

1 **Insights on aquatic microbiome of the Indian Sundarbans mangrove**
2 **areas**

3 **Paltu Kumar Dhal^{1*}, Germán A. Kopprío², Astrid Gärdes²**

4 ¹Department of Life Science and Biotechnology, Jadavpur University, ²Tropical Marine
5 Microbiology, Department of Biogeochemistry and Geology, Leibniz Center for Tropical Marine
6 Research, Bremen, Germany

7

8

9

10 ***Corresponding address:**

11 Dr. Paltu Kumar Dhal

12 Department of Life Science and Biotechnology,

13 Jadavpur University

14 188, Raja Subodh Chandra Mallick Road, West Bengal,

15 Kolkata 700032

16

17

18

19

20

21

22

23

24

25 **Insights on aquatic microbiome of the Indian Sundarbans mangrove**

26 **areas**

27 **Paltu Kumar Dhal^{1*}, Germán A. Kopprio², Astrid Gärdes²**

28 ¹Department of Life Science and Biotechnology, Jadavpur University, ²Tropical Marine

29 Microbiology, Department of Biogeochemistry and Geology, Leibniz Center for Tropical Marine

30 Research, Bremen, Germany

31 **ABSTRACT:**

32 **Background:** Anthropogenic perturbations have strong impact on water quality and ecological
33 health of mangrove areas of Indian Sundarbans. Diversity in microbial community composition
34 is important causes for maintaining the healthy of the mangrove ecosystem. However, microbial
35 communities of estuarine water in Indian Sundarbans mangrove areas and environmental
36 determinants that contribute to those communities were seldom studied.

37 **Methods:** Nevertheless, this study attempted first to report bacterial and archaeal communities
38 simultaneously in the water from Matla River and Thakuran River of Maipith coastal areas more
39 accurately using 16S rRNA gene-based amplicon approaches. Attempt also been made to assess
40 the capability of the environmental parameters for explaining the variation in microbial
41 community composition.

42 **Results:** Our investigation indicates the dominancy of halophilic marine bacteria from families
43 *Flavobacteriaceae* and OM1 clade in the water with lower nutrient load collected from costal
44 regions of a small Island of Sundarban Mangroves (ISM). At higher eutrophic conditions,
45 changes in bacterial communities in Open Marine Water (OMW) were detected, where some of

46 the marine hydrocarbons degrading bacteria under families *Oceanospirillaceae* and
47 *Spongiibacteraceae* were dominated. While most abundant bacterial family *Rhodobacteracea*
48 almost equally (18% of the total community) dominated in both sites. Minor variation in the
49 composition of archaeal community was also observed between OMW and ISM. Redundancy
50 analysis indicates a combination of total nitrogen and dissolved inorganic nitrogen for OMW and
51 for ISM, salinity and total nitrogen was responsible for explaining the changes in their respective
52 microbial community composition.

53 **Conclusions:** Our study can serve as baseline approaches, which should focused on how do
54 multiple environmental/anthropogenic stressors (salinity, pollution, eutrophication, land-use)
55 affect the estuary water and consequently the microbial communities in concert. However,
56 systematic approaches with more samples for evaluating the effect of environmental pollutions
57 on mangrove microbial communities are recommended.

58 **Keywords:** Indian Sundarbans, mangrove water, Microbial community structure, 16S Illumina
59 sequencing.

60

61 **1. INTRODUCTION:**

62 Sundarbans, the largest mangrove forest of the world, is situated in the joint delta of Ganges,
63 Brahmaputra and Meghna rivers at Bay of Bengal (Ghosh and Bhadury 2017). This UNESCO
64 World Heritage site comprises the Indian state of West Bengal and southwest Bangladesh
65 (Bhattacharyya *et al.* 2015). Livelihood and well being of millions of people live in and around
66 of Sundarbans, depends on its status and ecological services. Despite its high ecological and
67 economical values, Sundarbans is seriously threatened by different anthropogenic activities.
68 Since the early 19th century, landscapes of Sundarbans have also been changing due to saline and

69 freshwater imbalances. Water quality of this ecosystem is largely affected by sewage pollutant
70 originated from industries located upstream and urban areas of West Bengal. Sewage entering
71 into coastal water contains diverse chemical and microbiological pollutants and a wide variety of
72 organic and inorganic wastes (Mitra *et al.* 2009, Banerjee *et al.* 2017), driving changes on its
73 ecological and physiological health.

74 Microbial communities of mangroves are responsible for nutrient cycling and play a vital role in
75 productivity, conservation and rehabilitation of mangrove ecosystems (Holguin *et al.* 2006).
76 Therefore, understanding their responses to environmental changes is essential to predict changes
77 in service-provisioning (Mishra *et al.* 2012). Several recent studies described the microbial
78 community compositions of surface sediments and water of Indian Sundarban mangrove areas.
79 Surface sediments this area dominated with *Deltaproteobacteria* followed by
80 *Gammaproteobacteria*, *Alphaproteobacteria*, *Betaproteobacteria*, and *Epsilonproteobacteria*
81 under phylum *Proteobacteria*. Abundant bacterial orders are *Desulfobacteriales*,
82 *Desulfuromonadales*, *Myxococcales*, and *Bdellovibrionales*. (Basak *et al.* 2015 ; Chakraborty *et*
83 *al.* 2015; Basak *et al.* 2016). While bacterioplankton communities in the water of this region
84 were found to be abundant with *Gammaproteobacteria* and *Alphaproteobacteria*. At the family
85 level dominancy of *Hyphomicrobiaceae*, *Rhodobacteraceae*, *Pseudomonadaceae*,
86 *Erythrobacteraceae*, *Kordiimonadaceae*, *Hyphomonadaceae*, and *Ruminococcaceae* were
87 observed (Ghosh and Bhadury 2017; Ghosh and Bhadury 2018; Ghosh and Bhadury 2019).
88 However, sampling locations of those studies on microbial communities in the Indian Sundarban
89 mangrove water mainly restricted near to an island (Sagar Island), therefore the major
90 conclusions of these studies were made based on a single site specific with limited number of
91 samples. Moreover, not much effort has been made to investigate the archaeal community of this

92 region except single report by Bhattacharyya *et al.* (2015) on surface and subsurface sediments
93 of Indian Sundarban mangrove forest. Moreover, the above studies have rarely analyzed the
94 bacterial and archaeal community structures of the same samples at the same time. Therefore,
95 our knowledge on those communities as well as information on how they are controlled by
96 environmental parameters is limited. In order to assess the microbial communities of marine
97 ecosystem via high-throughput sequencing of amplified 16S rRNA genes with high resolution
98 and fidelity, it is extremely important to select the proper primer set that can't underestimated or
99 overestimated any common marine taxa (Parada *et al.* 2016). However, this will be the first
100 attempt to visualize the accurate and well-resolved picture of bacterial and archaeal communities
101 simultaneously of marine water in Sundarbans mangrove using next-generation amplicon
102 sequencing of the 16S rRNA gene using recently developed 515F-Y/926R primers. We also tried
103 to explore the environmental determinants that contribute to the variation of their microbial
104 communities. This study will provide baseline knowledge on microbial ecology of the World
105 Heritage site and serve as a baseline for monitoring programs and predicting changes at impacted
106 sites.

107 **2. MATERIAL AND METHODS**

108 **2.1 Study sites and Sample collection:**

109 In the present study sampling was conducted in the Sundarbans mangrove ecosystem that shared
110 between India and Bangladesh and lies in the Ganga-Brahmaputra-Meghna (GBM) delta. This
111 mangrove ecosystem contains over 102 islands with a network complex of many rivers, rivulets,
112 and creeks (Das 2017). Sampling was carried out at two different locations on Thakuran River -
113 Matla River estuarine complex of Maipith coastal areas in the Indian Sundarbans mangroves

114 during March 2017 (Fig 1). They are designated as Island of Sundarban Mangroves (ISM) and
115 Open Marine Water (OMW). ISM is an uninhabited small island with lesser anthropogenic
116 disturbance situated in Thakuran and Matla river complex in low-lying costal plain. This river
117 has no reports for perennial fresh water source (Das 2017). Water from three different costal
118 regions of this island is selected for sampling. OMW is an open marine site around same regions,
119 which is supposedly continually influenced by the wastewaters from upstream regions of Matla
120 River. Three independent replicated water samples (1 L) from each of three different sampling
121 sites of both ISM and OMW were collected collected in sterile containers and immediately
122 stored at a chilled box until further laboratory analysis.

123 **2.2 Environmental Parameters and nutrients analysis:**

124 Physiochemical parameters (salinity and pH) of all collected samples were measured using
125 Eureka 2 Manta multiprobe (Eureka Environmental Engineering, Texas, USA). Total 50 mL of
126 each sample was filtered through a 0.7 μm syringe filter and poisoned with 200 μL of 3.5 g/100
127 mL HgCl_2 solution for nutrient analysis. Dissolved organic carbon (DOC) and total nitrogen
128 (TN) as well as dissolved inorganic nutrients (DIN), that including nitrate and nitrite (NO_x),
129 phosphate (PO_4^{3-}), and silicate (Si) were analyzed using a continuous flow analyzer (Flowsys by
130 Unity Scientific, Brookfield, USA) (<https://doi.pangaea.de/10.1594/PANGAEA.889699>). Each
131 of the samples was filtered through a 0.7 μm pore size GF/F filter (GE Healthcare Bio-Sciences,
132 Pittsburgh, PA, USA) before DOC and nutrients measurements. The inorganic nutrients (nitrite,
133 nitrate, phosphate, and silicate) were measured using spectro-photometrical analysis with a
134 Flowsys continuous flow analyzer (Systea, Anagni, Italy). For measurements of DOC, the
135 filtered samples were acidified with concentrated HCl ($\text{pH} < 2$) and analyzed by high-temperature
136 oxic combustion method using a TOC- V_{CPN} analyzer (Shimadzu, Mandel, Canada). Seawater

137 standards (Hansell laboratory, RSMAS University Miami, USA) were used for calibration and
138 quality control, and ultrapure water as a blank.

139 **2.3 DNA Extraction, PCR Amplification, and Illumina MiSeq osequencing**

140 From each site, water (1 L) was filtered (0.2 μ m) and DNA was extracted using the
141 PowerWater® DNA Isolation Kit according to the manufacturer's instructions (MoBio
142 Laboratories Inc., Carlsbad, CA, USA). DNA concentrations and purity were measured
143 spectrophotometrically. Presence of bacterial and archebacterial 16S rRNA gene was in the
144 extracted metagenome was verified following previous method (Dhal and Sar, 2014). In order to
145 classify taxonomically both bacterial and archaeal community structure simultaneously,
146 sequencing of V4–V5 hypervariable regions of 16S rRNA gene were generated using primers
147 515F-Y (5'-GTGYCAGCMGCCGCGGTAA-3') and 926R (5'-
148 CCGYCAATTYMTTTRAGTTT-3') (Parada *et al.*, 2016) on the Illumina MiSeq platform
149 (CeBiTec Bielefeld, Germany), in a 2 \times 300 bp paired-end run.

150 **2.4 High Throughput sequencing data processing :**

151 Primer sequences were removed using *cutadapt* from the raw paired-end reads (Martin, 2011).
152 The primer-trimmed sequences are available on Sequence Read Archive (SRA) (accession no.
153 SRP144285). Sequences were quality trimmed with *trimmomatic* v0.32 (Bolger *et al.* 2014) using
154 a sliding window of 4 bases and a minimum average quality of 15, and merged with PEAR
155 v0.9.5 (Zhang *et al.* 2014). Quality-filtered sequences were clustered into OTUs with *swarm*
156 algorithm using default parameters (Mahe *et al.* 2014). One single representative sequence per
157 OTU was taxonomically classified with SINA (SILVA Incremental Aligner; v1.2.11; Silva
158 reference database release 132) at a minimum alignment similarity of 0.9, and a last common
159 ancestor consensus of 0.7 (Pruesse *et al.* 2012). OTUs that were unclassified on the domain level

160 and those matching to chloroplast and mitochondrial sequences were excluded from the analysis
161 using well-standardized r script (Kunda *et al.* 2018). The final OTU tables are accessible at
162 (<https://doi.pangaea.de/10.1594/PANGAEA.890757>).

163 **2.5 Statistical analysis:**

164 Principal component analysis (PCA) was performed to cluster the sampling sites based on their
165 environmental parameters. Differences in environmental parameters among ISM and OMW were
166 assessed using general linear mixed models (GLMM) with sampling station as a random factor
167 (Kuznetsova *et al.* 2017).

168 Alpha-Diversity indices were calculated to assess richness and evenness of the microbial
169 communities (Hill 1973) in the studied samples, based on repeated random subsampling of the
170 amplicon data sets after randomly rarefying the data set to the minimum library size (50517
171 sequences). Significant differences in alpha-diversity indices between the studied stations were
172 determined by using the non-parametric Kruskal test followed by p-value adjusted Wilcoxon
173 tests (Hassenruck *et al.* 2016).

174 To assess the differences in community structure between two sampling sites (beta-diversity),
175 Bray–Curtis dissimilarities were calculated using the relative OTU abundances and also non-
176 metric multidimensional scaling (NMDS) plot was produced. Analysis of similarity (ANOSIM)
177 was calculated to assess the separation of bacterial communities between the two sites. P-values
178 of all multiple pairwise comparisons were adjusted using the false discovery rate (fdr) correction
179 method by Benjamini and Hochberg (1995). In order to evaluate the environmental parameters as
180 drivers of the variations in community compositions, redundancy analysis (RDA) was used with
181 centered log ratio (clr)-transformed sequence counts using the R function `aldex.clr` of the
182 ALDEx2 package via median values of 128 Monte-Carlo instances (Fernandes *et al.* 2014). To

183 compare the explanatory power of all measured environmental parameters, additional RDA
184 models were constructed with environmental parameters as predictors. Forward model selection
185 was used after checking for variance inflation to determine which of parameters would be
186 included in the RDA models. When more than one parameter was included, pure effects were
187 also tested accounting for the variation explained by the other factors in the model. Collinearity
188 among predictors was determined via Variance inflation factors (VIFs) of the individual
189 parameters. All of the parameters in any of the RDA models displayed VIFs less than 10. The
190 adjusted R^2 is provided as goodness-fit-stat. All statistical analyses were conducted in R using
191 the core distribution, version 3.3.2 and R-Studio, version 1.0.153, with following packages: vegan
192 (Oksanen *et al.* 2016), lmerTest for the GLMM (Kuznetsova *et al.* 2017), ALDEx2 (Oksanen *et*
193 *al.* 2016) and multcomp (Fernandes *et al.* 2014).

194

195 **3. Result:**

196 **3.1 Environmental characterization**

197 Environmental parameters (pH, salinity) and nutrients (including DOC, TN, NO_x, nitrate, DIN,
198 phosphate, and silicate) concentrations for all samples were measured (**Table 1**). Samples were
199 slightly alkaline (pH 8.0 to 8.7) in nature. The GLMMs analysis indicated that the measured
200 water nutrients that differed significantly among the two sampling station were mainly TN, DIN
201 and PO₄³⁻ (Table 1; **Table S1**). The PCA ordination (**Fig 2**) showed that first two principal
202 components (PC1 and PC2) represented 74.4% of data variation among sites. PC1 alone
203 represents 60.1% of total variation and influenced by most of the measured parameters, while pH
204 showed a strong correlation with PC2 (**Fig S1**). Noteworthy to mention, the samples were
205 separated into two clusters by PC1. One cluster is mainly composed with the samples from ISM

206 (except TH2.3 of OMW) and other cluster accommodating samples collected from OMW. This
207 ordination probably indicates elevated eutrophication in samples from the OMW compared to
208 ISM.

209 **3.2. Microbial communities**

210 Total numbers of reads generated per sample ranged between 50517 to 90468 (after merged)
211 corresponding to 3,644 to 6,470 swarmed, non-singleton 16S OTUs (**Fig S2**). After rarefaction,
212 numbers of bacterial and archaeal OTUs ranged between 3390 to 5415 and 37 to 91, respectively
213 (**Fig S2**). None of the measured diversity indices (Average Shannon diversity and inverse
214 Simpson indices) were found significance differences in between OMW and ISM (**Fig 3; Table**
215 **S2**) indicated by Kruskal - Wilcoxon test ($p > 0.5$), although values varied considerably.

216 The microbial community of marine estuary water from Sundarbans was dominated with bacteria
217 occupying more than 96 % of total community and archaea represented only 4 %. Bacterial
218 assemblage of two different sites this area showed to have a distinct community. In class level,
219 among the dominant bacterial groups, *Flavobacteria* (ISM: 15.9 % vs OMW: 8.6%),
220 *Alphaproteobacteria* (ISM: 29.5% vs OMW: 28%), and *Acidimicrobiia* (ISM: 6.6% vs OMW:
221 5.0%) were dominant in ISM while OMW was dominated with mainly with
222 *Gammaproteobacteria* (ISM: 22.6% vs OMW: 35.3%) (**Fig S3**). At higher taxonomic resolution
223 levels (**Fig 4**), bacterial communities were composed with a total of 474 and 915 different
224 bacterial family and genus, respectively. The most dominant bacterial family was
225 *Rhodobacteraceae* (18.6%), almost equally distributed between studied two sites. Other
226 dominant bacterial families of ISM were *Flavobacteriaceae* (14.8%) and OM1 clade (5.2%)
227 whereas in OMW, *Oceanospirillaceae* (16%) and *Spongiibacteraceae* (4 %) were the most
228 abundant.

229 We observed dominance of Marine Group (MG I) (currently known as *Thaumarchaeota*) and
230 *Euryarchaea* MG II in archaeal community assemblages with 78.5% and 16.9% of relative
231 abundance, respectively while the presence of *Woesearchaeota* (2.1%) was also evident (**Fig 5**).
232 MG I was found in relatively higher abundant at ISM constituting on average 82.1% of
233 sequences as opposed to 75% at OMW. *Euryarchaeota* MG II comprised about 20.3% at OMW
234 compared to 13.7% at ISM. Among total twenty-nine (29) archaeal genera, *Candidatus*
235 *Nitrosopumilus* and *Candidatus Nitrosopelagicus* accounted for the 40.7% and 21.4% of total
236 relative abundance, respectively (**Fig S4**).

237 **3.3 Environmental drivers of bacterial communities**

238 At OTU resolution level also, distinct microbial communities were observed between OMW and
239 ISM based on changes in community structure (beta diversity) which is quantified by non-metric
240 multidimensional scaling (NMDS) plot by calculating Bray–Curtis dissimilarity (**Fig 6**). This
241 pattern is confirmed by the ANOSIM test that indicated a significant difference in microbial
242 community structure between ISM and OMW (ANOSIM, $R = 0.24$, $p < 0.001$). Redundancy
243 analyses attempted to identify the water quality parameters that had strong explanatory power for
244 microbial communities. We observed that total nitrogen (TN) and dissolved inorganic nitrogen
245 (DIN) accounted for almost 10% of the variability in microbial community of OMW where TN
246 alone explain 6% variation of microbial community (RDA, $R^2 = 0.06$, $F_{(1,7)} = 1.34$, $p < 0.05$). In
247 contrast, salinity and TN explained approximately 9% of the variability in community
248 composition of ISM (RDA, $R^2 = 0.06$, $F_{(1,7)} = 1.34$, $p < 0.05$) and alone salinity responsible for
249 explaining 7% microbial variation of this site (**Table 2**).

250 **4. Discussion:**

251 The pH values (8.0- 8.7) indicates the water of ISM and OMW slightly alkaline in nature which
252 supports the previous findings in similar samples from Sundarbans Mangrove forest areas (
253 Sarkar and Bhattacharya, 2010). Such ranges of pH may be attributed by the buffering capacity
254 of water that support high biological activity (Balasubramanian and Kannan 2005; Sarkar and
255 Bhattacharya 2010). The water of Sundarbans is characterized by elevated salinity values in line
256 with previous reports (Balasubramanian and Kannan 2005). The long-term changes in water
257 properties in the eastern part of Sundarbans, sampling regions of our study, indicating increased
258 trends on salinity and pH (Bhattacharyya *et al.* 2015). Our result shows differences in measured
259 environmental parameters between two sites represented by three sampling stations and nine
260 samples each leading to their segregation into two clusters (in PCA analysis) along with their
261 sampling sites. This ordination as a result of different nutrient loads and this is reflected by
262 potential eutrophication in water from OWM. Influences from the Thakuran and Matla rivers
263 reported to have a strong impact on the estuary water quality represented by OWM (Mitra *et al.*
264 2009, Banerjee *et al.* 2017), that also reflected in our study. The perennial discharge in Thakuran
265 and Matla Estuary from upstream regions brings in a high suspended matter load throughout the
266 year (Sarkar and Bhattacharya 2010). Those estuaries severely contaminated with huge organic
267 load and sediment flux originated from upstream domestic sewage, aquaculture, intensive
268 trawling activities, agricultural runoff as well as soil erosion (Sarkar *et al.* 2004; Mukherjee *et al.*
269 2009).

270 Because of the relevance of microbial community of Indian Sundarbans, several investigators
271 attempted on surface sediments samples (Ghosh *et al.* 2010; Basak *et al.* 2015; Bhattacharyya *et*
272 *al.* 2015; Eloie-Fadrosch *et al.* 2015; Basak *et al.* 2016; Mallick *et al.* 2018) as well as recently on
273 water column of this regions (Ghosh and Bhadury 2018; Ghosh and Bhadury 2019) using 16S

274 rRNA gene metagenomic approaches. Unlike previous studies, this investigation attempted to
275 assess both the bacterial and archaeal community at a same time of water from relative less
276 anthropogenic disturbance sites using an efficient primer set to target V4-V5 variable region of
277 16S rRNA gene in order to avoid the problems of underestimated or overestimated common
278 marine taxa (Parada et al. 2016), therefore our investigation gives more accurate and well-
279 resolved picture of microbial communities of these sites.

280 Although insignificant differences, elevated trends of α -diversity of the marine estuary water
281 samples (OMW) might be an indications of relatively rich bacterial community compared to ISM
282 of Sundarbans might be attributed toward their elevated eutrophication. This observation was
283 supported by previous reports that indicate a higher diversity and equitability in the human
284 impacted estuary because of proliferation of several different microorganisms (Nogales et al.,
285 2011; Borin et al., 2009).

286 At phyla level, bacterial assemblages of the studied samples (specially OMW) showed similarity
287 with the previously reported bacterial community of marine sediments and water samples of
288 Sundarban Mangrove areas (Basak *et al.* 2015; Basak *et al.* 2016; Ghosh and Bhadury 2017;
289 Ghosh and Bhadury 2018). The most dominant bacterial family presents both the station with
290 almost equal proportion is *Rhodobacteraceae*. Dominancy of members of this family in marine
291 water microbial community previously reported and known to be associated with marine
292 phytoplankton blooms where it plays important role in transforming phytoplankton-derived
293 organic matter (Ghosh *et al.* 2010; Buchan *et al.* 2010; Simon *et al.* 2017). The abundant OTUs
294 of *Rhodobacteraceae* are classified as anoxygenic phototrophs *Nautella*, reported to serve an
295 indicator of marine eutrophication, predominantly found in higher eutrophic OMW samples. The
296 other dominant one under the same family is marine heterotrophs *Ruegerias* (almost equally

297 distributed among both ISM and OMW) serve as the model of marine sulfur and carbon cycle
298 (Buchan *et al.*, 2005, Dang *et al.* 2008; Doberva *et al.* 2017).

299 Interestingly enrichment of several OTUs from the *Flavobacteria* is observed in the oligotrophic
300 ISM compared to the eutrophic OMW. They are specialized in utilization of biopolymers and
301 organic substances in oligotrophic environment i.e., when organic substances present even at
302 very low concentrations. Higher abundance of these polymers degrading bacteria biopolymers
303 promotes the growth of heterotrophic bacteria at oligotrophic environments (Lauro et al 2009;
304 Jessen *et al.* 2013; Kegler *et al.* 2018), might plays central role in microbial ecology of ISM by
305 creating the supporting environment for heterotrophs/copiotrophic organisms in relatively
306 oligotrophic conditions. The dominant OTUs under family *Flavobacteriaceae* were mainly
307 classified as *Aureimarina* and NS5 marine group genus. Roles of *Aureimarina* in marine
308 biogeochemistry has not been investigated much although few studies reported on their
309 presences in coastal seawater and saline estuarine (Teeling *et al.* 2012; Campbell *et al.* 2015).
310 However, this is the first report of their abundance in marine water of Sundarbans. The NS5
311 marine group which are equally dominated in both the studied sites are reported to be ubiquitous
312 in the seawater-related samples and known for phytoplankton-derived macromolecules (Tanaka
313 *et al.* 2008; Dupont *et al.* 2012 Seo et al, 2017).

314 This investigation identified *Actinobacteria* constituted a predominant fraction both in OMW
315 and ISM but elevated amount in the later samples. Bacteria under this group are consisted of
316 both copiotrophic and oligotrophic members with abundance in oligotrophic marine
317 environments (Ho et al 2017; Nouioui et al 2018). As marine *Actinobacteria* are the richest
318 sources of secondary metabolites thus, have been well reported as potential sources of bioactive
319 compounds (Manivasagan et al 2014). Therefore, their abundance in our studied sites (specially

320 ISM) would be potential hotspot for isolating bioactive molecules from Indian Sundarban
321 mangrove forest. The OM1 clade (dominating in ISM), an uncultured Actinobacterial clade,
322 frequently recovered from various marine environments with higher abundance at near coastal
323 sites than open marine areas however supports our reports (Giovannoni and Stingl 2005; Morris
324 *et al.* 2012; Ngugi and Stingl 2018). The dominant OTUs of this family were classified as
325 *Candidatus Actinomarina*. Those photoheterotrophs are one of smallest free-living prokaryotes
326 are reported to be ubiquitous in marine systems. Not many reports are described their role in the
327 marine biogeochemical cycle and first reports on their presence in Sundarban mangrove areas.

328 The OTUs affiliated to families *Oceanospirillaceae* and *Spongiibacteraceae* of
329 *Gammaproteobacteria* showed increase abundance in the impacted site OMW. Bacteria from
330 these families are known to be present in eutrophic marine environments. They are known as
331 polymer degraders and can utilize polyhydroxy alkananoate compounds and proteorhodopsin, for
332 harvesting an additional energy, supports their living in eutrophic water samples (Mizuno *et al.*
333 2015; Hoffmann *et al.* 2017; Ribicic *et al.* 2018). The dominant OTUs of *Oceanospirillaceae* are
334 affiliated to chemoheterotrophic genus *Marinobacterium*. Their presence in mangroves as well as
335 surface seawater have already been described in previous studies and known to be associated
336 with hydrocarbon biodegradation (Dos Santos *et al.* 2011; Spring *et al.* 2015). The other
337 dominated bacterial family in the samples from OMW is *Spongiibacteraceae*. They comprise
338 mainly marine bacteria known as Oligotrophic Marine *Gammaproteobacteria* (OMG) group
339 (Cho and Giovannoni 2004; Yilmaz *et al.* 2016). We recorded the dominant OTUs of this family
340 are affiliated with BD1-7 which is a cosmopolitan group of *Gammaproteobacteria* is mostly
341 autochthonous, reported to inhabits at diverse marine habitats (Huggett and Rappé 2012; Spring
342 *et al.* 2015; Zhou *et al.* 2018). In line with previous reports this investigation, therefore, indicates

343 proliferation of bacterial groups under *Gammaproteobacteria* with respond to increased nutrient
344 concentrations in estuary (Nogales et al 2011).

345 However, in contrast to sediments reported in previous investigations, an archaeal community of
346 marine waters in the Sundarban mangroves is dominated with *Thaumarchaeota* Marine Group
347 (MG I) and *Euryarchaea* MG II. The chemolithoautotrophic MG I which are in higher in number
348 on ISM are responsible for oxidation of ammonia and showed ability in inorganic carbon fixation
349 (Haro-Moreno et al 2017) thus important players in global Carbon (C) and Nitrogen (N)
350 biogeochemical cycles. In contrast enrichment of heterotrophic MG II, which is more dominate
351 in OMW, also are observed in the marine aquatic environment (Liu et al 2017). Their abilities in
352 organic carbon degradation and in the photic zone, they acquired energy in presences of light.
353 Dominant OTUs of MGI group are affiliated with ammonia-oxidizing archaeal, *Candidatus*
354 *Nitrosopumilus* and *Candidatus Nitrosopelagicus*, play important roles in nitrogen and carbon
355 cycling of marine ecosystem (Bhattacharyya *et al.* 2015) however, this investigation reports first
356 on their present in this areas. Therefore, the biological and geochemical processes at estuary
357 water habitats in the Indian Sundarban Mangrove areas have likely influenced the archaeal
358 community structure.

359 **5. CONCLUSION**

360 This investigation provides the first details description of bacterial and archeal communities
361 concurrently of Thakuran and Matla river complex of Maipith coastal areas in the Indian
362 Sundarbans mangroves areas. Our study indicates along with the elevated level of average pH
363 and salinity, the open marine water (OMW) showed eutrophication probably leads to an
364 observed bacterial shift toward more copiotrophic and photoheterotrophic bacterial
365 (*Oceanospirillaceae* and *Spongiibacteraceae*) and archaeal community (*Euryarchaea* MG II)

366 and compared to the more oligotrophic microbial community (*Aureimarina*, NS5 marine group,
367 OM1 clade and *Thaumarchaeota* MG I) of costal water of a small Island of Sundarban
368 Mangroves (ISM). These microbial assembles thus might represent key players in
369 biogeochemical cycle of this mangrove and the studied areas represent a hotspot for bacterial
370 having potential to produce the commercially important secondary metabolites. This
371 investigation also reports that total nitrogen and dissolved inorganic nitrogen are the major
372 environmental contributors on determining the microbial communities for OMW and for ISM it
373 is combination of total nitrogen and salinity.

374 However, given the rising burden on Indian mangrove coastal ecosystems, this study suggests
375 that sewages from urban areas lacking proper treatment can alter microbial communities that
376 may play vital role in biogeochemical cycle of mangrove ecosystem and consequently may
377 impact on the climate in the tropical country.

378 **6. ACKNOWLEDGEMENTS**

379 Foremost PKD would like to thank NAM S&T Centre New Delhi, India and ZMT, Bremen for
380 fellowship selection. PKD also acknowledge to Jadavpur University for a travel grant and
381 granting the leave to avail this fellowship. We sincerely thank the faculty and technical staff of
382 ZMT Bremen for their generous help to carry this research work. Many thanks to Dr Halina
383 Tegetmeyer of CeBiTec Bielefeld, Germany for sequences. The assistance of conducting analyses
384 at the laboratories at the ZMT in Bremen by Matthias Birkicht, Sonja Peters and Achim Meyer
385 also acknowledge.

386

387 **6. FUNDING**

388 This work was part of the Joint NAM S&T Centre – ZMT Bremen Fellowship in Tropical
389 Coastal Marine Research that was supported by the Leibniz Zentrum für Marine Tropenforschung
390 (ZMT), Bremen, Germany.

391

392 7. REFERENCES:

393

- 394 1. Balasubramanian R, Kannan L. Physico-chemical characteristics of the coral reef environs of the
395 Gulf of Mannar Biosphere Reserve, India. *Int J Ecol Environ Sci* 2005;31: 273-8.
- 396 2. Banerjee K, Gatti RC, Mitra A. Climate change-induced salinity variation impacts on a stenocious
397 mangrove species in the Indian Sundarbans. *Ambio* 2017;46: 492-9.
- 398 3. Basak P, Majumder NS, Nag S et al. Spatiotemporal analysis of bacterial diversity in sediments of
399 Sundarbans using parallel 16S rRNA gene tag sequencing. *Microb Ecol* 2015;69: 500-11.
- 400 4. Basak P, Pramanik A, Sengupta S et al. Bacterial diversity assessment of pristine mangrove
401 microbial community from Dhulibhashani, Sundarbans using 16S rRNA gene tag sequencing.
402 *Genom Data* 2016;7: 76-8.
- 403 5. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to
404 multiple testing. *Journal of the royal statistical society Series B (Methodological)* 1995: 289-300.
- 405 6. Bhattacharyya A, Majumder NS, Basak P et al. Diversity and Distribution of Archaea in the
406 Mangrove Sediment of Sundarbans. *Archaea* 2015;2015: 968582.
- 407 7. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
408 *Bioinformatics* 2014;30: 2114-20.
- 409 8. Borin S, Brusetti L, Daffonchio D, Delaney E & Baldi F. Biodiversity of prokaryotic communities
410 in sediments of different sub-basins of the Venice lagoon. *Res Microbiol* 2009;160:307–314
- 411 9. Buchan A, González JM, Moran MA. Overview of the marine Roseobacter lineage. *Applied and
412 environmental microbiology* 2005;71: 5665-77.
- 413 10. Campbell AM, Fleisher J, Sinigalliano C et al. Dynamics of marine bacterial community diversity of
414 the coastal waters of the reefs, inlets, and wastewater outfalls of southeast Florida.
415 *MicrobiologyOpen* 2015;4: 390-408.
- 416 11. Chakraborty A, Bera A, Mukherjee A et al. Changing bacterial profile of Sundarbans, the world
417 heritage mangrove: impact of anthropogenic interventions. *World Journal of Microbiology and
418 Biotechnology* 2015;31: 593-610.
- 419 12. Cho J-C, Giovannoni SJ. Cultivation and growth characteristics of a diverse group of oligotrophic
420 marine Gammaproteobacteria. *Applied and Environmental Microbiology* 2004;70: 432-40.
- 421 13. Dang H, Zhang X, Sun J et al. Diversity and spatial distribution of sediment ammonia-oxidizing
422 crenarchaeota in response to estuarine and environmental gradients in the Changjiang Estuary and
423 East China Sea. *Microbiology* 2008;154: 2084-95.
- 424 14. Das, Gautam. (2016). Geomorphic Environments of the Thakuran River of the Sunderbans, India.
425 *Earth Science India*. 9. 974-8350. 10.31870/ESI.09.3.2016.9.
- 426 15. Dhal, P. K.; Sar, P. Microbial communities in uranium mine tailings and mine water sediment from
427 Jaduguda U mine, India: A culture independent analysis *J. Environ. Sci. Health, Part A:
428 Toxic/Hazard. Subst. Environ. Eng.* 2014, 49, 694– 709 DOI: 10.1080/10934529.2014.865458
- 429 16. Doberva M, Stien D, Sorres J et al. Large Diversity and Original Structures of Acyl-Homoserine
430 Lactones in Strain MOLA 401, a Marine Rhodobacteraceae Bacterium. *Frontiers in microbiology*
431 2017;8: 1152.
- 432 17. Dos Santos HF, Cury JC, Do Carmo FL et al. Mangrove bacterial diversity and the impact of oil
433 contamination revealed by pyrosequencing: bacterial proxies for oil pollution. *PLoS One* 2011;6:

- 434 e16943.
- 435 18. Dupont CL, Rusch DB, Yooseph S et al. Genomic insights to SAR86, an abundant and uncultivated
436 marine bacterial lineage. *The ISME journal* 2012;6: 1186.
- 437 19. Fernandes AD, Reid JN, Macklaim JM et al. Unifying the analysis of high-throughput sequencing
438 datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by
439 compositional data analysis. *Microbiome* 2014;2: 15.
- 440 20. Ghosh A, Bhadury P. Insights into bacterioplankton community structure from Sundarbans
441 mangrove ecoregion using Sanger and Illumina MiSeq sequencing approaches: A comparative
442 analysis. *Genom Data* 2017;11: 39-42.
- 443 21. Ghosh A, Bhadury P. Exploring biogeographic patterns of bacterioplankton communities across
444 global estuaries. *MicrobiologyOpen*. 2019;8:e741.
- 445 22. Ghosh A, Bhadury P. Investigating monsoon and post-monsoon variabilities of bacterioplankton
446 communities in a mangrove ecosystem. *Environmental Science and Pollution Research* 2018;25:
447 5722-39.
- 448 23. Ghosh A, Dey N, Bera A et al. Culture independent molecular analysis of bacterial communities in
449 the mangrove sediment of Sundarban, India. *Saline Systems* 2010;6: 1.
- 450 24. Giovannoni SJ, Stingl U. Molecular diversity and ecology of microbial plankton. *Nature* 2005;437:
451 343.
- 452 25. Hassenruck C, Fink A, Lichtschlag A et al. Quantification of the effects of ocean acidification on
453 sediment microbial communities in the environment: the importance of ecosystem approaches.
454 *FEMS Microbiol Ecol* 2016;92: fiw027.
- 455 26. Haro-Moreno JM, Rodriguez-Valera F, López-García P, Moreira D, Martin-Cuadrado A-B.
456 (2017). New insights into marine group III Euryarchaeota, from dark to light. *ISME J* 11: 1102–
457 1117.
- 458 27. Hill MO. Diversity and evenness: a unifying notation and its consequences. *Ecology* 1973;54: 427-
459 32.
- 460 28. Ho Adrian, Paolo Di Lonardo D., Bodelier Paul L. E. Revisiting life strategy concepts in
461 environmental microbial ecology, *FEMS Microbiology Ecology*, Volume 93, Issue 3, March 2017.
- 462 29. Hoffmann K, Hassenruck C, Salman-Carvalho V et al. Response of Bacterial Communities to
463 Different Detritus Compositions in Arctic Deep-Sea Sediments. *Front Microbiol* 2017;8: 266.
- 464 30. Holguin G, Gonzalez-Zamorano P, de-Bashan LE et al. Mangrove health in an arid environment
465 encroached by urban development--a case study. *Sci Total Environ* 2006;363: 260-74.
- 466 31. Huggett MJ, Rappé MS. Genome sequence of strain HIMB30, a novel member of the marine
467 Gammaproteobacteria. *Journal of bacteriology* 2012;194: 732-3.
- 468 32. Jessen C, Lizcano JFV, Bayer T et al. In-situ effects of eutrophication and overfishing on
469 physiology and bacterial diversity of the Red Sea coral *Acropora hemprichii*. *PLoS One* 2013;8:
470 e62091.
- 471 33. Kegler HF, Hassenrück C, Kegler P et al. Small tropical islands with dense human population:
472 differences in water quality of near-shore waters are associated with distinct bacterial communities.
473 *PeerJ* 2018;6: e4555.
- 474 34. Kunda P, Dhal PK, Mukherjee A. Endophytic bacterial community of rice (*Oryza sativa* L.) from
475 coastal saline zone of West Bengal: 16S rRNA gene based metagenomics approach. *Meta Gene*
476 2018;18: 79-86.
- 477 35. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects
478 models. *Journal of Statistical Software* 2017;82.
- 479 36. Lauro, F. M. *et al.* The genomic basis of trophic strategy in marine bacteria. *Proc. Natl Acad. Sci.*
480 *USA* 106, 15527–15533 (2009)
- 481 37. Liu H, Zhang CL, Yang C, Chen S, Cao Z, Zhang Z et al. Marine Group II dominates planktonic
482 Archaea in water column of the Northeastern South China Sea. *Front Microbiol.* 2017;8:1098
- 483 38. Mahe F, Rognes T, Quince C et al. Swarm v2: highly-scalable and high-resolution amplicon
484 clustering. *PeerJ* 2015;3: e1420.

- 485 39. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing
486 reads. *EMBnet.journal*, 17(1), pp. 10-12. doi:<https://doi.org/10.14806/ej.17.1.200>
- 487 40. Mallick I, Bhattacharyya C, Mukherji S et al. Effective rhizoinoculation and biofilm formation by
488 arsenic immobilizing halophilic plant growth promoting bacteria (PGPB) isolated from mangrove
489 rhizosphere: A step towards arsenic rhizoremediation. *Sci Total Environ* 2018;610-611: 1239-50.
- 490 41. Manivasagan P, Kang KH, Sivakumar K, Li-Chan EC, Oh HM, Kim SK. Marine actinebacteria: an
491 important source of bioactive natural products. *Environ Toxicol Pharmacol* 2014; 38(1): 172-88.
- 492 42. Mishra RR, Swain MR, Dangar TK et al. Diversity and seasonal fluctuation of predominant
493 microbial communities in Bhitarkanika, a tropical mangrove ecosystem in India. *Rev Biol Trop*
494 2012;60: 909-24.
- 495 43. Mitra A, Gangopadhyay A, Dube A et al. Observed changes in water mass properties in the Indian
496 Sundarbans (northwestern Bay of Bengal) during 1980–2007. *Current Science* 2009: 1445-52.
- 497 44. Mizuno CM, Rodriguez-Valera F, Ghai R. Genomes of planktonic acidimicrobiales: widening
498 horizons for marine actinobacteria by metagenomics. *MBio* 2015;6: e02083-14.
- 499 45. Morris RM, Frazar CD, Carlson CA. Basin-scale patterns in the abundance of SAR11 subclades,
500 marine Actinobacteria (OM1), members of the Roseobacter clade and OCS116 in the South
501 Atlantic. *Environmental microbiology* 2012;14: 1133-44.
- 502 46. Mukherjee D, Mukherjee A, Kumar B. Chemical fractionation of metals in freshly deposited marine
503 estuarine sediments of sundarban ecosystem, India. *Environmental geology* 2009;58: 1757-67.
- 504 47. Ngugi DK, Stingl U. High-Quality Draft Single-Cell Genome Sequence of the NS5 Marine Group
505 from the Coastal Red Sea. *Genome announcements* 2018;6: e00565-18.
- 506 48. Nogales, B., Lanfranconi, M.P., Piña-Villalonga, J.M., and Bosch, R. Anthropogenic perturbations
507 in marinemicrobial communities. *FEMS Microbiol Rev* 2011;35:275–298.
- 508 49. Nouioui, I., Carro, L., Garcia-Lopez, M., Meier-Kolthoff, J. P., Woyke, T., Kyrpides, N. C., et al.
509 (2018). Genome-based taxonomic classification of the phylum *Actinobacteria*. *Front.*
510 *Microbiol.* 9:2007. doi: 10.3389/fmicb.2018.02007
- 511 50. Oksanen J, Blanchet FG, Kindt R et al. 2016. *Vegan: Community Ecology Package*.
- 512 51. Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers
513 for marine microbiomes with mock communities, time series and global field samples. *Environ*
514 *Microbiol* 2016;18: 1403-14.
- 515 52. Pruesse E, Peplies J, Glockner FO. SINA: accurate high-throughput multiple sequence alignment of
516 ribosomal RNA genes. *Bioinformatics* 2012;28: 1823-9.
- 517 53. Ribicic D, Netzer R, Hazen TC et al. Microbial community and metagenome dynamics during
518 biodegradation of dispersed oil reveals potential key-players in cold Norwegian seawater. *Marine*
519 *pollution bulletin* 2018;129: 370-8.
- 520 54. Sarkar KS and Bhattacharya DB (2010) Water Quality analysis of the Coastal Region of Sunderban
521 Mangrove Wetland, India, and Using Multivariate Statistical Techniques. *Sciyo, Croatia*, pp 258..
- 522 55. Sarkar SK, Frančišković-Bilinski S, Bhattacharya A et al. Levels of elements in the surficial
523 estuarine sediments of the Hugli River, northeast India and their environmental implications.
524 *Environment international* 2004;30: 1089-98.
- 525 56. Seo J-H, Kang I, Yang S-J, Cho J-C. 2017. Characterization of spatial distribution of the bacterial
526 community in the South Sea of Korea. *PLoS One* 12:e0174159. doi:10.1371/journal.pone.0174159
- 527 57. Simon M, Scheuner C, Meier-Kolthoff JP et al. Phylogenomics of Rhodobacteraceae reveals
528 evolutionary adaptation to marine and non-marine habitats. *The ISME journal* 2017;11: 1483.
- 529 58. Spring S, Scheuner C, Göker M et al. A taxonomic framework for emerging groups of ecologically
530 important marine gammaproteobacteria based on the reconstruction of evolutionary relationships
531 using genome-scale data. *Frontiers in microbiology* 2015;6: 281.
- 532 59. Tanaka D, Tanaka S, Yamashiro Y et al. Distribution of oil-degrading bacteria in coastal seawater,
533 Toyama Bay, Japan. *Environmental Toxicology: An International Journal* 2008;23: 563-9.
- 534 60. Teeling H, Fuchs BM, Becher D et al. Substrate-controlled succession of marine bacterioplankton
535 populations induced by a phytoplankton bloom. *Science* 2012;336: 608-11.

- 536 61. Yilmaz P, Yarza P, Rapp JZ et al. Expanding the world of marine bacterial and archaeal clades.
537 Frontiers in microbiology 2016;6: 1524.
538 62. Zhang J, Kobert K, Flouri T et al. PEAR: a fast and accurate Illumina Paired-End reAd mergeR.
539 Bioinformatics 2014;30: 614-20.
540 63. Zhou J, Richlen ML, Sehein TR et al. Microbial community structure and associations during a
541 marine dinoflagellate bloom. Frontiers in microbiology 2018;9.

542
543
544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562 Table 1: Environmental conditions in Indian Sundarban mangroves and details of the sampling sites.

Station id	Site Id	GPS (DD COORDINATES)	Sample ID	Salinity	pH	DOC (μM)	TN (μM)	NO _x (μM)	NO ₃ (μM)	DIN (μM)	PO ₄ (μM)	Si (μM)
Open Marine Water (OMW)	TH 2	N 21.82389 E 88.50585	TH2.5	26.6	8.2	112.6	12.5	1	1	1	0.4	11.7
			TH2.2	27.0	8.2	111	11.4	1.4	1.4	1.5	0.2	13.1
			TH2.3	25.7	8.3	66.2	6.4	0	0	0	0.1	7.7
	KL 2	N 21.8172 E 88.53658	KL2.3	26.6	8.3	129.5	13.6	2.6	2.5	2.6	0.2	14.3
			KL2.6	26.9	8.1	162.4	10	1.2	1.2	1.3	0.1	26.3
			KL2.1	26.3	8.2	116.3	12.1	1.5	1.4	1.5	0.6	15.9
	BL 1	N 21.78962 E 88.50534	BL1.6	26.8	8.5	121.1	14.8	1	1	1.1	0.2	16.3
			BL1.5	26.9	8.4	114.1	13.2	1.7	1.7	1.8	0.2	12.6
			BL1.2	26.5	8.7	111.4	12.7	2.3	2.2	2.3	0.2	16.1
Island of Sundarban Mangroves (ISM)	KL 1	N 21.85189 E 88.51168	KL1.3	23.6	8.3	56.1	4.4	0.2	0.2	0.3	0.1	6.2
			KL1.6	22.9	8.0	58	6	0.9	0.8	0.9	0.1	6.1
			KL1.2	24.3	8.1	62	5.1	0.2	0.2	0.2	0.1	6.5
	KP 1	N 21.85604 E 88.51191	KP1.6	25.8	8.1	78.5	5.7	0	0	0.1	0.1	12.2
			KP1.1	25.3	8.1	92.5	9	0	0	0.1	0.2	15.4
			KP1.3	24.9	8.1	81.2	7	0.4	0.4	0.5	0.1	12.3
	TH1	N 21.85706 E 88.51638	TH1.4	21.5	8.3	91.9	8.3	0	0	0.1	0.2	8.7
			TH1.3	21.8	8.3	86.8	8.9	0.1	0	0.1	0.1	8.7
			TH1.2	21.2	8.3	84.8	7.4	0.6	0.6	0.6	0.1	9.5

563

564

565 Table 2: Contribution of environmental parameters including nutrient content of six sampling sites to
566 explaining the variation in microbial community composition based on redundancy analysis (RDA).

Sampling station	Explanatory variable	Adjusted R ²	F	df	p-Value
OMW	TN + DIN	6 %	1.24	2, 6	0.022*
	TN	4 %	1.34	1, 7	0.029*
	DIN	-	1.40	1, 7	0.125
ISM	Salinity + TN	7%	1.31	2,6	0.043*
	Salinity	9%	1.77	1,7	0.004**
	TN	0.6%	1.05	1,7	0.326

570

571 p-Values defined as significant at a threshold of 0.05 are highlighted in asterisks mark,
572 Adjusted R² are provided as goodness-of-fit metrics
573 df degrees of freedom (numerator, denominator).
574

575

576 **Figure legends:**

577 Figure 1: Map of the sampling area: water samples were collected from three stations (KL 1, KP
578 1 and TH1) of a small Island are named as ISM and open marine water samples named as OWM
579 (TH 2, KL 2 and BL 1). Three biological replicates from each of the six stations; therefore, total
580 eighteen (18) samples (nine from ISM and another nine from OWM) were collected and further
581 processed for this investigation.

582 Figure 2: Principal component analysis (PCA) to ordinate the eighteen collected water samples
583 collected samples from ISM and OWM based on their environmental parameters. The arrows
584 show the direction of the environmental parameters. DIN, dissolved inorganic nitrogen; TN, total
585 nitrogen; DOC, dissolved organic carbon.

586 Figure 3: Alpha diversity of the water microbial community at two different sites (ISM and
587 OWM) of Sundarban mangrove forest areas. Values are calculated based on repeated random
588 subsampling to the lowest number of sequences per sample. The median per group presented in
589 black line.

590 Figure 4: Taxonomic composition of dominant bacterial taxa on family level across eighteen
591 samples under sites ISM and OWM (nine samples each). Ten (10) most abundant bacterial
592 families for each of the samples were reported here and rests less dominant members are label as
593 “other”.

594 Figure 5: Taxonomic compositions of dominate archaeal phyla across eighteen samples
595 represents two sites ISM and OWM (nine samples each). Ten (10) most abundant phyla for each
596 of the samples were reported here and rests less dominant members are label as “other”

597 Figure 6: Non-metric multidimensional scaling (NMDS) plot of bacterial community
598 composition of the bacterial communities of each sampled at the inhabited island (ISM) and open
599 marine areas (OWM).

600 **Supporting Information:**

601 Figure S1: Heatmap of pairwise correlations between the different environmental parameters and
602 three principal components. Levels of correlations are indicated with different color bar.

603 Figure S2: The rarefaction curve of the eighteen (18) samples, indicated by the number of OTUs
604 as a function of the number of reads. The curve approaching plateau indicates that the number of
605 reads are enough to describe the OTUs representing the community.

606 Figure S3: Taxonomic composition of the ten (10) most abundant bacterial phyla in the studied
607 two sites ISM and OWM represented by eighteen samples (nine samples each).

608 Figure S4: Taxonomic composition of the ten (10) most abundant archaeal genus in the studied
609 two sites ISM and OWM represented by eighteen samples (nine samples each)

610 Table S1: Kruskal-Wallis test for Environmental parameters at the six sampling sites

611 Table S2: Kruskal-Wallis test for alpha diversity of the six sampling sites.

612

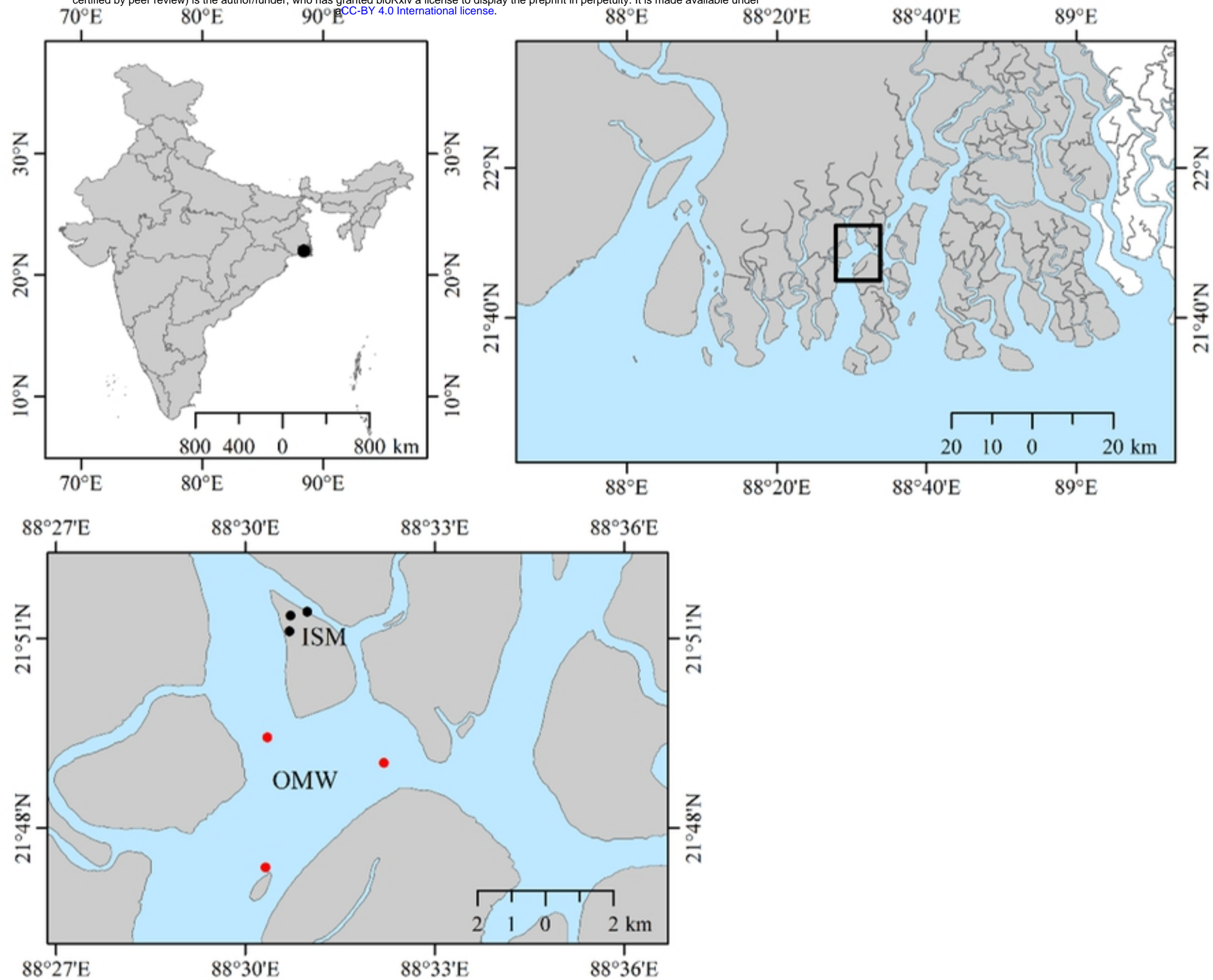


Figure 1

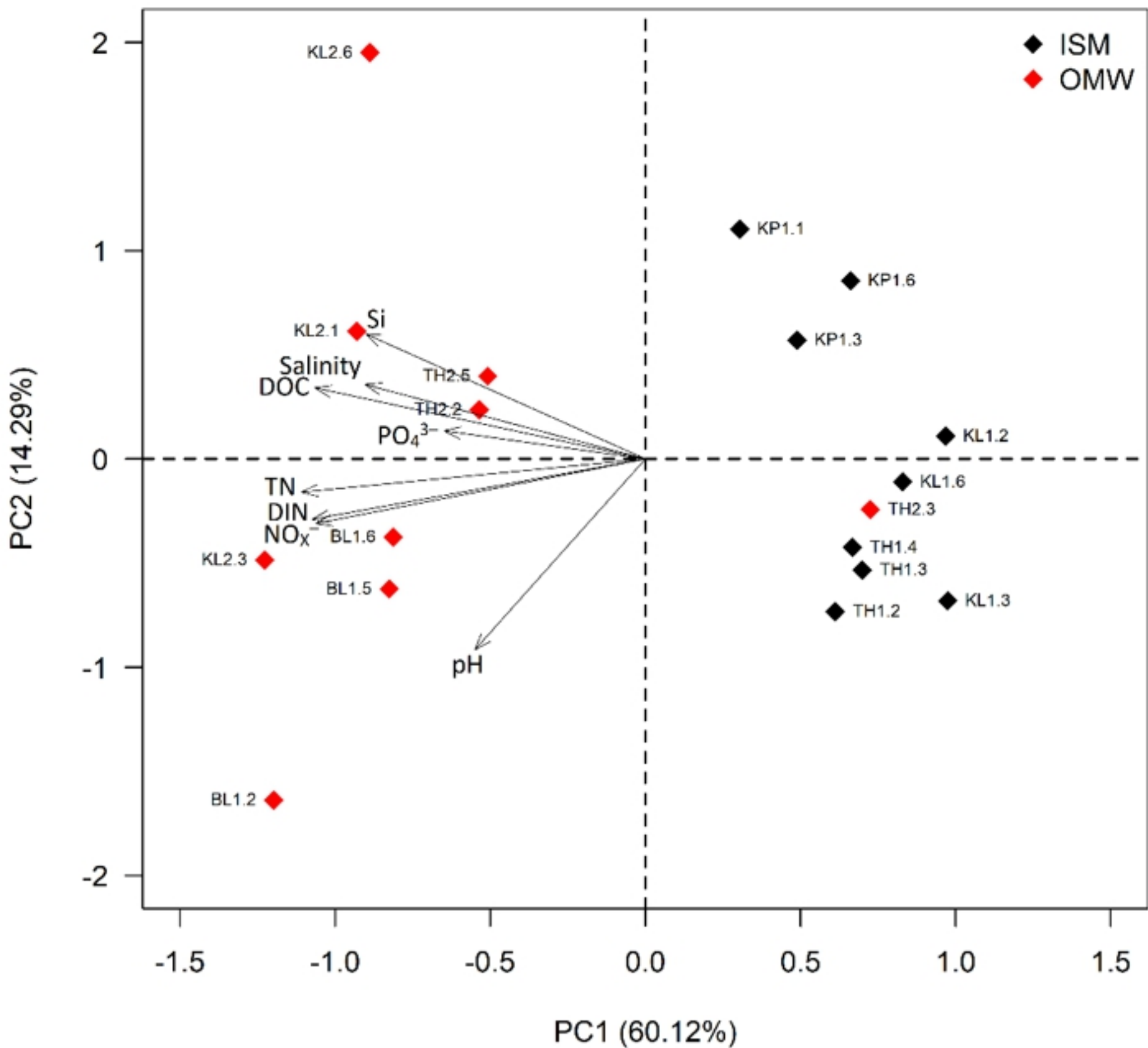
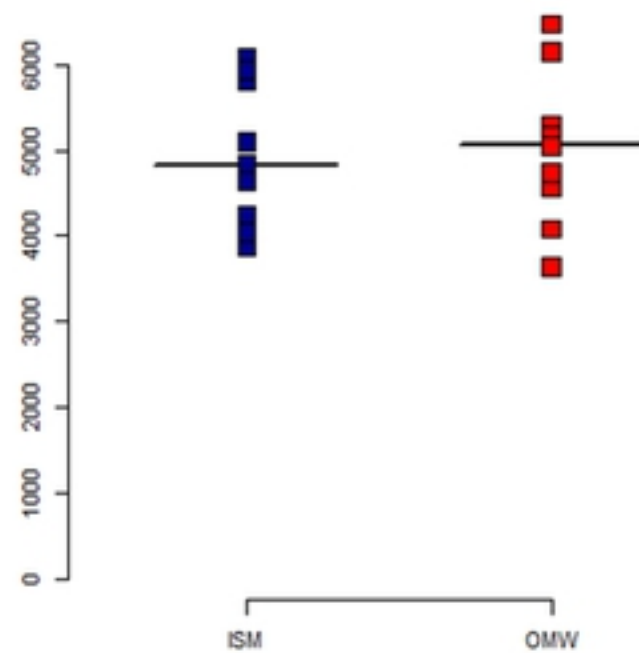
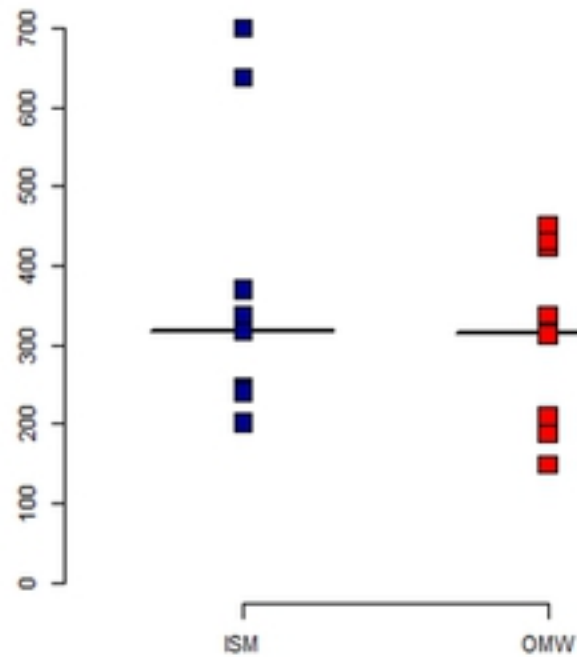


Figure 2

Species richness



The exponential of Shannon entropy



The inverse of Simpson index

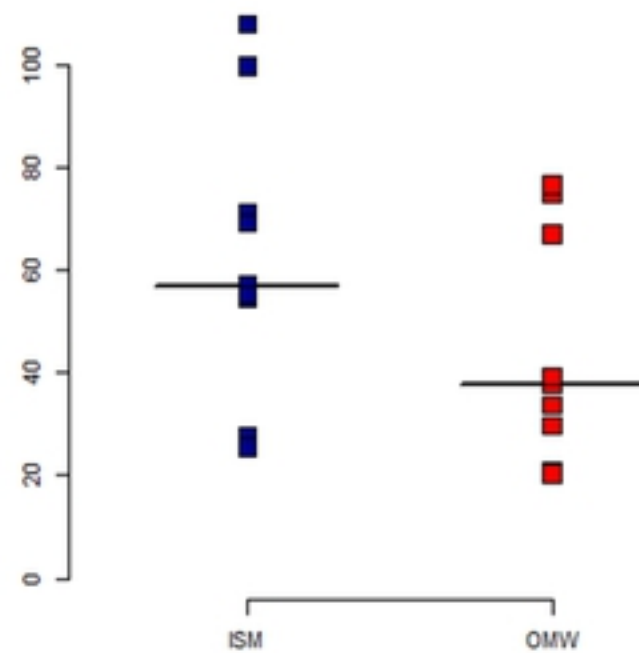


Figure 3

bioRxiv preprint doi: <https://doi.org/10.1101/732842>; this version posted August 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

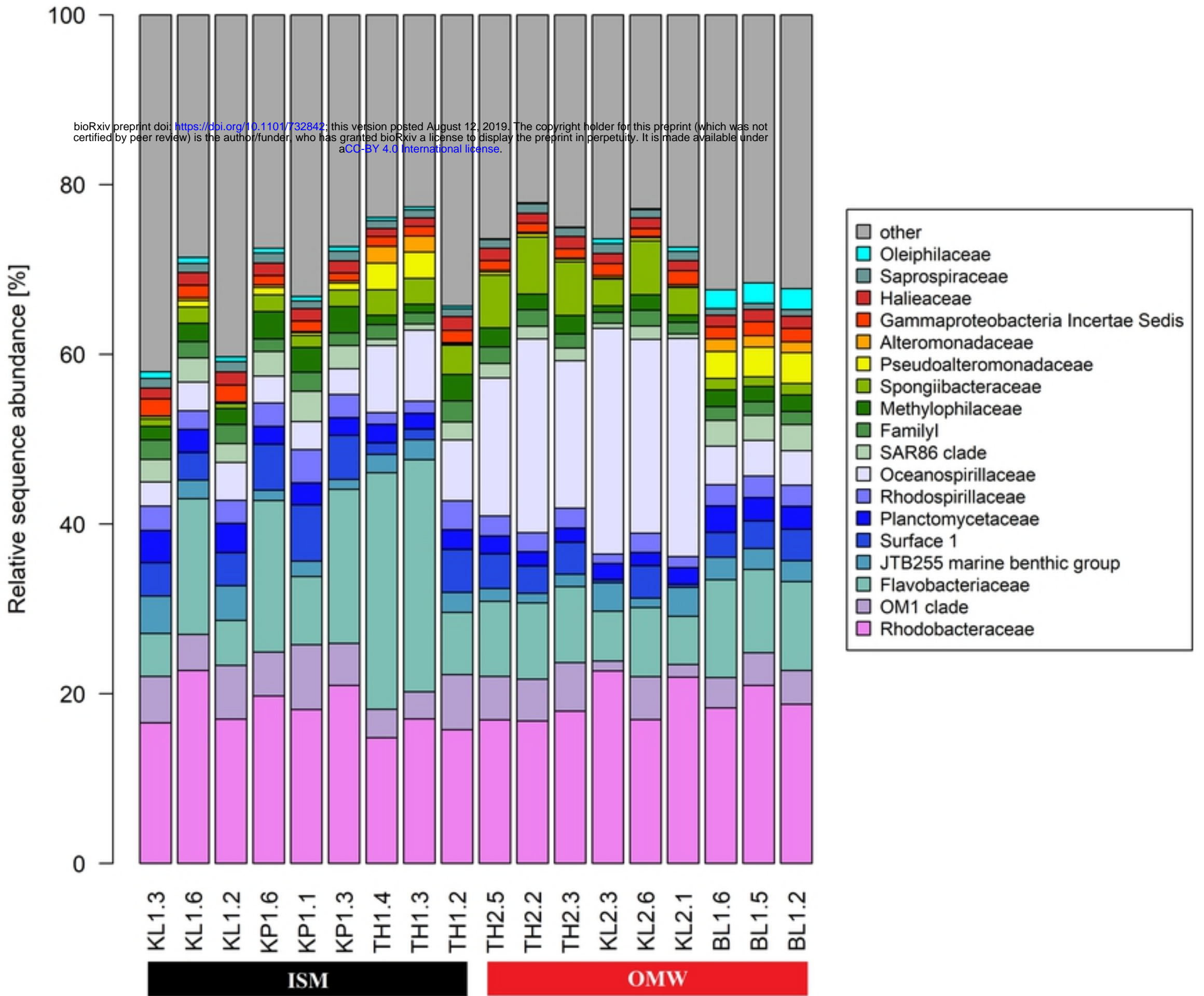


Figure 4

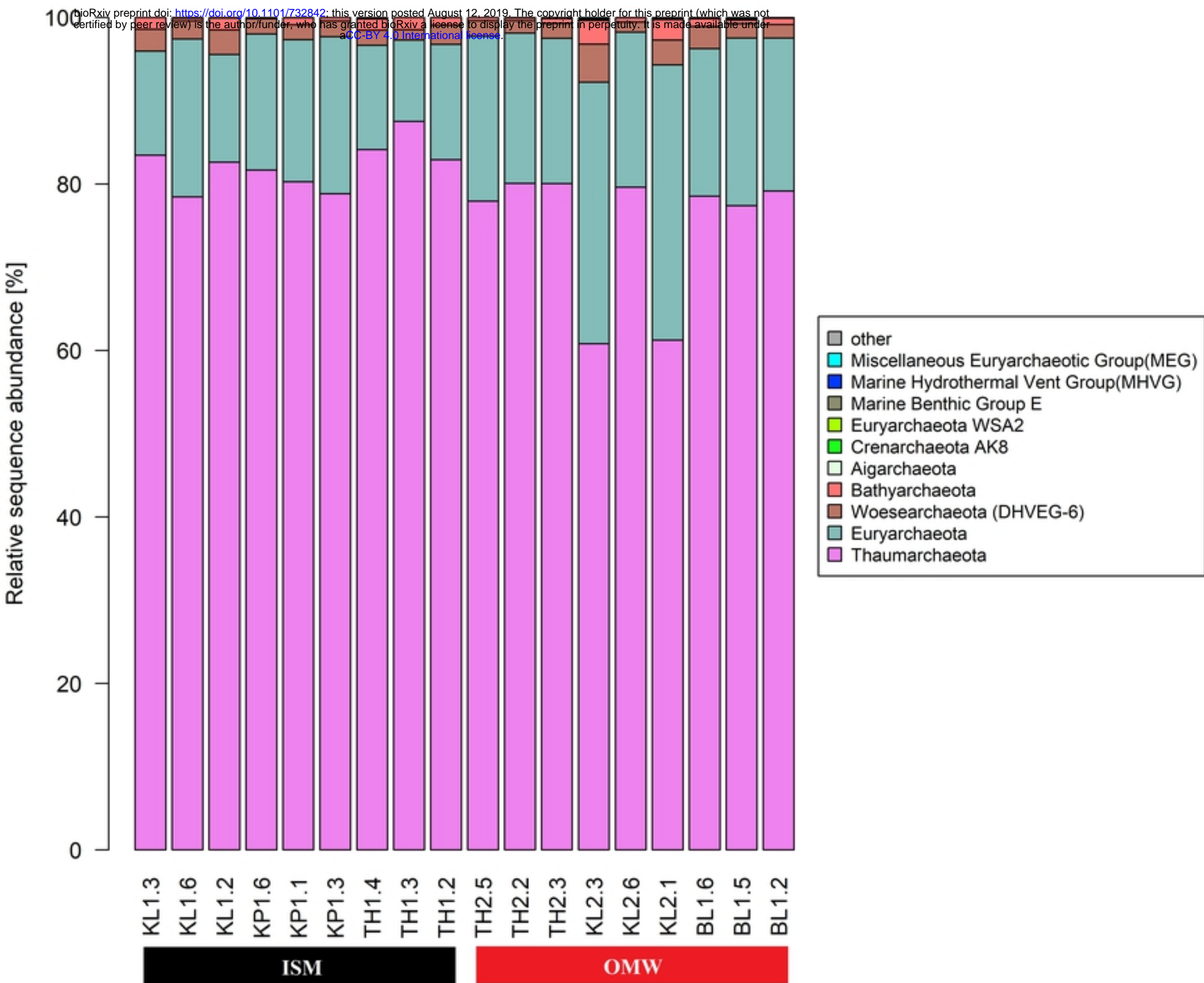


Figure 5

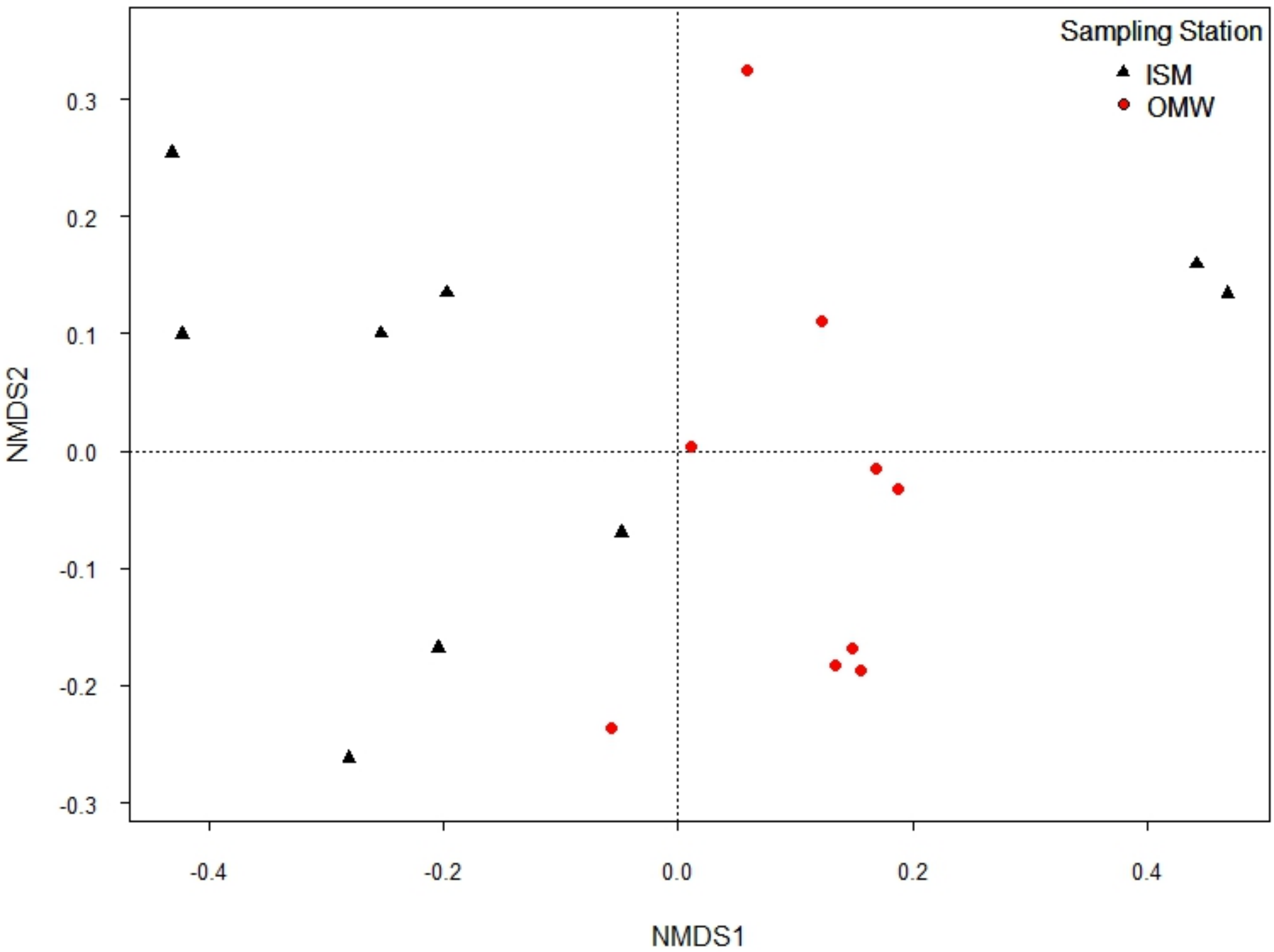


Figure 6