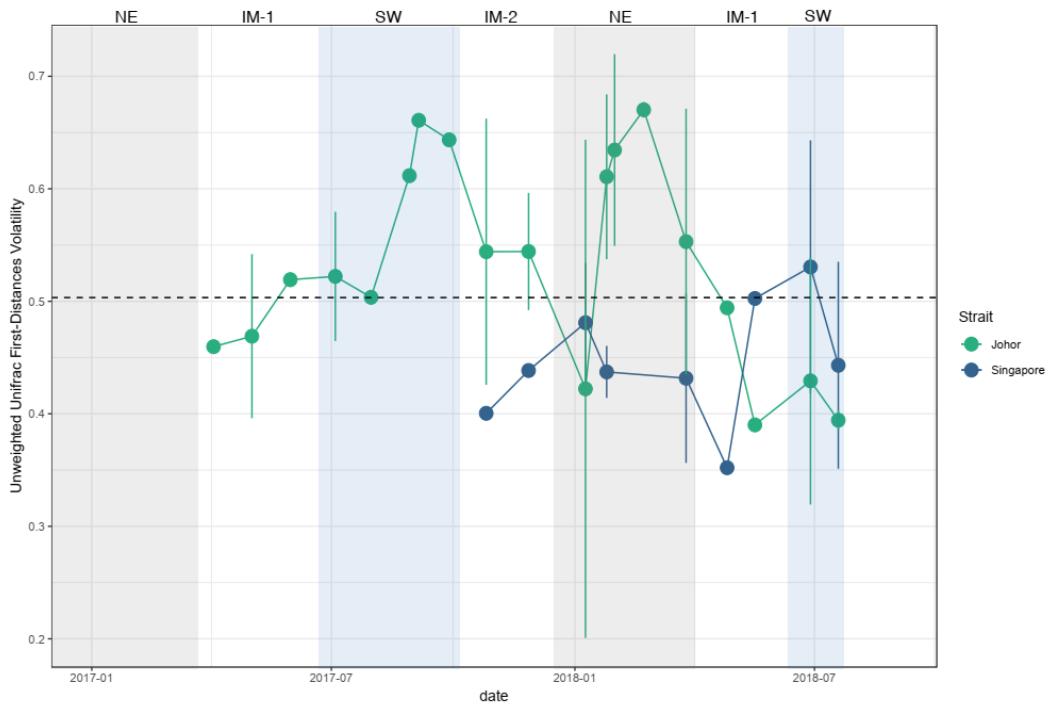


Figure S10. First-distances volatility values computed on the unweighted Unifrac distance matrix of the two Straits. The dotted line indicates the global average and error bars the individual sample dispersion from the Strait average

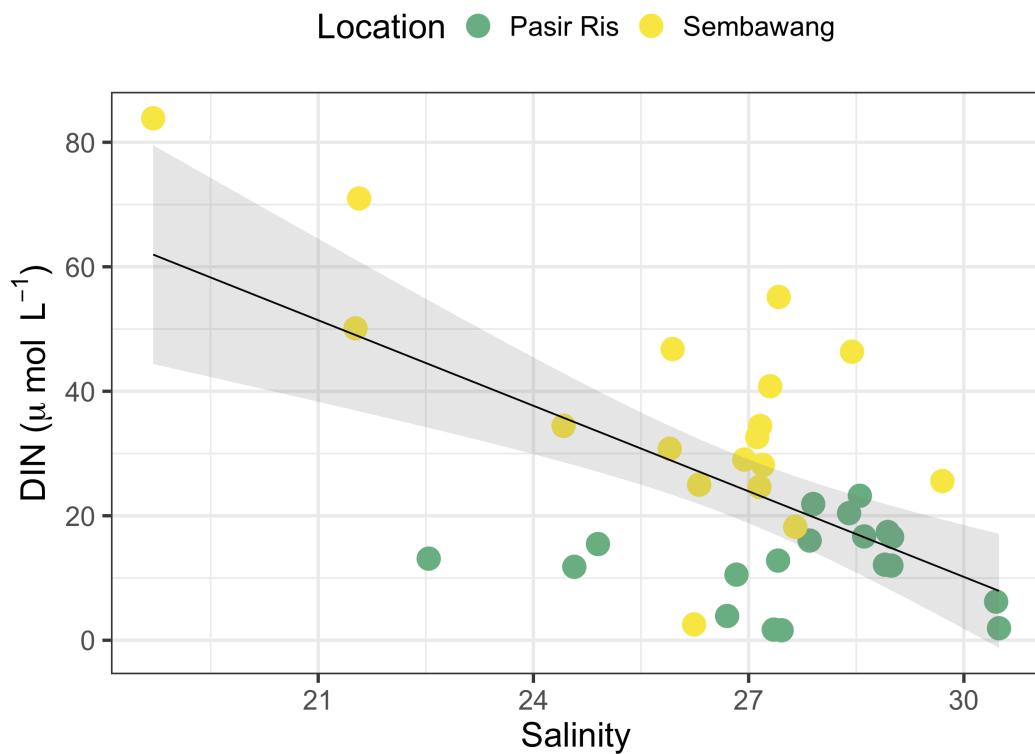


Figure S11. Correlation between DIN and salinity in Johor strait.