

# Selective DNA and Protein Isolation from Marine Macrophyte Surfaces

Marino Korlević<sup>1\*</sup>, Marsej Markovski<sup>1</sup>, Zihao Zhao<sup>2</sup>, Gerhard J. Herndl<sup>2,3</sup>, Mirjana Najdek<sup>1</sup>

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

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

3. NIOZ, 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 Studies of unculturable microbes often combine methods based on DNA, such as 16S rRNA  
3 sequencing and metagenomics, with methods that allow insight into the metabolic status, such as  
4 metaproteomics. To apply these techniques to the microbial community inhabiting the surfaces of  
5 marine macrophytes it is advisable to perform, prior to the analysis, a selective DNA and protein  
6 isolation so that the host material, present in higher quantities, is not hampering the analysis.  
7 Two protocols, for DNA and protein isolation, were adapted for selective extractions of DNA  
8 and proteins from epiphytic communities inhabiting the surfaces of two marine macrophytes: the  
9 seagrass *Cymodocea nodosa* and the macroalga *Caulerpa cylindracea*. Protocols showed an almost  
10 complete removal of the epiphytic community regardless of the sampling season, station, settlement  
11 or host species. The obtained DNA was suitable for metagenomic and 16S rRNA sequencing, while  
12 isolated proteins could be identified by mass spectrometry, showing that protocols can be used in 16S  
13 rRNA, metagenomic and metaproteomic analysis. Low presence of host DNA and proteins, observed  
14 in isolated samples, indicated a selective nature of the protocols. Furthermore, the procedures are  
15 based on universally available laboratory chemicals making the protocols widely applicable. Taken  
16 together, the adapted protocols ensure an almost complete removal of the macrophyte epiphytic  
17 community, are selective for microbes inhabiting macrophyte surfaces and are providing DNA and  
18 proteins applicable in 16S rRNA sequencing, metagenomics and metaproteomics.

19 **Introduction**

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

26 Biological material (i.e. proteins and DNA) from pelagic microbial communities is usually  
27 isolated by collecting cells onto filters and subsequently isolating the aimed compound (Gilbert *et*  
28 *al.*, 2009). If a specific microbial size fraction is aimed sequential filtration is applied (Massana *et*  
29 *al.*, 1997; Andersson *et al.*, 2010). In contrast, obtaining biological materials from microorganism  
30 inhabiting surfaces requires either a cell detachment procedure prior to isolation or the host material  
31 is coextracted together with the targeted material. Methods for separating microbial cells form the  
32 host include shaking of host tissue (Gross *et al.*, 2003; Nõges *et al.*, 2010), scraping of macrophyte  
33 surfaces (Uku *et al.*, 2007) or the application of ultrasonication (Weidner *et al.*, 1996; Cai *et al.*,  
34 2014). It was shown that shaking alone is not sufficient to remove microbial cells, at least from  
35 plant root surfaces (Richter-Heitmann *et al.*, 2016). Manual separation methods, such as scraping  
36 and brushing, are time consuming and subjective, as the detachment efficiency depends on host  
37 tissue and the person performing the procedure (Cai *et al.*, 2014). Ultrasonication was proposed as  
38 an alternative method as it is providing better results in terms of detachment efficiency (Cai *et al.*,  
39 2014; Richter-Heitmann *et al.*, 2016). The downside of this procedure is that complete cell removal  
40 was still not obtained and tissue disruption was observed especially after the application of probe  
41 ultrasonication (Richter-Heitmann *et al.*, 2016). An alternative to these cell detachment procedures  
42 is the isolation of targeted epiphytic compounds together with host materials (Staufenberger *et*  
43 *al.*, 2008; Jiang *et al.*, 2015). This procedure can lead to problems in the following processing  
44 steps such as mitochondrial and chloroplast 16S rRNA sequence contaminations from the host

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

48 An alternative to these procedures is a direct isolation of the targeted material by incubating  
49 macrophyte tissues in an extraction buffer. After the incubation, the undisrupted host tissue is  
50 removed and the isolation procedure continues, omitting host material contaminations. To our  
51 knowledge the only procedure describing a direct and selective epiphytic DNA isolation from the  
52 surfaces of marine macrophytes was provided by Burke *et al.* (2009). In contrast to previously  
53 described methods this protocol enables an almost complete removal of the surface community and  
54 it was used for 16S rRNA gene clone library construction (Burke and Thomas *et al.*, 2011) and  
55 metagenome sequencing (Burke and Peter Steinberg *et al.*, 2011). This method, though providing a  
56 selective isolation procedure, uses a rapid multienzyme cleaner (3M) that is not available worldwide  
57 and without a known composition (Burke *et al.*, 2009). Also to our knowledge, no selective isolation  
58 protocol for proteins from epiphytic communities inhabiting marine macrophytes was established.

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

65 **Results**

66 To assess the removal efficiency of DNA and protein isolation procedures leaves and thalli were  
67 examined under a confocal microscope before and after treatments were performed. Developed  
68 procedures showed an almost complete removal of the surface community of both, *C. nodosa* and  
69 *C. cylindracea*. In addition, a similar removal efficiency was observed for communities sampled  
70 in contrasting seasons, December 2017 (Fig. 1) and June 2018 (Fig. 2). Also, no effect of station,  
71 settlement or isolation procedure (DNA or protein) on the removal efficiency was observed (Figs. 1  
72 and 2).

73 To evaluate if the obtained DNA is suitable to determine the microbial community structure  
74 an Illumina sequencing of the V4 16S rRNA region was performed. Sequencing yielded a total  
75 of 336,944 sequences after quality curation and exclusion of eukaryotic, mitochondrial and no  
76 relative sequences. The number of sequences classified as chloroplast was 97,334. After excluding  
77 these sequences the total number of retrieved reads dropped to 239,610, ranging from 22,596 to  
78 52,930 sequences per sample (Table S1). Even when the highest sequencing effort was applied the  
79 rarefaction curves did not level off which is commonly observed in high-throughput 16S rRNA  
80 amplicon sequencing procedures (Fig. S1). Sequence clustering at a similarity level of 97 % yielded a  
81 total of 8,355 different OTUs. Taxonomic classification of reads allowed for a macrophyte associated  
82 epiphytic community determination that was mainly composed of: *Alphaproteobacteria* ( $14.9 \pm 3.5$   
83 %), *Bacteroidota* ( $12.5 \pm 2.4$  %), *Gammaproteobacteria* ( $11.6 \pm 4.3$  %), *Desulfobacterota* ( $7.8 \pm$   
84  $7.5$  %), *Cyanobacteria* ( $6.5 \pm 4.7$  %) and *Planctomycetota* ( $2.9 \pm 1.7$  %) (Fig. 3).

85 Primers used in this study, as in many other 16S rRNA amplicon procedures, also amplified  
86 chloroplast 16S rRNA genes. We observed a high proportion of chloroplast sequences in all  
87 analyzed samples ( $33.4 \pm 9.4$  %) (Fig. 3). To determine if chloroplast sequences originate from  
88 hosts or eukaryotic epiphytic organisms, we exported SILVA classified chloroplast sequences  
89 and reclassified them using the RDP (Ribosomal Database Project) training set that allows for a

more detailed chloroplast classification. The largest proportion of sequences were classified as Bacillariophyta ( $89.7 \pm 5.7\%$ ) indicating that the DNA removal procedure did not coextract larger quantities of host DNA (Fig. 4). Chloroplast sequences classified as Streptophyta constituted  $3.3 \pm 2.8\%$  of all chloroplast sequences originating from *C. nodosa* samples, while sequences classified as Chlorophyta comprised only  $0.02 \pm 0.01\%$  of all chloroplast sequences associated with *C. cylindracea* samples.

To determine if the extracted DNA can be used for metagenomic sequencing four samples containing epiphytic DNA were selected and shotgun sequenced using an Illumina platform. Metagenomic sequencing yielded from 207,149,524 to 624,029,930 sequence pairs (Table S2). Obtained sequences were successfully assembled into contigs whose L50 ranged from 654 to 1,011 bp. In addition, predicted coding sequences were successfully functionally annotated (9,066,667 – 20,256,215 annotated sequences; Fig. 5) and taxonomically classified. Functional annotation allowed for an assessment of the relative contribution of each COG (Clusters of Orthologous Groups) functional category to the total number of annotated coding sequences (Fig. 5a). Functional categories containing the highest number of sequences were: C (Energy production and conversion), E (Amino acid transport and metabolism), M (Cell wall/membrane/envelope biogenesis), L (Replication, recombination and repair) and P (Inorganic ion transport and metabolism). If host DNA was coextracted with epiphytic it should be detected in larger proportions in sequenced metagenomes. Indeed, no higher proportions of coding sequences classified into phylum Streptophyta or Chlorophyta were detected (Table S3). Sequenced metagenomic DNA originating from the surface of *C. nodosa* contained 1.3 % of coding sequences classified into the phylum Streptophyta in December 2017 and 0.7 % in June 2018. Furthermore, the summed RPKM (Reads Per Kilobase Million) of these sequences constituted 1.7 % of total RPKM of all successfully classified sequences in December 2017 and 1.1 % in June 2018. Similar low proportions of host coding sequences were detected in metagenomic samples originating from the surfaces of *C. cylindracea*. Of all successfully classified coding sequences 0.2 % were classified into Chlorophyta in December 2017 and 0.1 % in June 2018. A relatively higher proportion of these sequences'

<sup>117</sup> RPKM than in the case of *C. nodosa* was observed, indicating a higher coextraction of host DNA in  
<sup>118</sup> *C. cylindracea*. In December, the proportion of RPKM of sequences classified into Chlorophyta  
<sup>119</sup> was 8.2 %, while in June 2018 it reached 13.6 %.

<sup>120</sup> To evaluate if the procedure for protein extraction is suitable for metaproteomic analysis,  
<sup>121</sup> obtained proteins were trypsin digested and sequenced using a mass spectrometer. Obtained  
<sup>122</sup> MS/MS spectra were searched against a protein database from sequenced metagenomes. From  
<sup>123</sup> 14,219 to 16,449 proteins were identified in isolated protein samples (Fig. 5b). In addition,  
<sup>124</sup> successful identification of proteins allowed for an assessment of the relative contribution of  
<sup>125</sup> each COG functional category to the total number of identified proteins (Fig. 5b). Functional  
<sup>126</sup> categories containing the highest number of identified proteins were: C (Energy production and  
<sup>127</sup> conversion), G (Carbohydrate transport and metabolism), P (Inorganic ion transport and metabolism),  
<sup>128</sup> O (Posttranslational modification, protein turnover, chaperones) and E (Amino acid transport and  
<sup>129</sup> metabolism). Isolated proteins could derive from epiphytic organisms inhabiting the macrophyte  
<sup>130</sup> surface or from macrophyte tissue underlying them. The contribution of proteins originating from  
<sup>131</sup> host tissues was evaluated by identifying all the proteins predicted to belong to any taxonomic group  
<sup>132</sup> within the phyla Streptophyta and Chlorophyta and by calculating their contribution to the number  
<sup>133</sup> and abundance of all identified proteins. On average, proteins isolated from the surface of *C. nodosa*  
<sup>134</sup> contained  $1.8 \pm 0.06$  % of proteins associated with Streptophyta, contributing to  $2.2 \pm 0.8$  % of  
<sup>135</sup> total proteins. Similar to metagenomes, proteins associated with Chlorophyta contributed more to  
<sup>136</sup> *C. cylindracea* than proteins associated with Streptophyta to *C. nodosa*. Chlorophyta associated  
<sup>137</sup> proteins comprised  $5.2 \pm 0.06$  % of all identified proteins in *C. cylindracea*, contributing to  $19.2 \pm$   
<sup>138</sup> 1.5 % of all protein abundances.

139 **Discussion**

140 The study of marine macrophyte epiphytic communities using culture independent techniques,  
141 such as 16S rRNA analysis, metagenomics and metaproteomics, requires an initial step of biological  
142 material isolation. Methods that have been developed for selective isolation of epiphytic biological  
143 material can be divided into two groups: (i) procedures involving, prior to extraction, a cell  
144 detachment step such as shaking (Gross *et al.*, 2003; Nõges *et al.*, 2010), scraping (Uku *et al.*, 2007)  
145 and ultrasonication (Weidner *et al.*, 1996; Cai *et al.*, 2014) and (ii) procedures involving a host  
146 tissue incubation aiming at direct lysis of epiphytic microbial cells (Burke *et al.*, 2009). Protocols  
147 that include a cell detachment procedure do not provide complete cell removal (Cai *et al.*, 2014;  
148 Richter-Heitmann *et al.*, 2016) and in the case of probe ultrasonication tissue disruption can also be  
149 observed (Richter-Heitmann *et al.*, 2016). Though, host tissue incubation procedures provide an  
150 almost complete cell removal and selective isolation, existing protocols like the one developed for  
151 DNA isolation by Burke *et al.* (2009) use in the extraction buffer a rapid multienzyme cleaner (3M)  
152 not available worldwide and without a known composition (Burke *et al.*, 2009). To circumvent  
153 these problems we developed and tested two protocols for selective DNA and protein isolation from  
154 marine macrophyte epiphytic communities.

155 To test if the developed DNA and protein isolation procedures could be applied on a variety  
156 of macrophyte species we selected *C. nodosa* and *C. cylindracea* as representatives of seagrass  
157 and macroalgal species, on which the procedures were tested. These species especially differ  
158 morphologically. While *C. nodosa* leaves are flat, *C. cylindracea* thallus is composed of uneven  
159 surfaces (Kuo and den Hartog, 2001; Verlaque *et al.*, 2003). The developed procedures showed an  
160 almost complete removal of epiphytic cells from the surfaces of both species comparable to the  
161 result of Burke *et al.* (2009) and indicating that structural differences do not impact the removal  
162 efficiency. In addition, isolation protocols were tested in two contrasting season as it is known that  
163 macrophytes are harboring more algal epiphytes during autumn and winter (Reyes and Sansón,  
164 2001). No differences in the removal efficiency was observed between seasons suggesting that

<sup>165</sup> protocols can be used on macrophyte samples retrieved throughout the year. Also, no removal  
<sup>166</sup> differences were observed on samples derived from the same host but from different localities or  
<sup>167</sup> settlements demonstrating the stability of the protocol in cell removal efficiency.

<sup>168</sup> Successful amplification and sequencing of the V4 16S rRNA gene region proved that the  
<sup>169</sup> isolated DNA can be used to estimate the microbial epiphytic diversity. Taxonomic groups detected  
<sup>170</sup> in this step can also be often found in epiphytic communities associated with other macrophytes  
<sup>171</sup> (Burke and Thomas *et al.*, 2011; Morrissey *et al.*, 2019). A problem often encountered in studies  
<sup>172</sup> focusing on epiphytic communities is the presence of large proportions of chloroplast 16S rRNA  
<sup>173</sup> sequences in the pool of amplified molecules, especially if the epiphytic DNA was isolated without  
<sup>174</sup> prior selection (Staufenberger *et al.*, 2008). These sequences can derive from host chloroplasts or  
<sup>175</sup> from eukaryotic epiphytic chloroplast DNA. Although the proportion of obtained chloroplast 16S  
<sup>176</sup> rRNA sequences in our samples was substantial they derived almost exclusively from eukaryotic  
<sup>177</sup> epiphytes. High proportion of chloroplast 16S rRNA sequences in studies applying selective  
<sup>178</sup> procedures that include direct cellular lysis on host surfaces were observed before (Michelou *et*  
<sup>179</sup> *al.*, 2013). It is possible that chloroplast specific sequences even in these studies are originating  
<sup>180</sup> from eukaryotic epiphytic cells and not host chloroplast. Indeed, it is common during 16S rRNA  
<sup>181</sup> profiling of pelagic microbial communities to observe high proportions of chloroplast sequences  
<sup>182</sup> (Gilbert *et al.*, 2009; Korlević *et al.*, 2016). In addition, a very low proportion of chloroplast 16S  
<sup>183</sup> rRNA sequences in samples originating from *C. cylindracea* in comparison to *C. nodosa* could be  
<sup>184</sup> explained by the presence of three introns in the gene for 16S rRNA in some members of the genus  
<sup>185</sup> *Caulerpa* that could hamper the amplification process (Lam and Lopez-Bautista, 2016).

<sup>186</sup> Beside 16S rRNA sequencing high quality DNA is needed for metagenomics. The obtained  
<sup>187</sup> number of metagenomic sequences and assembly statistics were comparable to metagenomes and  
<sup>188</sup> metatranscriptomes derived from similar surface associated communities (Crump *et al.*, 2018; Cúcio  
<sup>189</sup> *et al.*, 2018). In addition, functional annotation of predicted coding sequences to COG functional  
<sup>190</sup> categories showed that the obtained metagenomes can be used to determine the metabolic capacity

191 of surface associated communities (Leary *et al.*, 2014; Cúcio *et al.*, 2018). The proportion of  
192 coding sequences, including their RPKM, originating from *C. nodosa* metagenomes and classified  
193 as Streptophyta was low indicating a good selectivity of the isolation procedure towards epiphytic  
194 cells. In the case of DNA samples isolated from the surface of *C. cylindracea* the proportion of  
195 Chlorophyta coding sequences was also low but their RPKM was higher than in the case of *C.*  
196 *nodosa*. One of the causes for this elevated RPKM of Chlorophyta sequences in *C. cylindracea*  
197 could lay in tissue structure differences between these two host species. While *C. nodosa* leaves  
198 are composed of individual cells, the thallus of *C. cylindracea* is, like in other siphonous algal  
199 species, composed of a single large multinucleate cell (Coneva and Chitwood, 2015). The absence  
200 of individual cells in *C. cylindracea* could cause a leakage of genetic material into the extraction  
201 buffer causing an elevated presence of host sequences in metagenomic data.

202 To obtain an insight into the metabolic status of uncultivated members, a metaproteomic  
203 approach is required (Saito *et al.*, 2019). The applied protocol for epiphytic protein isolation  
204 followed by a metaproteomic analysis identified between 14,219 and 16,449 proteins, which is  
205 higher than previously reported for e.g. soils (Chourey *et al.*, 2010; Hultman *et al.*, 2015), seawater  
206 (Williams *et al.*, 2012) and biofilms (Leary *et al.*, 2014). The functional annotation of identified  
207 proteins into COG functional categories showed that the protein isolation procedure can be used to  
208 assess the metabolic status of the macrophyte epiphytic community (Leary *et al.*, 2014). Similar  
209 to metagenomes, the number of identified proteins, including their abundances, associated with  
210 Streptophyta in *C. nodosa* samples were low indicating that the procedure is selective for epiphytic  
211 cell proteins. In addition, a higher number of identified proteins, and especially their abundances,  
212 associated with Chlorophyta was observed in *C. cylindracea* samples. The cause of this elevated  
213 presence of Chlorophyta associated proteins can be, as in the case of the DNA isolation procedure,  
214 explained by the absence of individual cells in this siphonous alga (Coneva and Chitwood, 2015).

215 In conclusion, the developed protocols for DNA and protein isolation from macrophyte surfaces  
216 are providing an almost complete removal of the epiphytic community and are shown to ensure

*removal from both, *C. nodosa* and *C. cylindracea*, in different seasons. Also, the obtained DNA and proteins are suitable for 16S rRNA sequencing, metagenomics and metaproteomics, while the obtained material contains low quantities of host DNA or proteins making the procedures epiphyte selective. Furthermore, the procedures are based on universally available laboratory chemicals making the protocols widely applicable.*

222 **Experimental procedures**

223 **Sampling**

224 Leaves of *C. nodosa* were sampled in a *C. nodosa* meadow in the Bay of Saline (45°7'5''  
225 N, 13°37'20'' E) and in a *C. nodosa* meadow invaded by *C. cylindracea* in the proximity of the  
226 village of Funtana (Bay of Funtana; 45°10'39'' N, 13°35'42'' E). Thalli of *C. cylindracea* were  
227 sampled in the same *C. nodosa* invaded meadow in the Bay of Funtana and on a locality of  
228 only *C. cylindracea* located in the proximity of the invaded meadow. Leaves and thalli for 16S  
229 rRNA analysis, metagenomics and metaproteomics were collected in two contrasting seasons, on 4  
230 December 2017 (16S rRNA analysis and metaproteomics), 14 December 2017 (metagenomics) and  
231 18 June 2018 (16S rRNA analysis, metagenomics and metaproteomics). During spring 2018 the *C.*  
232 *nodosa* meadow in the Bay of Saline decayed to an extent that no leaves could be retrieved (Najdek  
233 *et al.*, 2020). In addition, as not enough DNA for both metagenomic and 16S RNA analysis were  
234 obtained during the sampling on 4 December 2017, an additional sampling on 14 December 2017  
235 was carried out in the Bay of Funtana. Leaves and thalli were collected by diving and transported to  
236 the laboratory in containers placed on ice and filled with site seawater. Upon arrival to the laboratory,  
237 *C. nodosa* leaves were cut into sections of 1 – 2 cm, while *C. cylindracea* thalli were cut into 5 – 8  
238 cm long sections. Leaves and thalli were washed three times with sterile artificial seawater (ASW)  
239 to remove loosely attached microbial cells.

240 **DNA isolation**

241 The DNA was isolated according to the protocol for isolation from filters described in Massana  
242 *et al.* (1997). This protocol was modified and adapted for microbial DNA isolation from macrophyte  
243 surfaces as described below. 5 ml of lysis buffer (40 mM EDTA, 50 mM Tris-HCl, 0.75 M sucrose;  
244 pH 8.3) was added to 1 g wet weight of leaves or 2 g wet-weight of thalli. For every sample

245 duplicate isolations were performed. Lysozyme was added (final concentration 1 mg ml<sup>-1</sup>) and the  
246 mixture was incubated at 37 °C for 30 min. Subsequently, proteinase K (final concentration 0.5  
247 mg ml<sup>-1</sup>) and SDS (final concentration 1 %) were added and the samples were incubated at 55 °C  
248 for 2 h. Following the incubation, tubes were vortexed for 10 min and the mixture containing lysed  
249 epiphytic cells was separated from host leaves or thalli by transferring the solution into a clean tube.  
250 The lysate was extracted twice with a mixture of phenol:chloroform:isoamyl alcohol (25:24:1; pH  
251 8) and once with chloroform:isoamyl alcohol (24:1). After each organic solvent mixture addition  
252 tubes were slightly vortexed and centrifuged at 4,500 × g for 10 min. Following each centrifugation  
253 aqueous phases were retrieved. After the final extraction 1/10 of chilled 3 M sodium acetate (pH  
254 5.2) was added. DNA was precipitated by adding 1 volume of chilled isopropanol, incubating  
255 the mixtures overnight at -20 °C and centrifuging at 16,000 × g and 4 °C for 20 min. The pellet  
256 was washed twice with 1 ml of chilled 70 % ethanol and centrifuged after each washing step at  
257 20,000 × g and 4 °C for 10 min. After the first washing step duplicate pellets form the same sample  
258 were pooled and transferred to a clean 1.5 ml tube. The dried pellet was resuspended in 100 µl of  
259 deionized water.

260 **Illumina 16S rRNA sequencing**

261 An aliquot of isolated DNA was treated with RNase A (final concentration 200 µg ml<sup>-1</sup>)  
262 for 2 h at 37 °C. The DNA concentration was determined using the Quant-iT PicoGreen  
263 dsDNA Assay Kit (Invitrogen, USA) according to the manufacturer's instructions and  
264 diluted to 1 ng µl<sup>-1</sup>. The V4 region of the 16S rRNA gene was amplified using a two-step  
265 PCR procedure. In the first PCR the 515F (5'-GTGYCAGCMGCCGCGGTAA-3') and  
266 806R (5'-GGACTACNVGGGTWTCTAAT-3') primers from the Earth Microbiome Project  
267 (<http://press.igsb.anl.gov/earthmicrobiome/protocols-and-standards/16s/>) were used to amplify  
268 the target region (Caporaso *et al.*, 2012; Apprill *et al.*, 2015; Parada *et al.*, 2016). These primers  
269 contained on their 5' end a tagged sequence. Each sample was amplified in four parallel 25 µl

reactions of which each contained: 1 × Q5 Reaction Buffer , 0.2 mm of dNTPmix, 0.7 mg ml<sup>-1</sup> BSA (Bovine Serum Albumin), 0.2 µm of forward and reverse primers, 0.5 U of Q5 High-Fidelity DNA Polymerase (New England Biolabs, USA) and 5 ng of DNA template. Cycling conditions were: initial denaturation at 94 °C for 3 min, 20 cycles of denaturation at 94 °C for 45 s, annealing at 50 °C for 60 s and elongation at 72 °C for 90 s, finalized by an elongation step at 72 °C for 10 min. The four parallel reactions volumes were pooled and PCR products were purified using the GeneJET PCR Purification Kit (Thermo Fisher Scientific, USA) according to the manufacturer's instructions and following the protocol that included isopropanol addition for better small DNA fragment yield. The column was eluted in 30 µl of deionized water. Purified PCR products were sent for Illumina MiSeq sequencing (2 × 250 bp) at IMGM Laboratories, Martinsried, Germany. Before sequencing at IMGM, the second PCR amplification of the two-step PCR procedure was performed using primers targeting the tagged region incorporated in the first PCR. In addition, these primers contained adapter and sample-specific index sequences. The second PCR was carried out for 8 cycles. Beside samples, a positive and negative control were sequenced. A negative control was comprised of four parallel PCR reactions without DNA template, while for a positive control a mock community composed of evenly mixed DNA material originating from 20 bacterial strains (ATCC MSA-1002, ATCC, USA) was used. Partial 16S rRNA sequences obtained in this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession numbers SAMEA6786270, SAMEA6648792 – SAMEA6648794, SAMEA6648809 – SAMEA6648811.

Obtained sequences were analyzed on the computer cluster Isabella (University Computing Center, University of Zagreb) using mothur (version 1.43.0) (Schloss *et al.*, 2009) according to the MiSeq Standard Operating Procedure (MiSeq SOP; [https://mothur.org/wiki/MiSeq\\_SOP](https://mothur.org/wiki/MiSeq_SOP)) (Kozich *et al.*, 2013) and recommendations given from the Riffomonas project to enhance data reproducibility (<http://www.riffomonas.org/>). For alignment and classification of sequences the SILVA SSU Ref NR 99 database (release 138; <http://www.arb-silva.de>) was used (Quast *et al.*, 2013; Yilmaz *et al.*, 2014). Sequences classified as chloroplasts by SILVA were exported and reclassified using

297 mothur and the RDP (<http://rdp.cme.msu.edu/>) training set (version 16) reference files adapted for  
298 mothur (Cole *et al.*, 2014). In comparison to SILVA, RDP allows a more detailed classification of  
299 chloroplast sequences. Based on the ATCC MSA-1002 mock community included in the analysis a  
300 sequencing error rate of 0.009 % was determined, which is in line with previously reported values  
301 for next-generation sequencing data (Kozich *et al.*, 2013; Schloss *et al.*, 2016). In addition, the  
302 negative control processed together with the samples yielded only 2 sequences after sequence  
303 quality curation.

304 **Metagenomics**

305 Four DNA samples were selected and sent on dry ice to IMGM Laboratories, Martinsried,  
306 Germany for metagenomic sequencing. DNA was purified using AMPure XP Beads (Beckman  
307 Coulter, USA) applying a bead:DNA ratio of 1:1 (v/v), quantified with a Qubit dsDNA HS Assay  
308 Kit (Thermo Fisher Scientific, USA) and check for integrity on a 1 % agarose gel. Metagenomic  
309 sequencing libraries were generated from 300 ng of input DNA using a NEBNext Ultra II FS  
310 DNA Library Prep Kit for Illumina (New England Biolabs, USA) according to the manufacturer's  
311 instructions. Fragments were selected (500 – 700 bp) using AMPure XP Beads, PCR enriched for 3  
312 – 5 cycles and quality controlled. Libraries generated from different DNA samples were pooled and  
313 sequenced on an Illumina NovaSeq 6000 sequencing system (2 × 250 bp).

314 Obtained sequences were analyzed on a Life Science Compute Cluster (LiSC) (CUBE –  
315 Computational Systems Biology, University of Vienna). Individual sequences were assembled using  
316 MEGAHIT (version 1.1.2) (Li *et al.*, 2015) under default settings. Putative genes were predicted  
317 from contings longer than 200 bp using Prodigal (version 2.6.3) (Hyatt *et al.*, 2010) in metagenome  
318 mode (-p meta). Abundances of predicted genes were expressed as Reads Per Kilobase Million  
319 (RPKM) and calculated using the BWA algorithm (version 0.7.16a) (Li and Durbin, 2009). All  
320 predicted genes were functionally annotated using the eggNOG-mapper (Huerta-Cepas *et al.*, 2017)  
321 and eggNOG database (version 5.0) (Huerta-Cepas *et al.*, 2019). Sequence taxonomy classification

322 was determined using the lowest common ancestor algorithm adapted from DIAMOND (version  
323 0.8.36) (Buchfink *et al.*, 2015) and by searching against the NCBI non-redundant database (NR). To  
324 determine phylogeny the top 10 % hits with an e-value  $< 1 \times 10^{-5}$  were used (--top 10). Sequence  
325 renaming, coverage information computing and metagenomic statistics calculations were performed  
326 using software tools from BBTools (<https://jgi.doe.gov/data-and-tools/bbtools>). Metagenomic  
327 sequences obtained in this study have been deposited in the European Nucleotide Archive (ENA) at  
328 EMBL-EBI under accession numbers SAMEA6648795, SAMEA6648797, SAMEA6648809 and  
329 SAMEA6648811.

330 **Protein isolation**

331 Proteins were isolated according to the protocol for protein isolation from soil described in  
332 Chourey *et al.* (2010) and modified by Hultman *et al.* (2015). This protocol was further modified  
333 and adapted for microbial protein isolation from macrophyte surfaces as described below. 20 ml  
334 of protein extraction buffer (4 % SDS, 100 mM Tris-HCl [pH 8.0]) was added to 5 g wet weight  
335 of leaves or 10 g wet weight of thalli. The mixture was incubated in boiling water for 5 min,  
336 vortexed for 10 min and incubated again in boiling water for 5 min. After a brief vortex the lysate  
337 was transferred to a clean tube separating the host leaves or thalli from the mixture containing  
338 lyzed epiphytic cells. Dithiothreitol (DTT; final concentration 24 mM) was added and proteins were  
339 precipitated with chilled 100 % trichloroacetic acid (TCA; final concentration 20 %) overnight at  
340  $-20^{\circ}\text{C}$ . Precipitated proteins were centrifuged at  $10,000 \times g$  and  $4^{\circ}\text{C}$  for 40 min. The obtained  
341 protein pellet was washed three times with chilled acetone. During the first washing step the pellet  
342 was transferred to a clean 1.5 ml tube. After each washing step samples were centrifuged at 20,000  
343  $\times g$  and  $4^{\circ}\text{C}$  for 5 min. Dried pellets were stored at  $-80^{\circ}\text{C}$  until further analysis.

344 **Metaproteomics**

345 Isolated proteins were whole trypsin digested using the FASP (filter-aided sample preparation)  
346 Protein Digestion Kit (Expedeon, UK) according to the manufacturer's instructions with small  
347 modifications (Wiśniewski *et al.*, 2009). Before the solution was loaded to the column, protein  
348 pellets were solubilized in a urea sample buffer included in the kit amended with DTT (final  
349 concentration 100 mM) for 45 min at room temperature and centrifuged at 20,000 × g for 2 – 5 min  
350 at room temperature to remove larger particles. The first washing step after protein solution loading  
351 was repeated twice. In addition, centrifugation steps were prolonged if the column was clogged.  
352 Trypsin digestion was performed on column filters at 37 °C overnight for 18 h. The final filtrate  
353 containing peptides was acidified with 1 % (final concentration) trifluoroacetic acid, freezed at –80  
354 °C for 15 min, lyophilized and sent to the Vienna Metabolomics Center (University of Vienna) for  
355 metaproteomic analysis. Peptides were resuspended in 1 % (final concentration) trifluoroacetic  
356 acid (TFA), desalted using the Pierce C18 Tips (Thermo Fisher Scientific, USA) according to the  
357 manufacturer's instructions and sequenced on a Q Exactive Hybrid Quadrupole-Orbitrap Mass  
358 Spectrometer (Thermo Fisher Scientific, USA). Obtained MS/MS spectra were searched against a  
359 protein database composed of combined sequenced metagenomes using SEQUEST-HT engines  
360 and validated with Percolator in Proteome Discoverer 2.1 (Thermo Fisher Scientific, USA). The  
361 target-decoy approach was used to reduce the probability of false peptide identification. Results  
362 whose false discovery rate at the peptide level was <1 % were kept. For protein identification a  
363 minimum of two peptides and one unique peptide were required. For protein quantification, a  
364 chromatographic peak area-based free quantitative method was applied.

365 **Data processing and visualization**

366 Processing and visualization of 16S rRNA, metagenomic and metaproteomic data was done  
367 using R (version 3.6.0) (R Core Team, 2019), package tidyverse (version 1.3.0) (Wickham *et*

<sup>368</sup> *al.*, 2019) and multiple other packages (Xie, 2014, 2015, 2019, 2020; Neuwirth, 2014; Xie *et*  
<sup>369</sup> *al.*, 2018; Wilke, 2018; Allaire *et al.*, 2019; Zhu, 2019; Bengtsson, 2020). The detailed analysis  
<sup>370</sup> procedure including the R Markdown file for this paper are available as a GitHub repository  
<sup>371</sup> ([https://github.com/MicrobesRovinj/Korlevic\\_SelectiveRemoval\\_EnvironMicrobiol\\_2020](https://github.com/MicrobesRovinj/Korlevic_SelectiveRemoval_EnvironMicrobiol_2020)).

## <sup>372</sup> **Confocal microscopy**

<sup>373</sup> Host leaves and thalli from DNA and protein isolation steps were washed seven times in  
<sup>374</sup> deionized water and fixed with formaldehyde (final concentration ~ 3 %). In addition, nontreated  
<sup>375</sup> leaves and thalli, washed three times in ASW to remove loosely attached microbial cells, were fixed  
<sup>376</sup> in the same concentration of formaldehyde and used as a positive control. For long therm storage,  
<sup>377</sup> fixed leaves and thalli were immersed in a mixture of phosphate-buffered saline (PBS) and ethanol  
<sup>378</sup> (1:1) and stored at –20 °C. Treated and untreated segemnts of leaves and thalli were stained in a 2  
<sup>379</sup> × solution of SYBR Green I and examined under a Leica TCS SP8 X FLIM confocal microscope  
<sup>380</sup> (Leica Microsystems, Germany).

## <sup>381</sup> **Acknowledgements**

<sup>382</sup> This work was founded by the Croatian Science Foundation through the MICRO-SEAGRASS  
<sup>383</sup> project (IP-2016-06-7118). We would like to thank Margareta Buterer for technical support, Paolo  
<sup>384</sup> Paliaga for help during sampling, Lucija Horvat for technical support in confocal microscopy and  
<sup>385</sup> Dušica Vujaklija for help in accessing the confocal microscope. In addition, we would like to thank  
<sup>386</sup> the University Computing Center of the University of Zagreb for access to the computer cluster  
<sup>387</sup> Isabella and CUBE – Computational Systems Biology of the University of Vienna for access to the  
<sup>388</sup> Life Science Compute Cluster (LiSC).

389 **References**

- 390 Allaire, J.J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., et al. (2019)
- 391 rmarkdown: Dynamic documents for R.
- 392 Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and *in situ*  
393 detection of individual microbial cells without cultivation. *Microbiol Rev* **59**: 143–169.
- 394 Andersson, A.F., Riemann, L., and Bertilsson, S. (2010) Pyrosequencing reveals contrasting  
395 seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *ISME J* **4**: 171–181.
- 396 Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015) Minor revision to V4 region SSU  
397 rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat Microb*  
398 *Ecol* **75**: 129–137.
- 399 Bengtsson, H. (2020) matrixStats: Functions that apply to rows and columns of matrices (and  
400 to vectors).
- 401 Buchfink, B., Xie, C., and Huson, D.H. (2015) Fast and sensitive protein alignment using  
402 DIAMOND. *Nat Methods* **12**: 59–60.
- 403 Burke, C., Kjelleberg, S., and Thomas, T. (2009) Selective extraction of bacterial DNA from  
404 the surfaces of macroalgae. *Appl Environ Microbiol* **75**: 252–256.
- 405 Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011) Bacterial  
406 community assembly based on functional genes rather than species. *Proc Natl Acad Sci U S A* **108**:  
407 14288–14293.
- 408 Burke, C., Thomas, T., Lewis, M., Steinberg, P., and Kjelleberg, S. (2011) Composition,  
409 uniqueness and variability of the epiphytic bacterial community of the green alga *Ulva australis*.  
410 *ISME J* **5**: 590–600.

- 411 Cai, X., Gao, G., Yang, J., Tang, X., Dai, J., Chen, D., and Song, Y. (2014) An ultrasonic  
412 method for separation of epiphytic microbes from freshwater submerged macrophytes. *J Basic*  
413 *Microbiol* **54**: 758–761.
- 414 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., et al.  
415 (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq  
416 platforms. *ISME J* **6**: 1621–1624.
- 417 Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., et al.  
418 (2010) Direct cellular lysis/protein extraction protocol for soil metaproteomics. *J Proteome Res* **9**:  
419 6615–6622.
- 420 Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., et al. (2014) Ribosomal  
421 Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res* **42**:  
422 D633–D642.
- 423 Coneva, V. and Chitwood, D.H. (2015) Plant architecture without multicellularity: Quandaries  
424 over patterning and the soma-germline divide in siphonous algae. *Front Plant Sci* **6**: 287.
- 425 Crump, B.C., Wojahn, J.M., Tomas, F., and Mueller, R.S. (2018) Metatranscriptomics and  
426 amplicon sequencing reveal mutualisms in seagrass microbiomes. *Front Microbiol* **9**: 388.
- 427 Cúcio, C., Overmars, L., Engelen, A.H., and Muyzer, G. (2018) Metagenomic analysis shows  
428 the presence of bacteria related to free-living forms of sulfur-oxidizing chemolithoautotrophic  
429 symbionts in the rhizosphere of the seagrass *Zostera marina*. *Front Mar Sci* **5**: 171.
- 430 Egan, S., Harder, T., Burke, C., Steinberg, P., Kjelleberg, S., and Thomas, T. (2013) The  
431 seaweed holobiont: Understanding seaweed-bacteria interactions. *FEMS Microbiol Rev* **37**:  
432 462–476.
- 433 Gilbert, J.A., Field, D., Swift, P., Newbold, L., Oliver, A., Smyth, T., et al. (2009) The

<sup>434</sup> seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol* **11**:  
<sup>435</sup> 3132–3139.

<sup>436</sup> Gross, E.M., Feldbaum, C., and Graf, A. (2003) Epiphyte biomass and elemental composition  
<sup>437</sup> on submersed macrophytes in shallow eutrophic lakes. *Hydrobiologia* **506-509**: 559–565.

<sup>438</sup> Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C.,  
<sup>439</sup> and Bork, P. (2017) Fast genome-wide functional annotation through orthology assignment by  
<sup>440</sup> eggNOG-mapper. *Mol Biol Evol* **34**: 2115–2122.

<sup>441</sup> Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H.,  
<sup>442</sup> et al. (2019) eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology  
<sup>443</sup> resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**: D309–D314.

<sup>444</sup> Hultman, J., Waldrop, M.P., Mackelprang, R., David, M.M., McFarland, J., Blazewicz, S.J., et  
<sup>445</sup> al. (2015) Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature*  
<sup>446</sup> **521**: 208–212.

<sup>447</sup> Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010)  
<sup>448</sup> Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC*  
<sup>449</sup> *Bioinformatics* **11**: 119.

<sup>450</sup> Jiang, Y.-F., Ling, J., Dong, J.-D., Chen, B., Zhang, Y.-Y., Zhang, Y.-Z., and Wang, Y.-S. (2015)  
<sup>451</sup> Illumina-based analysis the microbial diversity associated with *Thalassia hemprichii* in Xincun Bay,  
<sup>452</sup> South China Sea. *Ecotoxicology* **24**: 1548–56.

<sup>453</sup> Korlević, M., Šupraha, L., Ljubešić, Z., Henderiks, J., Ciglenečki, I., Dautović, J., and Orlić, S.  
<sup>454</sup> (2016) Bacterial diversity across a highly stratified ecosystem: A salt-wedge Mediterranean estuary.  
<sup>455</sup> *Syst Appl Microbiol* **39**: 398–408.

<sup>456</sup> Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013)

457 Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon  
458 sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* **79**: 5112–5120.

459 Kuo, J. and den Hartog, C. (2001) Seagrass taxonomy and identification key. In *Global*  
460 *Seagrass Research Methods*. Short, F.T. and Coles, R.G. (eds). Amsterdam: Elsevier Science B.V.,  
461 pp. 31–58.

462 Lam, D.W. and Lopez-Bautista, J.M. (2016) Complete chloroplast genome for *Caulerpa*  
463 *racemosa* (Bryopsidales, Chlorophyta) and comparative analyses of siphonous green seaweed  
464 plastomes. *Cymbella* **2**: 23–32.

465 Leary, D.H., Li, R.W., Hamdan, L.J., IV, W.J.H., Lebedev, N., Wang, Z., et al. (2014)  
466 Integrated metagenomic and metaproteomic analyses of marine biofilm communities. *Biofouling*  
467 **30**: 1211–1223.

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

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

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

476 Massana, R., Murray, A.E., Preston, C.M., and DeLong, E.F. (1997) Vertical distribution and  
477 phylogenetic characterization of marine planktonic *Archaea* in the Santa Barbara Channel. *Appl*  
478 *Environ Microbiol* **63**: 50–56.

479 Michelou, V.K., Caporaso, J.G., Knight, R., and Palumbi, S.R. (2013) The ecology of microbial

- 480 communities associated with *Macrocystis pyrifera*. *PloS One* **8**: e67480.
- 481 Morrissey, K.L., Çavas, L., Willems, A., and De Clerck, O. (2019) Disentangling the influence  
482 of environment, host specificity and thallus differentiation on bacterial communities in siphonous  
483 green seaweeds. *Front Microbiol* **10**: 717.
- 484 Najdek, M., Korlević, M., Paliaga, P., Markovski, M., Ivančić, I., Iveša, L., et al. (2020)  
485 Dynamics of environmental conditions during the decline of a *Cymodocea nodosa* meadow.  
486 *Biogeosciences* **17**: 3299–3315.
- 487 Neuwirth, E. (2014) RColorBrewer: ColorBrewer palettes.
- 488 Nõges, T., Luup, H., and Feldmann, T. (2010) Primary production of aquatic macrophytes and  
489 their epiphytes in two shallow lakes (Peipsi and Võrtsjärv) in Estonia. *Aquat Ecol* **44**: 83–92.
- 490 Parada, A.E., Needham, D.M., and Fuhrman, J.A. (2016) Every base matters: Assessing small  
491 subunit rRNA primers for marine microbiomes with mock communities, time series and global field  
492 samples. *Environ Microbiol* **18**: 1403–1414.
- 493 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA  
494 ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic  
495 Acids Res* **41**: D590–D596.
- 496 R Core Team (2019) R: A language and environment for statistical computing, Vienna, Austria:  
497 R Foundation for Statistical Computing.
- 498 Reyes, J. and Sansón, M. (2001) Biomass and production of the epiphytes on the leaves of  
499 *Cymodocea nodosa* in the Canary Islands. *Bot Mar* **44**: 307–313.
- 500 Richter-Heitmann, T., Eickhorst, T., Knauth, S., Friedrich, M.W., and Schmidt, H. (2016)  
501 Evaluation of strategies to separate root-associated microbial communities: A crucial choice in  
502 rhizobiome research. *Front Microbiol* **7**: 773.

- 503 Saito, M.A., Bertrand, E.M., Duffy, M.E., Gaylord, D.A., Held, N.A., Hervey, W.J., et al.  
504 (2019) Progress and challenges in ocean metaproteomics and proposed best practices for data  
505 sharing. *J Proteome Res* **18**: 1461–1476.
- 506 Schloss, P.D., Jenior, M.L., Koumpouras, C.C., Westcott, S.L., and Highlander, S.K. (2016)  
507 Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. *PeerJ* **4**:  
508 e1869.
- 509 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al.  
510 (2009) Introducing mothur: Open-source, platform-independent, community-supported software for  
511 describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.
- 512 Staufenberger, T., Thiel, V., Wiese, J., and Imhoff, J.F. (2008) Phylogenetic analysis of bacteria  
513 associated with *Laminaria saccharina*. *FEMS Microbiol Ecol* **64**: 65–77.
- 514 Su, C., Lei, L., Duan, Y., Zhang, K.-Q., and Yang, J. (2012) Culture-independent methods for  
515 studying environmental microorganisms: Methods, application, and perspective. *Appl Microbiol*  
516 *Biotechnol* **93**: 993–1003.
- 517 Uku, J., Björk, M., Bergman, B., and Díez, B. (2007) Characterization and comparison of  
518 prokaryotic epiphytes associated with three East African seagrasses. *J Phycol* **43**: 768–779.
- 519 Verlaque, M., Durand, C., Huisman, J.M., Boudouresque, C.-F., and Le Parco, Y. (2003) On the  
520 identity and origin of the Mediterranean invasive *Caulerpa Racemosa* (Caulerpales, Chlorophyta).  
521 *Eur J Phycol* **38**: 325–339.
- 522 Weidner, S., Arnold, W., and Puhler, A. (1996) Diversity of uncultured microorganisms  
523 associated with the seagrass *Halophila stipulacea* estimated by restriction fragment length  
524 polymorphism analysis of PCR-amplified 16S rRNA genes. *Appl Environ Microbiol* **62**: 766–771.
- 525 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., et al. (2019)

- 526 Welcome to the tidyverse. *J Open Source Softw* **4**: 1686.
- 527 Wilke, C.O. (2018) cowplot: Streamlined plot theme and plot annotations for 'ggplot2'.
- 528 Williams, T.J., Long, E., Evans, F., DeMaere, M.Z., Lauro, F.M., Raftery, M.J., et al. (2012) A  
529 metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal  
530 surface waters. *ISME J* **6**: 1883–1900.
- 531 Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample  
532 preparation method for proteome analysis. *Nat Methods* **6**: 359–362.
- 533 Xie, Y. (2015) Dynamic Documents with R and knitr, 2nd ed. Boca Raton, Florida: Chapman  
534 and Hall/CRC.
- 535 Xie, Y. (2014) knitr: A comprehensive tool for reproducible research in R. In *Implementing  
536 Reproducible Computational Research*. Stodden, V., Leisch, F., and Peng, R.D. (eds). New York:  
537 Chapman and Hall/CRC, pp. 3–32.
- 538 Xie, Y. (2019) TinyTeX: A lightweight, cross-platform, and easy-to-maintain LaTeX  
539 distribution based on TeX Live. *TUGboat* **40**: 30–32.
- 540 Xie, Y. (2020) tinytex: Helper functions to install and maintain 'TeX Live', and compile  
541 'LaTeX' documents.
- 542 Xie, Y., Allaire, J.J., and Grolemund, G. (2018) R Markdown: The Definitive Guide, 1st ed.  
543 Boca Raton, Florida: Chapman and Hall/CRC.
- 544 Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014) The  
545 SILVA and "All-Species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res* **42**:  
546 D643–D648.
- 547 Zhu, H. (2019) kableExtra: Construct complex table with 'kable' and pipe syntax.

548 **Figure Captions**

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

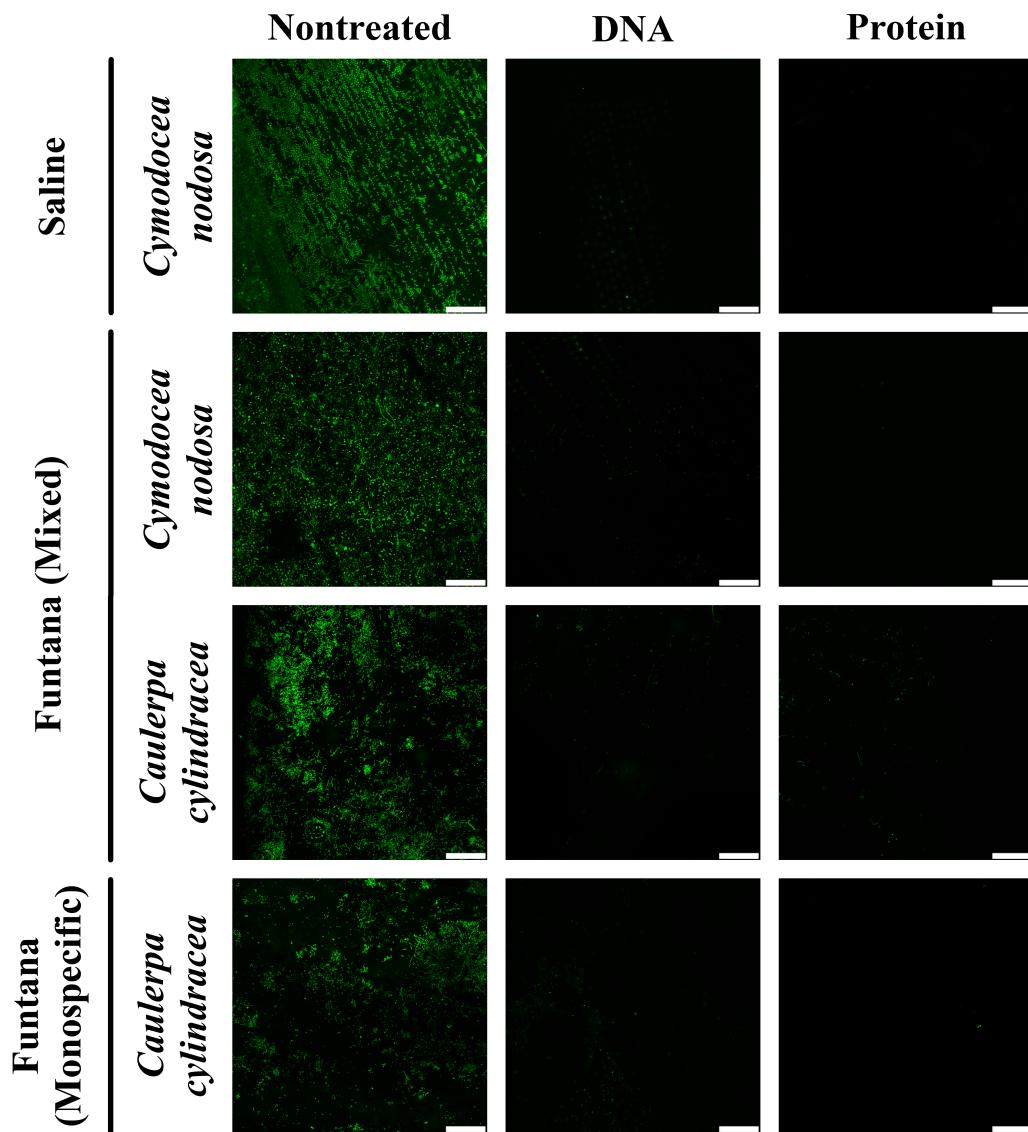
552 **Fig. 2.** Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of  
553 Funtana (Mixed and Monospecific Settlements) sampled on 19 June 2018 and stained with SYBR  
554 Green I. Scale bar in all images is 60 µm.

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

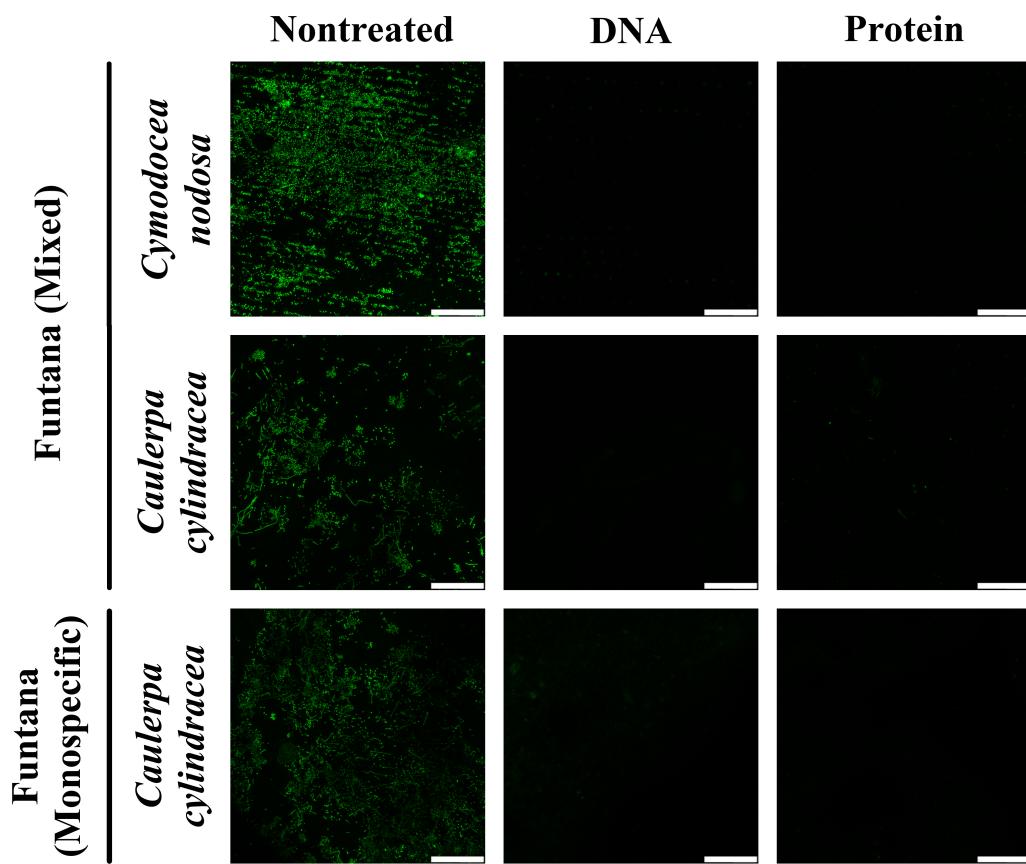
559 **Fig. 4.** Taxonomic classification and relative contribution of chloroplast sequences from surfaces  
560 of two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline and the  
561 Bay of Funtana (Mixed and Monospecific Settlements) and in two contrasting seasons (4 December  
562 2017 and 19 June 2018).

563 **Fig. 5.** Relative contribution of each COG category to the total number of annotated coding  
564 sequences (A) from metagenomes and identified proteins (B) from metaproteomes associated with  
565 surfaces of two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline  
566 and the Bay of Funtana (Mixed and Monospecific Settlements) and in two contrasting seasons (4/14  
567 December 2017 and 19 June 2018). Total number of annotated coding sequences and identified  
568 proteins is given above the corresponding bar.

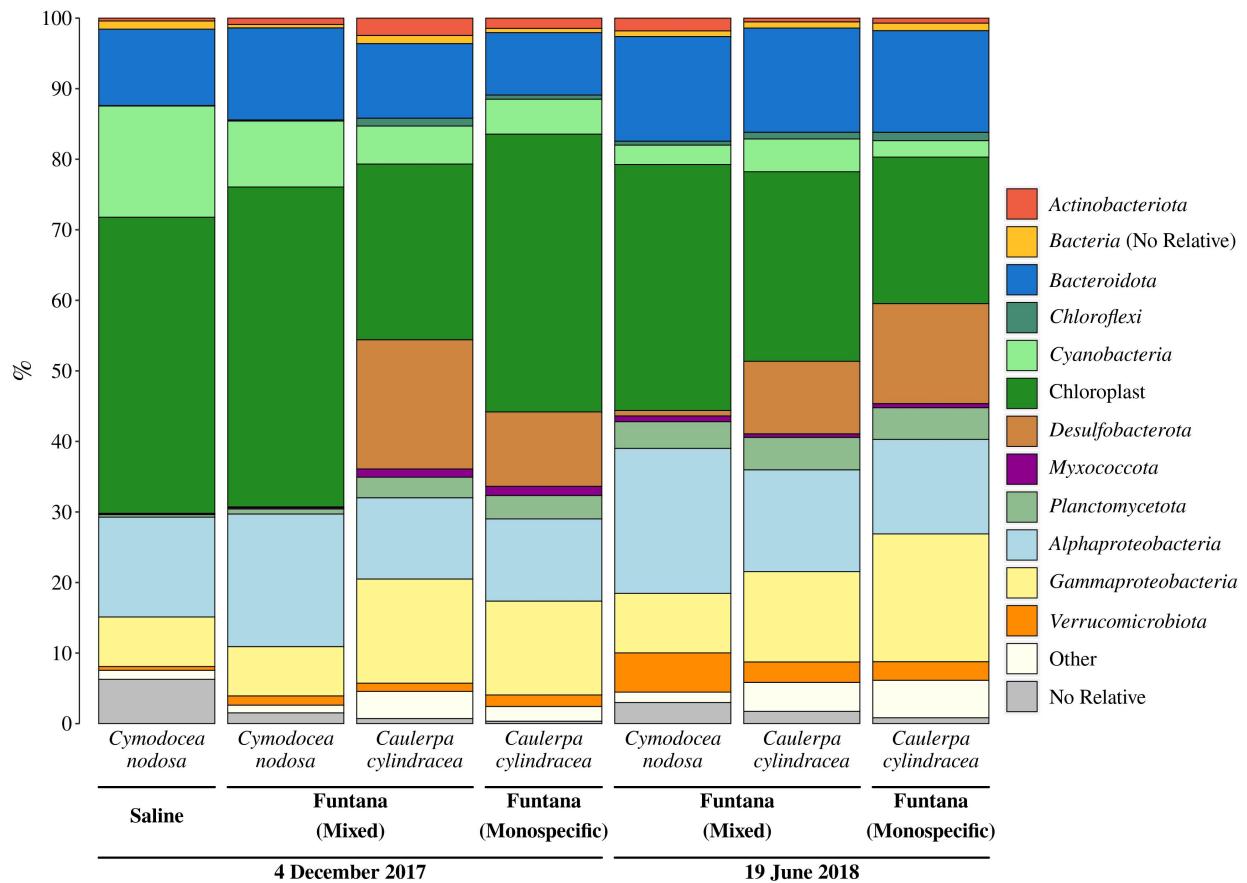
569 **Figures**



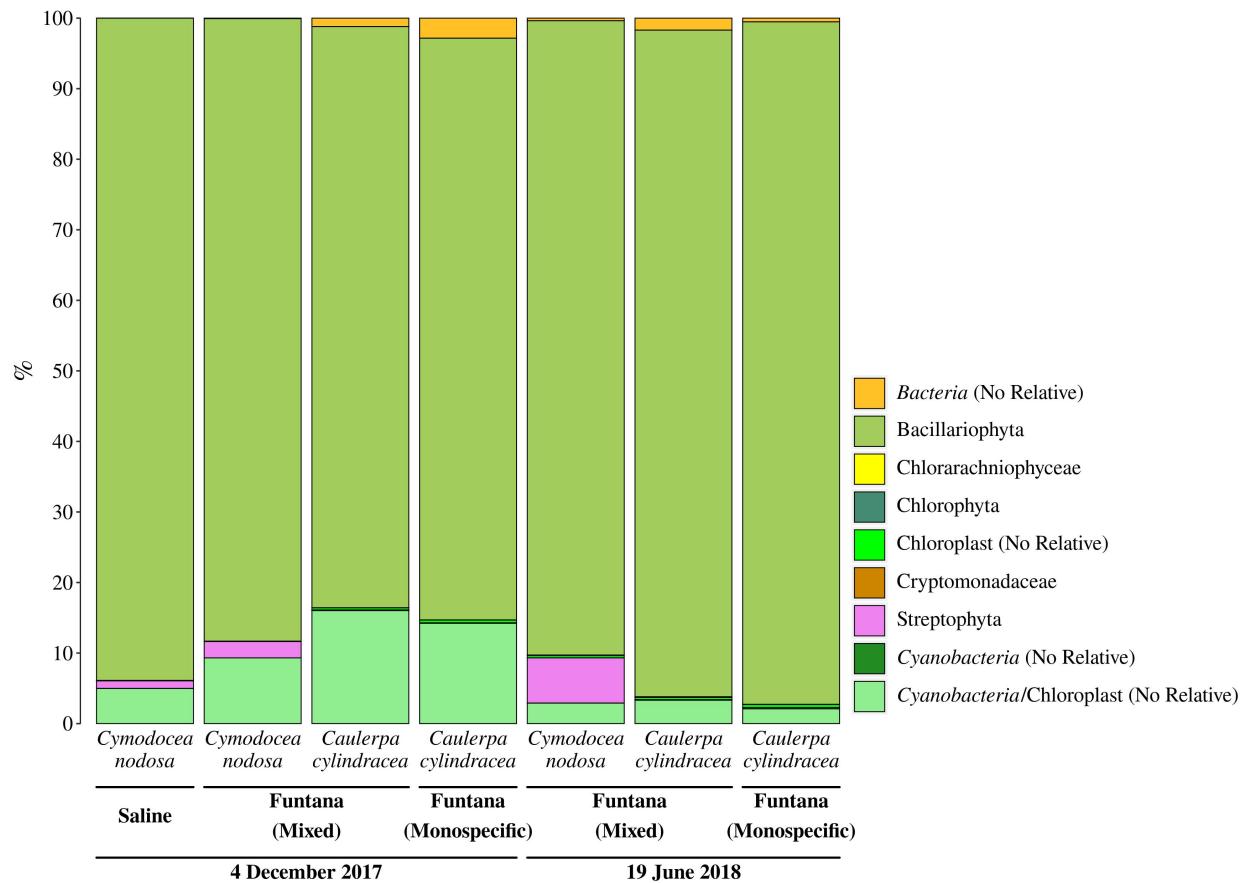
**Fig. 1.** Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of Saline and the Bay of Funtana (Mixed and Monospecific Settlements) sampled on 4 December 2017 and stained with SYBR Green I. Scale bar in all images is 60 µm.



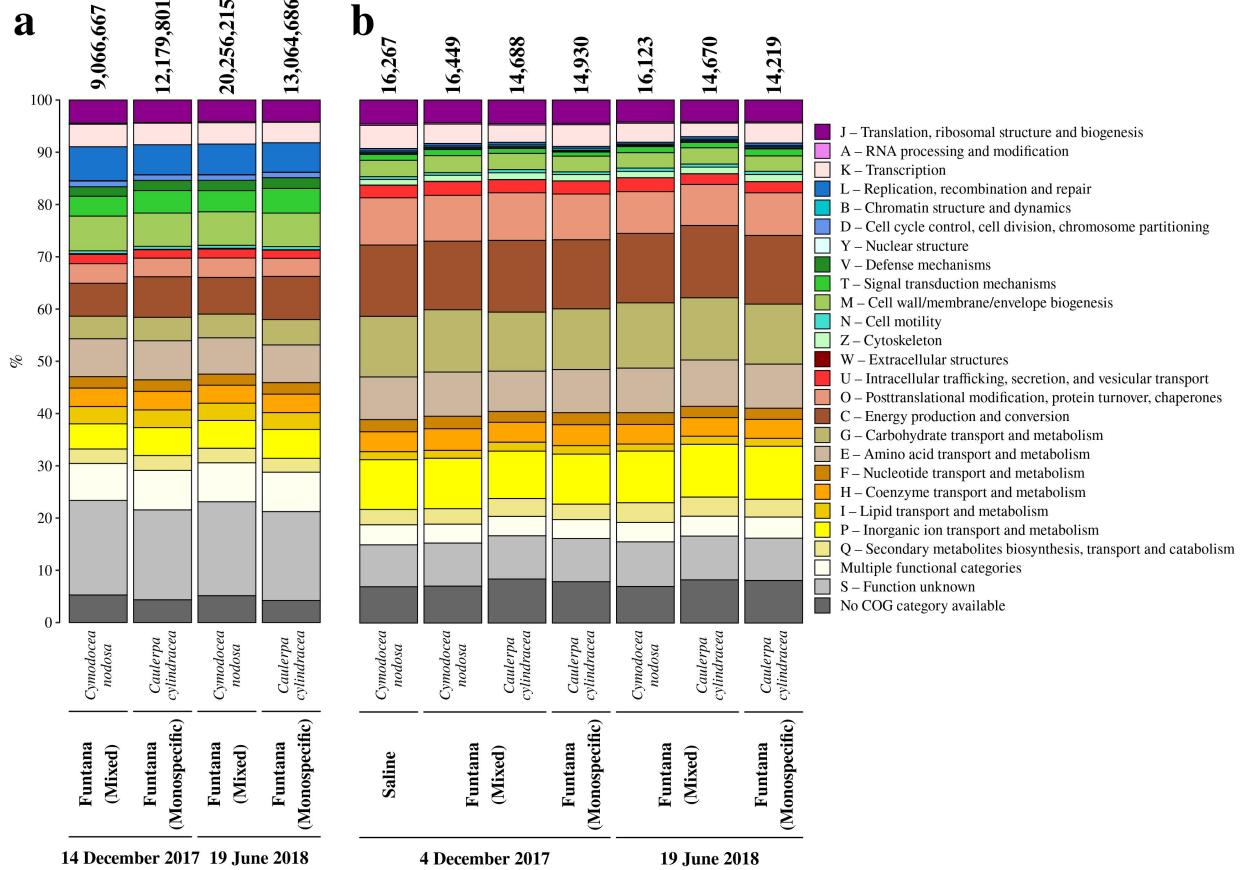
**Fig. 2.** Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of Funtana (Mixed and Monospecific Settlements) sampled on 19 June 2018 and stained with SYBR Green I. Scale bar in all images is 60  $\mu\text{m}$ .



**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).



**Fig. 5.** Relative contribution of each COG category to the total number of annotated coding sequences (A) from metagenomes and identified proteins (B) from metaproteomes associated with 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 (4/14 December 2017 and 19 June 2018). Total number of annotated coding sequences and identified proteins is given above the corresponding bar.