

Selective DNA and Protein Isolation from Marine Macrophyte Surfaces

Marino Korlević^{1*}, Marsej Markovski¹, Zihao Zhao², Gerhard J. Herndl^{2,3}, Mirjana Najdek¹

1. Center for Marine Research, Ruđer Bošković Institute, Croatia

2. Department of Functional and Evolutionary Ecology, University of Vienna, Austria

3. Department of Marine Microbiology and Biogeochemistry, Royal Netherlands Institute for Sea
Research, Utrecht University, The Netherlands

*To whom correspondence should be addressed:

Marino Korlević

G. Paliaga 5, 52210 Rovinj, Croatia

Tel.: +385 52 804 768

Fax: +385 52 804 780

e-mail: marino.korlevic@irb.hr

Running title: DNA and protein isolation from macrophyte surfaces

1 Summary

2 Introduction

3 Surfaces of marine macrophytes are inhabited by a diverse microbial community whose
4 structure and function are poorly understood (Egan *et al.*, 2013). As less than 1 % of all prokaryotic
5 species are culturable, to study these organisms, molecular methods such as 16S rRNA sequencing,
6 metagenomics and metaproteomics are indispensable (Amann *et al.*, 1995; Su *et al.*, 2012). Applying
7 these techniques requires an initial isolation step, with the purpose of obtaining high quality DNA
8 and proteins.

9 Biological material (i.e. proteins and DNA) from pelagic microbial communities is usually
10 isolated by collecting cells onto filters and subsequently isolating the aimed compound (Gilbert
11 *et al.*, 2009). If a microbial size fraction is aimed sequential filtration is applied (Massana *et al.*,
12 1997; Andersson *et al.*, 2010). In contrast, obtaining biological materials from microorganism
13 inhabiting surfaces requires either a cell detachment procedure prior to isolation or the host material
14 is coextracted together with the targeted material. Methods for separating microbial cells form the
15 host include shaking of host tissue (Gross *et al.*, 2003; Nõges *et al.*, 2010), scraping of macrophyte
16 surfaces (Uku *et al.*, 2007) or the application of ultrasonication (Weidner *et al.*, 1996; Cai *et al.*,
17 2014). It was shown that shaking alone is not sufficient to remove microbial cells, at least from
18 plant root surfaces (Richter-Heitmann *et al.*, 2016). Manual separation methods, such as scraping
19 and brushing, are time consuming and subjective, as the detachment efficiency depends on host
20 tissue and the person performing the procedure (Cai *et al.*, 2014). Ultrasonication was proposed as
21 an alternative method as it is providing better results in terms of detachment efficiency (Cai *et al.*,
22 2014; Richter-Heitmann *et al.*, 2016). The downside of this procedure is that complete cell removal
23 was still not obtained and tissue disruption was observed especially after the application of probe
24 ultrasonication (Richter-Heitmann *et al.*, 2016). An alternative to these cell detachment procedures
25 is the isolation of targeted epiphytic compounds together with host materials (Staufenberger *et*
26 *al.*, 2008; Jiang *et al.*, 2015). This procedure can lead to problems in the following processing
27 steps such as mitochondrial and chloroplast 16S rRNA sequence contaminations from the host

28 (Longford *et al.*, 2007; Staufenberger *et al.*, 2008). In addition, when performing metagenomics
29 and metaproteomics host material can cause biased results towards more abundant host DNA and
30 proteins.

31 An alternative to these procedures is a direct isolation of the targeted material by incubating
32 macrophyte tissues in an extraction buffer. After the incubation is done the undisrupted host tissue
33 is removed and the isolation procedure continues omitting host material contaminations. To our
34 knowledge the only procedure describing a direct and selective epiphytic DNA isolation from the
35 surfaces of marine macrophytes was provided by Burke *et al.* (2009). In contrast to previously
36 described methods this protocol enables an almost complete removal of the surface community
37 and was used for 16S rRNA gene clone libraries construction (Burke and Thomas *et al.*, 2011) and
38 metagenomes sequencing (Burke and Peter Steinberg *et al.*, 2011). This method, thought providing
39 a selective isolation procedure, is using in the extraction buffer a rapid multienzyme cleaner (3M)
40 that is not worldwide available and whose composition is not known (Burke *et al.*, 2009). Also to
41 our knowledge, no selective isolation protocol for proteins from epiphytic communities inhabiting
42 marine macrophytes was established.

43 In the present study, we adapted a protocol widely used for DNA isolation from filters (Massana
44 *et al.*, 1997) and a protocol used for protein isolation from sediments (Chourey *et al.*, 2010; Hultman
45 *et al.*, 2015) for selective extractions of DNA and proteins from epiphytic communities inhabiting
46 the surfaces of two marine macrophytes: the seagrass *Cymodocea nodosa* and the alga *Caulerpa*
47 *cylindracea*. In addition, we tested the removal efficiency of the protocol and the suitability of
48 obtained DNA and proteins for 16S rRNA sequencing and metaproteomics.

49 **Results**

50 To assess the removal efficiency of DNA and protein isolation procedures leaves and thalli were
51 stained with SYBR Green I and examined under a confocal microscope before and after treatments
52 were performed. As it is known that marine macrophytes are harboring more algal epiphytes
53 during Autumn and Winter months (Reyes and Sansón, 2001), the sampling was performed in two
54 contrasting seasons, in December 2017 when higher abundances of algal epiphytes are expected,
55 and in June 2018, when lower abundances could be expected. In addition, to further test the isolation
56 protocols on different macrophyte species, DNA and proteins were isolated from the surfaces of two
57 macrophytes, the seagrass *C. nodosa* and the macroalga *C. cylindracea*. *C. nodosa* was sampled in a
58 monospecific meadow (Bay of Saline) and *C. cylindracea* invaded meadow (Bay of Funtana), while
59 samples of *C. cylindracea* originated from a monospecific settlement and a *C. cylindracea* invaded
60 *C. nodosa* meadow in the Bay of Funtana. *C. nodosa* June samples from the Bay of Saline were
61 missing as the meadow went to a decline from Spring 2018 (M. Najdek, personal communication).
62 Procedures developed for DNA and protein isolation showed an almost complete removal of the
63 surface community on both, samples collected in December 2017 (Fig. 1) and June 2018 (Fig. 2).
64 In addition, a similar removal efficiency was observed for communities associated with *C. nodosa*
65 and *C. cylindracea*. Also, no effect of station, settlement or isolation procedure (DNA or protein)
66 on the removal efficiency was observed (Figs. 1 and 2).

67 To evaluate if the obtained DNA is suitable to determine the microbial community structure
68 an Illumina sequencing of the V4 16S rRNA region was performed. Sequencing yielded a total
69 of 336,937 sequences after quality curation and exclusion of eukaryotic, mitochondrial and no
70 relative sequences. The number of sequences classified as chloroplast was 97,328 (28.9 %) After
71 the exclusion of these sequences the total number of retrieved reads dropped to 239,609, ranging
72 from 22,587 and 52,958 sequences per sample (Table S1). Even when the highest sequencing effort
73 was applied the rarefaction curves did not level off that is commonly observed in high-throughput
74 16S rRNA amplicon sequencing procedures (Fig. S1). Sequence clustering at a similarity level

75 of 97 % yielded a total of 8,360 different OTUs. Taxonomic classification of reads allowed the
76 determination of the macrophyte associated epiphytic community that was mainly composed of:
77 *Alphaproteobacteria* ($14.9 \pm 3.5\%$), *Bacteroidota* ($12.5 \pm 2.4\%$), *Gammaproteobacteria* ($11.6 \pm$
78 4.3%), *Desulfobacterota* ($7.8 \pm 7.4\%$), *Cyanobacteria* ($6.5 \pm 4.7\%$) and *Planctomycetota* ($2.9 \pm$
79 1.7%) (Fig. 3).

80 Primers used in this study, as in many other 16S rRNA amplicon procedures, also amplify
81 chloroplast 16S rRNA genes. We observed a high proportion of chloroplast sequences in all analyzed
82 samples ($33.4 \pm 9.4\%$) (Fig. 3). To determine if chloroplast sequences originate from hosts or
83 eukaryotic epiphytic organisms, we exported SILVA classified chloroplast sequences and reclassified
84 them using the RDP (Ribosomal Database Project) training set that allows for a more detailed
85 chloroplast classification. The largest proportion of sequences was classified as *Bacillariophyta*
86 ($89.7 \pm 5.7\%$) indicating that the DNA removal procedure did not coextract larger quantities of
87 host DNA (Fig. 4).

88 **Discussion**

89 **Experimental procedures**

90 **Sampling**

91 Leaves of *C. nodosa* were sampled in a *C. nodosa* meadow in the Bay of Saline ($45^{\circ}7'5''$ N,
92 $13^{\circ}37'20''$ E) and in a *C. nodosa* meadow invaded by *C. cylindracea* in the proximity of the village
93 of Funtana ($45^{\circ}10'39''$ N, $13^{\circ}35'42''$ E). Thalli of *C. cylindracea* were sampled in the same *C.*
94 *nodosa* invaded meadow in Funtana and on a locality of only *C. cylindracea* located in the proximity
95 of the invaded meadow. Leaves and thalli were collected on the same day in two contrasting seasons,
96 on 4 December 2017 and 18 June 2018. During spring 2018 the *C. nodosa* meadow in the Bay
97 of Saline decayed to an extent that no leaves could be retrieved (Najdek *et al.*, unpublished data).
98 Leaves and thalli were collected by diving and transported to the laboratory in containers placed
99 on ice and filled with site seawater. Upon arrival to the laboratory, *C. nodosa* leaves were cut into
100 sections of 1 – 2 cm, while *C. cylindracea* thalli were cut into 5 – 8 cm long sections. Leaves and
101 thalli were washed three times with sterile artificial seawater (ASW) to remove loosely attached
102 microbial cells.

103 **DNA isolation**

104 The DNA was isolated according to the protocol for isolation from filters described in
105 Massana *et al.* (1997). This protocol was modified and adapted for DNA isolation from microbial
106 communities from macrophyte surfaces as described below. 5 ml of lysis buffer (40 mM EDTA,
107 50 mM Tris-HCl, 0.75 M sucrose; pH 8.3) was added to 1 g wet weight of leaves or g wet-weight
108 of thalli. Lysozyme was added (final concentration 1 mg ml^{-1}) and the mixture was incubated at
109 37°C for 30 min. Subsequently, proteinase K (final concentration 0.5 mg ml^{-1}) and SDS (final
110 concentration 1 %) were added and the samples were incubated at 55°C for 2 h. Following the
111 incubation, tubes were vortexed for 10 min and the mixture containing lysed epiphytic cells was

112 separated from host leaves or thalli by transferring the solution into a clean tube. The lysate was
113 extracted twice with a mixture of phenol:chloroform:isoamyl alcohol (25:24:1; pH 8) and once
114 with chloroform:isoamyl alcohol (24:1). After each organic solvent mixture addition tubes were
115 slightly vortexed and centrifuged at $4,500 \times g$ for 10 min. Following each centrifugation aqueous
116 phases were retrieved. After the final extraction 1/10 of chilled 3 M sodium acetate (pH 5.2) was
117 added. DNA was precipitated by adding 1 volume of chilled isopropanol, incubating the mixtures
118 overnight at -20°C and centrifuging at $16,000 \times g$ and 4°C for 20 min. The pellet was washed
119 twice with 1 ml of chilled 70 % ethanol and centrifuged after each washing step at $20,000 \times g$ and 4°C
120 for 10 min. After the first washing step duplicate pellets from the same sample were pooled and
121 transferred to a clean 1.5 ml tube. The dried pellet was resuspended in 100 μl of deionized water.

122 Illumina 16S rRNA sequencing

123 An aliquot of isolated DNA was treated with RNase A (final concentration $200 \mu\text{g ml}^{-1}$) for 2
124 h at 37°C . The DNA concentration was determined using the Quant-iT PicoGreen dsDNA Assay
125 Kit (Invitrogen, USA) according to the manufacturer's instructions and diluted to $1 \text{ ng } \mu\text{l}^{-1}$. The V4
126 region of the 16S rRNA gene was amplified using a two-step PCR procedure. In the first PCR the
127 515F (5'-GTGYCAGCMGCCGCGTAA-3') and 806R (5'-GGACTACNVGGGTWTCTAAT-3')
128 primers from the Earth Microbiome Project (<http://press.igsb.anl.gov/earthmicrobiome/protocols-and-standards/16s/>) were used to amplify the target region (Caporaso *et al.*, 2012; Apprill *et al.*,
129 2015; Parada *et al.*, 2016). These primers contained on their 5' end a tagged sequence. Each sample
130 was amplified in four parallel 25 μl reactions of which each contained: $1 \times$ Q5 Reaction Buffer,
131 0.2 mM of dNTPmix, 0.7 mg ml^{-1} BSA (Bovine Serum Albumin), 0.2 μM of forward and reverse
132 primers, 0.5 U of Q5 High-Fidelity DNA Polymerase (New England Biolabs, USA) and 5 ng of
133 DNA template. Cycling conditions were: initial denaturation at 94°C for 3 min, 20 cycles of
134 denaturation at 94°C for 45 s, annealing at 50°C for 60 s and elongation at 72°C for 90 s, finalized
135 by an elongation step at 72°C for 10 min. The four parallel reactions volumes were pooled and PCR
136 by an elongation step at 72°C for 10 min. The four parallel reactions volumes were pooled and PCR

137 products were purified using the GeneJET PCR Purification Kit (Thermo Fisher Scientific, USA)
138 according to the manufacturer's instructions and following the protocol that included isopropanol
139 addition for better small DNA fragment yield. The column was eluted in 30 µl of deionized water.
140 Purified PCR products were sent for Illumina MiSeq sequencing (2×250 bp) at IMGM Laboratories,
141 Martinsried, Germany. Before sequencing, at IMGM the second PCR amplification of the two-step
142 PCR procedure was performed using primers targeting the tagged region incorporated in the first
143 PCR. In addition, these primers contained adapter and sample-specific index sequences. The second
144 PCR was carried out for 8 cycles. Beside samples, a positive and negative control were sequenced.
145 A negative control was comprised of four parallel PCR reactions without DNA template, while for a
146 positive control a mock community composed of evenly mixed DNA material originating from 20
147 bacterial strains (ATCC MSA-1002, ATCC, USA) was used. Sequences obtained in this study have
148 been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession numbers
149 SAMEA6786270, SAMEA6648792 – SAMEA6648794, SAMEA6648809 – SAMEA6648811.

150 Obtained sequences were analyzed on the computer cluster Isabella (University Computing
151 Center, University of Zagreb) using mothur (version 1.43.0) (Schloss *et al.*, 2009) according to the
152 MiSeq Standard Operating Procedure (MiSeq SOP; https://mothur.org/wiki/MiSeq_SOP) (Kozich *et*
153 *al.*, 2013) and recommendations given from the Riffomonas project to enhance data reproducibility
154 (<http://www.riffomonas.org/>). For alignment and classification of sequences the SILVA SSU Ref
155 NR 99 database (release 138; <http://www.arb-silva.de>) was used (Quast *et al.*, 2013; Yilmaz *et*
156 *al.*, 2014). Sequences classified as chloroplasts by SILVA were exported and reclassified using
157 mothur and the RDP (<http://rdp.cme.msu.edu/>) training set (version 16) reference files adapted for
158 mothur (Cole *et al.*, 2014). In comparison to SILVA, RDP allows a more detailed classification of
159 chloroplast sequences. Pipeline data processing and visualization was done using R (version 3.6.0)
160 (R Core Team, 2019), package tidyverse (version 1.3.0) (Wickham *et al.*, 2019) and multiple other
161 packages (Xie, 2014, 2015, 2020; Neuwirth, 2014; Xie *et al.*, 2018; Y. Xie, 2019b, 2019a; Allaire *et*
162 *al.*, 2019; Zhu, 2019; Bengtsson, 2020). The detailed analysis procedure including the R Markdown
163 file for this paper are available as a GitHub repository (https://github.com/mkorlevic/Korlevic_

¹⁶⁴ SelectiveRemoval_EnvironMicrobiol_2020). Based on the ATCC MSA-1002 mock community
¹⁶⁵ included in the analysis a sequencing error rate of 0.009 % was determined, which is in line with
¹⁶⁶ previously reported values for next-generation sequencing data (Kozich *et al.*, 2013; Schloss *et*
¹⁶⁷ *al.*, 2016). In addition, the negative control processed together with the samples yielded only 2
¹⁶⁸ sequences after sequence quality curation.

¹⁶⁹ Metagenomics

¹⁷⁰ Four samples were selected and sent on dry ice to IMGM Laboratories, Martinsried, Germany
¹⁷¹ for metagenomic sequencing. DNA was purified using AMPure XP Beads (Beckman Coulter, USA)
¹⁷² applying a bead:DNA ratio of 1:1 (v/v), quantified with a Qubit dsDNA HS Assay Kit (Thermo
¹⁷³ Fisher Scientific, USA) and check for integrity on a 1 % agarose gel. Metagenomic sequencing
¹⁷⁴ libraries were generated from 300 ng of input DNA using a NEBNext Ultra II FS DNA Library
¹⁷⁵ Prep Kit for Illumina (New England Biolabs, USA) according to the manufacturer's instructions.
¹⁷⁶ Fragments were selected (500 – 700 bp) using AMPure XP Beads, PCR enriched for 3 – 5 cycles
¹⁷⁷ and quality controlled. Libraries generated from different DNA samples were pooled and sequenced
¹⁷⁸ on an Illumina NovaSeq 6000 sequencing system (2 × 250 bp).

¹⁷⁹ Obtained sequences were assembled individually using MEGAHIT (version 1.1.2) (Li *et al.*,
¹⁸⁰ 2015) under default settings. Putative genes were predicted from contings longer than 200 bp
¹⁸¹ using Prodigal (version 2.6.3) (Hyatt *et al.*, 2010) in metagenome mode (-p meta). Abundances
¹⁸² of predicted genes were expressed as Reads Per Kilobase Million (RPKM) and calculated using
¹⁸³ the BWA algorithm (version 0.7.16a) (Li and Durbin, 2009). All predicted genes were functionally
¹⁸⁴ annotated using the eggNOG-mapper (Huerta-Cepas *et al.*, 2017) and eggNOG database (version
¹⁸⁵ 5.0) (Huerta-Cepas *et al.*, 2019). The taxonomic classification of sequences was determined using
¹⁸⁶ the lowest common ancestor algorithm adapted from DIAMOND (version 0.8.36) (Buchfink *et al.*,
¹⁸⁷ 2015) and by searching against the non-redundant database (NR). To determine phylogeny the top
¹⁸⁸ 10 % hits with an e-value < 1×10^{-5} were used (--top 10).

189 **Protein isolation**

190 Proteins were isolated according to the protocol for isolation from soil described in Chourey
191 *et al.* (2010) and modified by Hultman *et al.* (2015). These protocols were further modified and
192 adapted for protein isolation from microbial communities form macrophyte surfaces as described
193 below. 20 ml of protein extraction buffer (4 % SDS, 100 mM Tris-HCl [pH 8.0]) was added to 5 g
194 wet weight of leaves or 10 g wet weight of thalli. The mixture was incubated in boiling water for 5
195 min, vortexed for 10 min and incubated again in boiling water for 5 min. After a brief vortex the
196 lysate was transferred to a clean tube separating the host leaves or thalli from the mixture containing
197 lyzed epiphytic cells. Dithiothreitol (DTT; final concentration 24 mM) was added and proteins were
198 precipitated with chilled 100 % trichloroacetic acid (TCA; final concentration 20 %) overnight at
199 –20 °C. Precipitated proteins were centrifuged at 10,000 × g and 4 °C for 40 min. The obtained
200 protein pellet was washed three times with chilled acetone. During the first washing step the pellet
201 was transferred to a clean 1.5 ml tube. After each washing step samples were centrifuged at 20,000
202 × g and 4 °C for 5 min. Dried pellets were stored at –80 °C until further analysis.

203 **Metaproteomics**

204 Isolated proteins were whole trypsin digested using the FASP (filter-aided sample preparation)
205 Protein Digestion Kit (Expedeon, UK) according to the manufacturer's instructions (Wiśniewski
206 *et al.*, 2009) with small modifications. Before loading the solution to the column, protein pellets
207 were solubilized in a urea sample buffer included in the kit amended with DTT (final concentration
208 100 mM) for 45 min at room temperature and centrifuged at 20,000 × g for 2 – 5 min at room
209 temperature to remove larger particles. The first washing step after protein solution loading was
210 repeated twice. In addition, centrifugation steps were prolonged if the column was clogged.
211 Trypsin digestion was performed on column filters at 37 °C overnight for 18 h. The final filtrate
212 containing digested proteins was acidified with 1 % (final concentration) trifluoroacetic acid,

213 freezed at –80 °C for 15 min, lyophilized and sent to the Vienna Metabolomics Center (University
214 of Vienna) for metaproteomic analysis. Peptides were resuspended in 1 % (final concentration)
215 trifluoroacetic acid, desalted using the Pierce C18 Tips (Thermo Fisher Scientific, USA) according
216 to the manufacturer's instructions and sequenced on a Q Exactive Hybrid Quadrupole-Orbitrap
217 Mass Spectrometer (Thermo Fisher Scientific, USA). Obtained MS/MS spectra were searched
218 against a protein database from metagenomic assembly published in Burke and Peter Steinberg *et*
219 *al.* (2011) using SEQUEST-HT engines and validated with Percolator in Proteome Discoverer 2.1
220 (Thermo Fisher Scientific, USA). The target-decoy approach was used to reduce the probability of
221 false peptide identification. Results whose false discovery rate at the peptide level was <1 % were
222 kept. For protein identification a minimum of two peptides and one unique peptide were required.
223 For protein quantification, a chromatographic peak area-based free quantitative method was applied.

224 **Confocal microscopy**

225 Host leaves and thalli from DNA and protein isolation steps were washed seven times in
226 deionized water and fixed with formaldehyde (final concentration ~ 3 %). In addition, nontreated
227 leaves and thalli, washed three times in ASW to remove loosely attached microbial cells, were fixed
228 in the same concentration of formaldehyde and used as a positive control. For long therm storage,
229 fixed leaves and thalli were immersed in a mixture of phosphate-buffered saline (PBS) and ethanol
230 (1:1) and stored at –20 °C. Treated and untreated leaves and thalli segments were stained in a 2
231 × solution of SYBR Green I and examined under a Leica TCS SP8 X FLIM confocal microscope
232 (Leica Microsystems, Germany).

233 **Acknowledgements**

234 **References**

- 235 Allaire, J.J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., et al. (2019)
- 236 rmarkdown: Dynamic documents for R.
- 237 Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and *in situ*
238 detection of individual microbial cells without cultivation. *Microbiol Rev* **59**: 143–169.
- 239 Andersson, A.F., Riemann, L., and Bertilsson, S. (2010) Pyrosequencing reveals contrasting
240 seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *ISME J* **4**: 171–181.
- 241 Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015) Minor revision to V4 region SSU
242 rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat Microb*
243 *Ecol* **75**: 129–137.
- 244 Bengtsson, H. (2020) MatrixStats: Functions that apply to rows and columns of matrices (and
245 to vectors).
- 246 Buchfink, B., Xie, C., and Huson, D.H. (2015) Fast and sensitive protein alignment using
247 DIAMOND. *Nat Methods* **12**: 59–60.
- 248 Burke, C., Kjelleberg, S., and Thomas, T. (2009) Selective extraction of bacterial DNA from
249 the surfaces of macroalgae. *Appl Environ Microbiol* **75**: 252–256.
- 250 Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011) Bacterial
251 community assembly based on functional genes rather than species. *Proc Natl Acad Sci U S A* **108**:
252 14288–14293.
- 253 Burke, C., Thomas, T., Lewis, M., Steinberg, P., and Kjelleberg, S. (2011) Composition,
254 uniqueness and variability of the epiphytic bacterial community of the green alga *Ulva australis*.
255 *ISME J* **5**: 590–600.

- 256 Cai, X., Gao, G., Yang, J., Tang, X., Dai, J., Chen, D., and Song, Y. (2014) An ultrasonic
257 method for separation of epiphytic microbes from freshwater submerged macrophytes. *J Basic*
258 *Microbiol* **54**: 758–761.
- 259 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., et al.
260 (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq
261 platforms. *ISME J* **6**: 1621–1624.
- 262 Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., et al.
263 (2010) Direct cellular lysis/protein extraction protocol for soil metaproteomics. *J Proteome Res* **9**:
264 6615–6622.
- 265 Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., et al. (2014) Ribosomal
266 Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res* **42**:
267 D633–D642.
- 268 Egan, S., Harder, T., Burke, C., Steinberg, P., Kjelleberg, S., and Thomas, T. (2013) The
269 seaweed holobiont: Understanding seaweed-bacteria interactions. *FEMS Microbiol Rev* **37**:
270 462–476.
- 271 Gilbert, J.A., Field, D., Swift, P., Newbold, L., Oliver, A., Smyth, T., et al. (2009) The
272 seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol* **11**:
273 3132–3139.
- 274 Gross, E.M., Feldbaum, C., and Graf, A. (2003) Epiphyte biomass and elemental composition
275 on submersed macrophytes in shallow eutrophic lakes. *Hydrobiologia* **506-509**: 559–565.
- 276 Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C.,
277 and Bork, P. (2017) Fast genome-wide functional annotation through orthology assignment by
278 eggNOG-mapper. *Mol Biol Evol* **34**: 2115–2122.

279 Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H.,
280 et al. (2019) eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology
281 resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**: D309–D314.

282 Hultman, J., Waldrop, M.P., Mackelprang, R., David, M.M., McFarland, J., Blazewicz, S.J., et
283 al. (2015) Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature*
284 **521**: 208–212.

285 Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010)
286 Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC*
287 *Bioinformatics* **11**: 119.

288 Jiang, Y.-F., Ling, J., Dong, J.-D., Chen, B., Zhang, Y.-Y., Zhang, Y.-Z., and Wang, Y.-S. (2015)
289 Illumina-based analysis the microbial diversity associated with *Thalassia hemprichii* in Xincun Bay,
290 South China Sea. *Ecotoxicology* **24**: 1548–56.

291 Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013)
292 Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon
293 sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* **79**: 5112–5120.

294 Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015) MEGAHIT: An ultra-fast
295 single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph.
296 *Bioinformatics* **31**: 1674–1676.

297 Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with BurrowsWheeler
298 transform. *Bioinformatics* **25**: 1754–1760.

299 Longford, S., Tujula, N., Crocetti, G., Holmes, A., Holmström, C., Kjelleberg, S., et al. (2007)
300 Comparisons of diversity of bacterial communities associated with three sessile marine eukaryotes.
301 *Aquat Microb Ecol* **48**: 217–229.

- 302 Massana, R., Murray, A.E., Preston, C.M., and DeLong, E.F. (1997) Vertical distribution and
303 phylogenetic characterization of marine planktonic *Archaea* in the Santa Barbara Channel. *Appl*
304 *Environ Microbiol* **63**: 50–56.
- 305 Neuwirth, E. (2014) RColorBrewer: ColorBrewer palettes.
- 306 Nõges, T., Luup, H., and Feldmann, T. (2010) Primary production of aquatic macrophytes and
307 their epiphytes in two shallow lakes (Peipsi and Võrtsjärv) in Estonia. *Aquat Ecol* **44**: 83–92.
- 308 Parada, A.E., Needham, D.M., and Fuhrman, J.A. (2016) Every base matters: Assessing small
309 subunit rRNA primers for marine microbiomes with mock communities, time series and global field
310 samples. *Environ Microbiol* **18**: 1403–1414.
- 311 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA
312 ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic*
313 *Acids Res* **41**: D590–D596.
- 314 R Core Team (2019) R: A language and environment for statistical computing, Vienna, Austria:
315 R Foundation for Statistical Computing.
- 316 Reyes, J. and Sansón, M. (2001) Biomass and production of the epiphytes on the leaves of
317 *Cymodocea nodosa* in the Canary Islands. *Bot Mar* **44**: 307–313.
- 318 Richter-Heitmann, T., Eickhorst, T., Knauth, S., Friedrich, M.W., and Schmidt, H. (2016)
319 Evaluation of strategies to separate root-associated microbial communities: A crucial choice in
320 rhizobiome research. *Front Microbiol* **7**: 773.
- 321 Schloss, P.D., Jenior, M.L., Koumpouras, C.C., Westcott, S.L., and Highlander, S.K. (2016)
322 Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. *PeerJ* **4**:
323 e1869.
- 324 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al.

- 325 (2009) Introducing mothur: Open-source, platform-independent, community-supported software for
326 describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.
- 327 Staufenberger, T., Thiel, V., Wiese, J., and Imhoff, J.F. (2008) Phylogenetic analysis of bacteria
328 associated with *Laminaria saccharina*. *FEMS Microbiol Ecol* **64**: 65–77.
- 329 Su, C., Lei, L., Duan, Y., Zhang, K.-Q., and Yang, J. (2012) Culture-independent methods for
330 studying environmental microorganisms: Methods, application, and perspective. *Appl Microbiol*
331 *Biotechnol* **93**: 993–1003.
- 332 Uku, J., Björk, M., Bergman, B., and Díez, B. (2007) Characterization and comparison of
333 prokaryotic epiphytes associated with three East African seagrasses. *J Phycol* **43**: 768–779.
- 334 Weidner, S., Arnold, W., and Puhler, A. (1996) Diversity of uncultured microorganisms
335 associated with the seagrass *Halophila stipulacea* estimated by restriction fragment length
336 polymorphism analysis of PCR-amplified 16S rRNA genes. *Appl Environ Microbiol* **62**: 766–771.
- 337 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., et al. (2019)
338 Welcome to the tidyverse. *J Open Source Softw* **4**: 1686.
- 339 Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample
340 preparation method for proteome analysis. *Nat Methods* **6**: 359–362.
- 341 Xie, Y. (2015) Dynamic Documents with R and knitr, 2nd ed. Boca Raton, Florida: Chapman
342 and Hall/CRC.
- 343 Xie, Y. (2014) knitr: A comprehensive tool for reproducible research in R. In *Implementing*
344 *Reproducible Computational Research*. Stodden, V., Leisch, F., and Peng, R.D. (eds). New York:
345 Chapman and Hall/CRC, pp. 3–32.
- 346 Xie, Y. (2019a) knitr: A general-purpose package for dynamic report generation in R.

³⁴⁷ Xie, Y. (2019b) TinyTeX: A lightweight, cross-platform, and easy-to-maintain LaTeX
³⁴⁸ distribution based on TeX Live. *TUGboat* **40**: 30–32.

³⁴⁹ Xie, Y. (2020) tinytex: Helper functions to install and maintain 'TeX Live', and compile
³⁵⁰ 'LaTeX' documents.

³⁵¹ Xie, Y., Allaire, J.J., and Grolemund, G. (2018) R Markdown: The Definitive Guide, 1st ed.
³⁵² Boca Raton, Florida: Chapman and Hall/CRC.

³⁵³ Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014) The
³⁵⁴ SILVA and "All-Species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res* **42**:
³⁵⁵ D643–D648.

³⁵⁶ Zhu, H. (2019) kableExtra: Construct complex table with 'kable' and pipe syntax.

357 **Figure Captions**

358 **Fig. 1.** Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of
359 Saline and the Bay of Funtana (Mixed and Monospecific Settlements) sampled on 4 December 2017
360 and stained with SYBR Green I. Scale bar in all images is 60 µm.

361 **Fig. 2.** Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of
362 Funtana (Mixed and Monospecific Settlements) sampled on 19 June 2018 and stained with SYBR
363 Green I. Scale bar in all images is 60 µm.

364 **Fig. 3.** Taxonomic classification and relative contribution of the most abundant (> 1 %) bacterial
365 and archaeal sequences from surfaces of two marine macrophytes (*C. nodosa* and *C. cylindracea*)
366 sampled in the Bay of Saline and the Bay of Funtana (Mixed and Monospecific Settlements) and in
367 two contrasting seasons (4 December 2017 and 19 June 2018).

368 **Fig. 4.** Taxonomic classification and relative contribution of chloroplast sequences from surfaces
369 of two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline and the
370 Bay of Funtana (Mixed and Monospecific Settlements) and in two contrasting seasons (4 December
371 2017 and 19 June 2018).

372 **Figures**

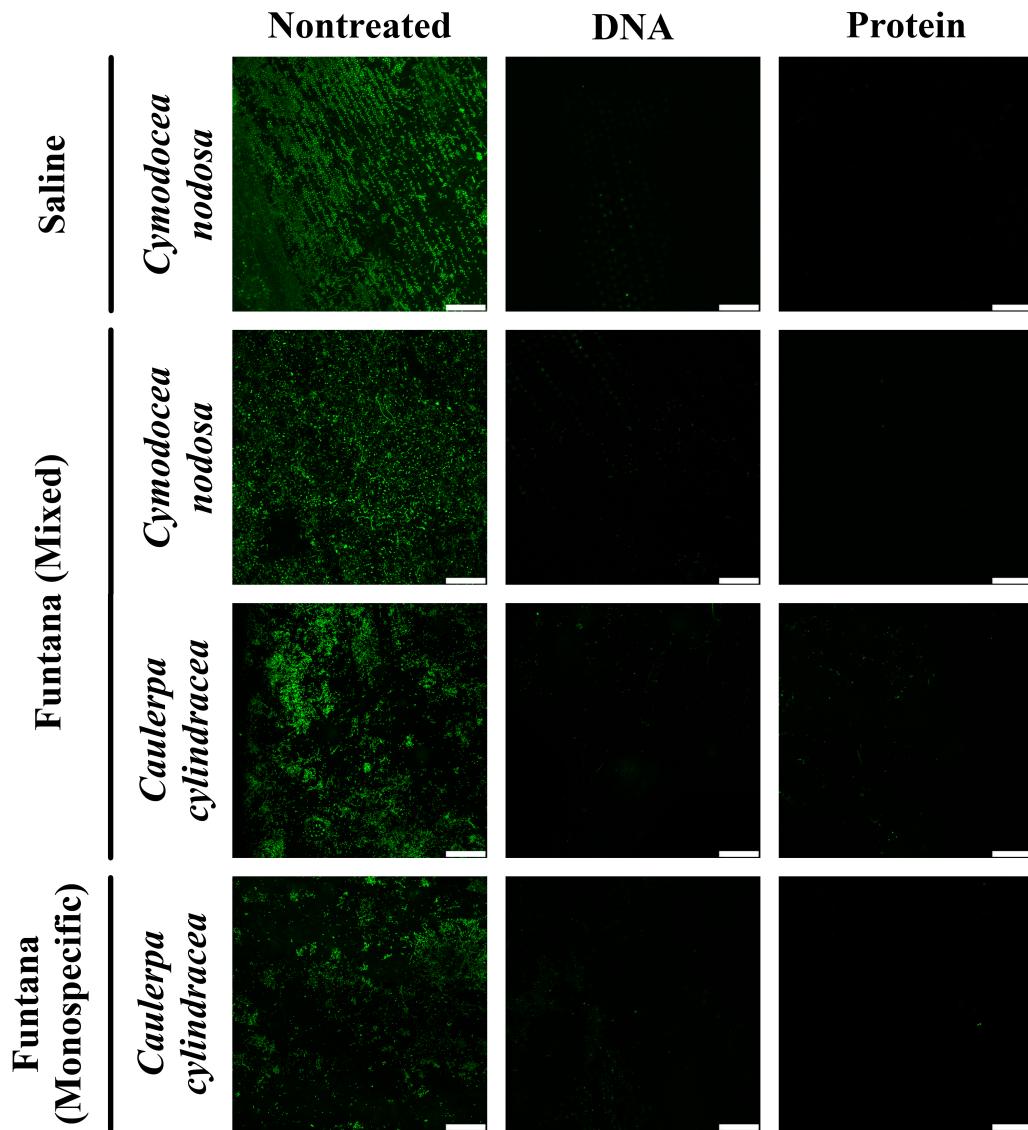


Fig. 1. Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of Saline and the Bay of Funtana (Mixed and Monospecific Settlements) sampled on 4 December 2017 and stained with SYBR Green I. Scale bar in all images is 60 µm.

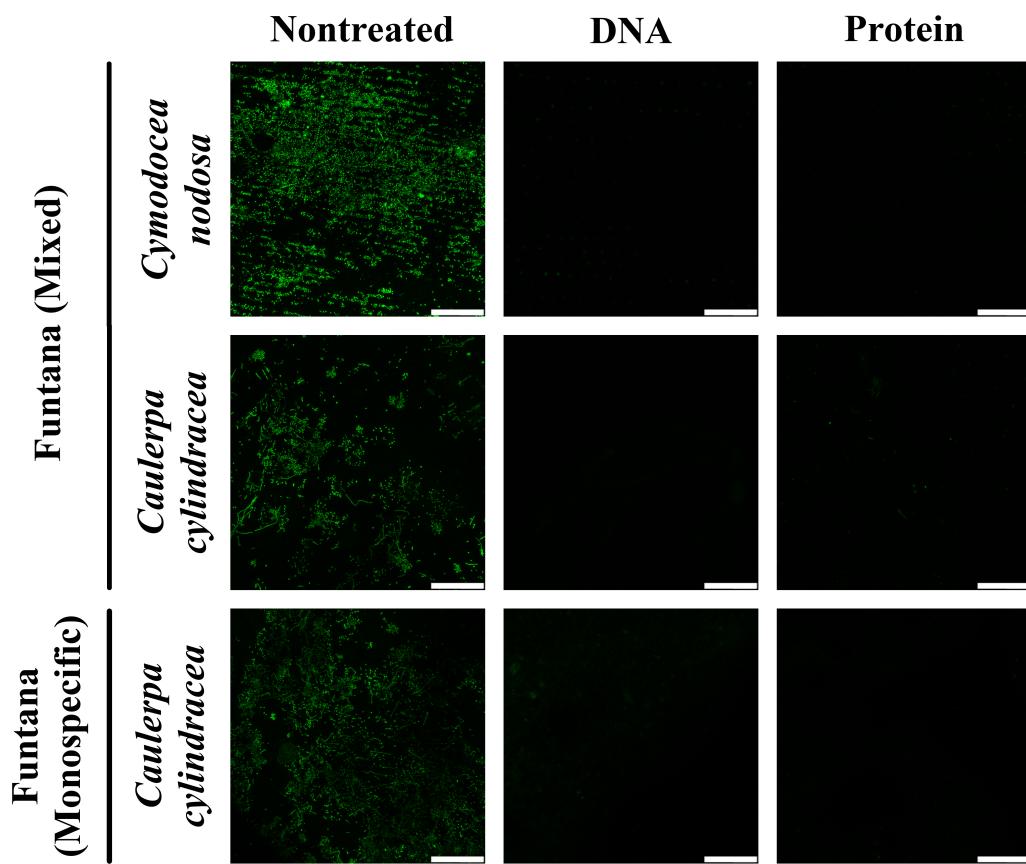


Fig. 2. Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of Funtana (Mixed and Monospecific Settlements) sampled on 19 June 2018 and stained with SYBR Green I. Scale bar in all images is 60 μm .

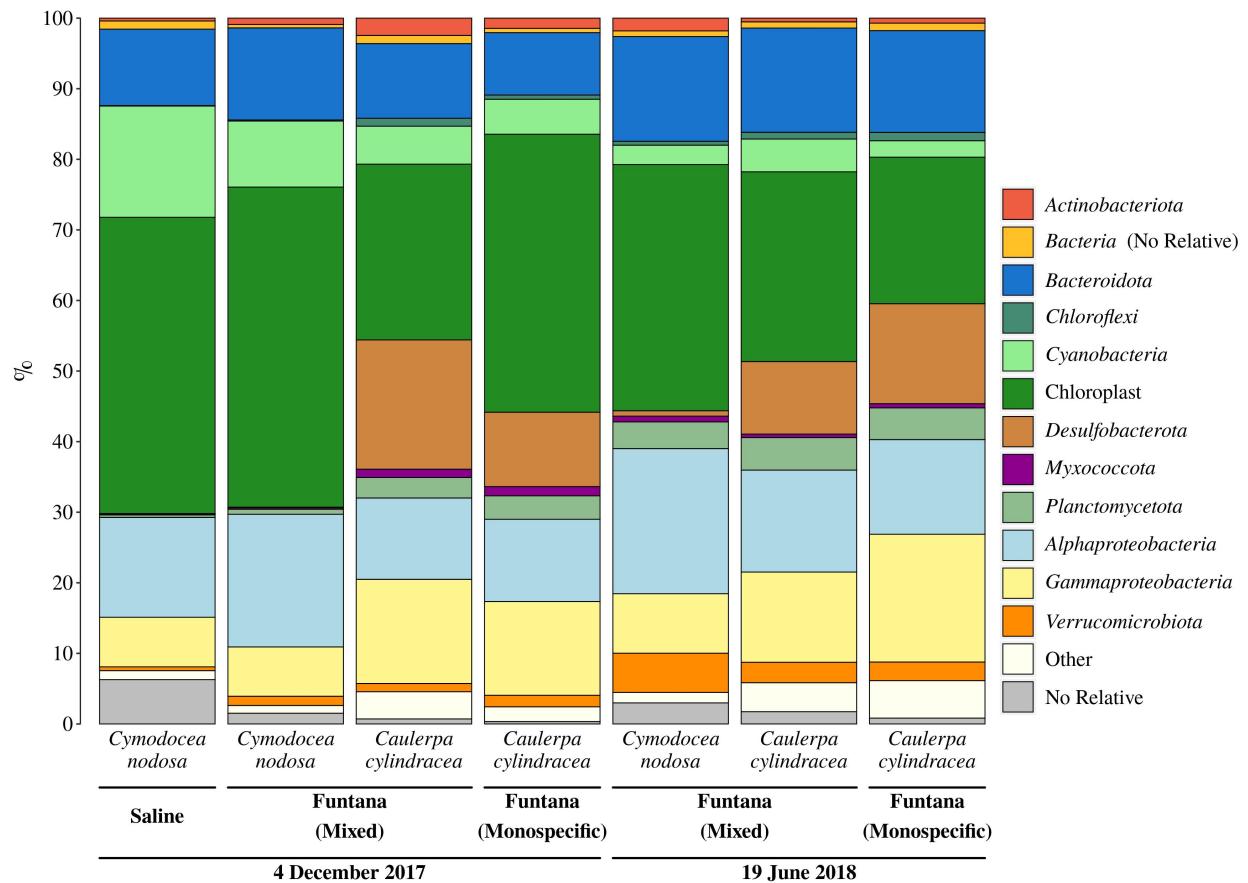


Fig. 3. Taxonomic classification and relative contribution of the most abundant (> 1 %) bacterial and archaeal sequences from surfaces of two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline and the Bay of Funtana (Mixed and Monospecific Settlements) and in two contrasting seasons (4 December 2017 and 19 June 2018).

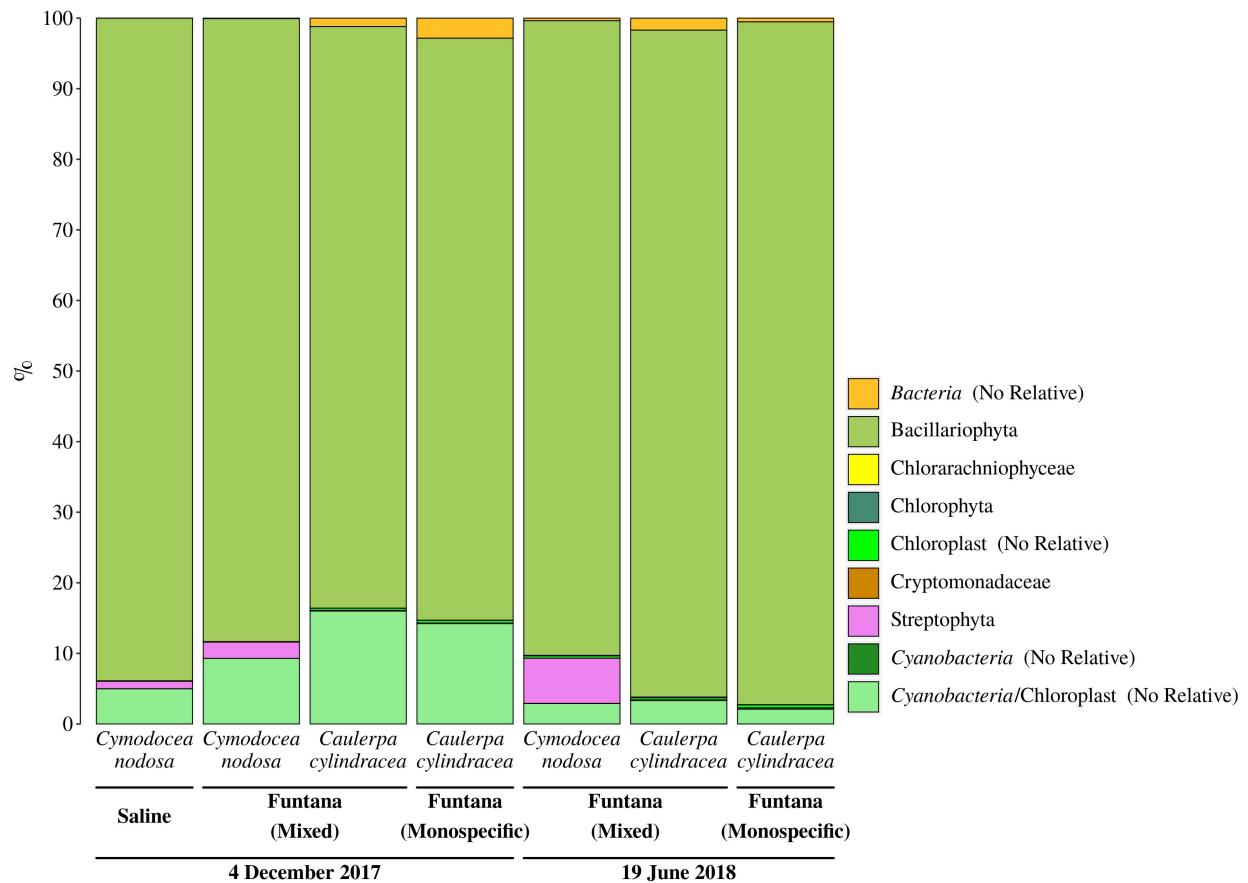


Fig. 4. Taxonomic classification and relative contribution of chloroplast sequences from surfaces of two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline and the Bay of Funtana (Mixed and Monospecific Settlements) and in two contrasting seasons (4 December 2017 and 19 June 2018).