

Selective DNA and Protein Isolation from Marine Macrophyte Surfaces

Marino Korlević^{1*}, Marsej Markovski¹, Zihao Zhao², Gerhard J. Herndl^{2,3,4}, Mirjana Najdek¹

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

4. Vienna Metabolomics Center, University of Vienna, Austria

*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 such as 16S rRNA sequencing,
3 metagenomics and metaproteomics. To apply these techniques to the microbial community
4 inhabiting the surfaces of marine macrophytes it is advisable to perform a selective DNA and
5 protein isolation prior to the analysis to avoid biases due to the host material being present in high
6 quantities. Two protocols for DNA and protein isolation were adapted for selective extractions of
7 DNA and proteins from epiphytic communities inhabiting the surfaces of two marine macrophytes,
8 the seagrass *Cymodocea nodosa* and the macroalga *Caulerpa cylindracea*. Protocols showed an
9 almost complete removal of the epiphytic community regardless of the sampling season, station,
10 settlement or host species. The obtained DNA was suitable for metagenomic and 16S rRNA
11 sequencing, while isolated proteins could be identified by mass spectrometry. Low presence of host
12 DNA and proteins in the samples indicated a high specificity of the protocols. The procedures are
13 based on universally available laboratory chemicals making the protocols widely applicable. Taken
14 together, the adapted protocols ensure an almost complete removal of the macrophyte epiphytic
15 community. The procedures are selective for microbes inhabiting macrophyte surfaces and provide
16 DNA and proteins applicable in 16S rRNA sequencing, metagenomics and metaproteomics.

17 **Introduction**

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

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

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

46 An alternative to these procedures is a direct isolation of the target material by incubating
47 macrophyte tissues in an extraction buffer. After the incubation, the undisrupted host tissue is
48 removed followed by the isolation procedure, omitting host material contaminations. To our
49 knowledge, the only procedure describing a direct and selective epiphytic DNA isolation from the
50 surfaces of marine macrophytes was described by Burke *et al.* (2009). In contrast to previously
51 described methods, this protocol enables an almost complete removal of the surface community.
52 It was used for 16S rRNA gene clone library construction (Burke and Thomas *et al.*, 2011) and
53 metagenome sequencing (Burke and Peter Steinberg *et al.*, 2011). This method, although providing a
54 selective isolation procedure, uses a rapid multi-enzyme cleaner (3M) that is not available worldwide
55 and the chemical constituents are unknown (Burke *et al.*, 2009). Also to our knowledge, no selective
56 isolation protocol to perform (meta)proteomics of epiphytic communities associated with marine
57 macrophytes has been developed yet.

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

65 **Results**

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

73 To evaluate whether the obtained DNA is suitable to determine the composition of the microbial
74 community Illumina sequencing of the V4 region of the 16S rRNA was performed. Sequencing
75 yielded a total of 336,944 sequences after quality curation and exclusion of eukaryotic, mitochondrial
76 and no relative sequences. The number of sequences classified as chloroplasts was 97,334. After
77 excluding these sequences, the total number of retrieved reads was 239,610, ranging from 22,596
78 to 52,930 sequences per sample (Table S1). Even when the highest sequencing effort was applied,
79 the rarefaction curves did not level off which is commonly observed in high-throughput 16S rRNA
80 amplicon sequencing (Fig. S1). Sequences clustering at a similarity level of 97 % yielded a total of
81 8,355 different OTUs. Taxonomic classification of reads revealed a macrophyte-associated epiphytic
82 community that mainly composed of *Alphaproteobacteria* ($14.9 \pm 3.5\%$), *Bacteroidota* ($12.5 \pm 2.4\%$),
83 *Gammaproteobacteria* ($11.6 \pm 4.3\%$), *Desulfobacterota* ($7.8 \pm 7.5\%$), *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 analysed
87 samples ($33.4 \pm 9.4\%$) (Fig. 3). To determine whether chloroplast sequences originate from
88 the host 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

90 a more detailed chloroplast classification. The largest proportion of sequences was classified as
91 Bacillariophyta ($89.7 \pm 5.7\%$) indicating that the DNA removal procedure resulted in only minor
92 co-extracted quantities of host DNA (Fig. 4). Chloroplast sequences classified as Streptophyta
93 constituted $3.3 \pm 2.8\%$ of all chloroplast sequences originating from *C. nodosa* samples, while
94 sequences classified as Chlorophyta comprised only $0.02 \pm 0.01\%$ of all chloroplast sequences
95 associated with *C. cylindracea* samples.

96 To determine whether the extracted DNA can be used for metagenomic sequencing, four
97 samples containing epiphytic DNA were selected and shotgun sequenced using an Illumina platform.
98 Metagenomic sequencing yielded between 207,149,524 and 624,029,930 sequence pairs (Table S2).
99 Obtained sequences were successfully assembled into contigs whose L50 ranged from 654 to 1,011
100 bp. In addition, predicted coding sequences were functionally annotated (9,066,667 – 20,256,215
101 annotated sequences; Fig. 5a) and taxonomically classified. Functional annotation allowed for an
102 assessment of the relative contribution of each COG (Clusters of Orthologous Groups) functional
103 category to the total number of annotated coding sequences (Fig. 5a). Functional categories
104 containing the highest number of sequences were C (Energy production and conversion), E (Amino
105 acid transport and metabolism), M (Cell wall/membrane/envelope biogenesis), L (Replication,
106 recombination and repair) and P (Inorganic ion transport and metabolism). If host DNA is
107 co-extracted with epiphytes it should be detected in large proportions in sequenced metagenomes.
108 However, no large proportions of coding sequences classified as Streptophyta and Chlorophyta
109 were detected (Table S3). Sequenced metagenomic DNA originating from the surface of *C. nodosa*
110 contained 1.3 % of coding sequences classified as Streptophyta in December 2017 and 0.7 % in June
111 2018. Furthermore, the summed RPKM (Reads Per Kilobase Million) of these sequences constituted
112 1.7 % of total RPKM of all successfully classified sequences in December 2017 and 1.1 % in June
113 2018. Similar low proportions of host coding sequences were detected in metagenomic samples
114 originating from the surfaces of *C. cylindracea*. Of all successfully classified coding sequences 0.2
115 % were classified as Chlorophyta in December 2017 and 0.1 % in June 2018. A relatively higher
116 proportion of RPKM of these sequences than in the case of *C. nodosa* was observed, indicating

¹¹⁷ a higher co-extraction of host DNA in *C. cylindracea*. In December, the proportion of RPKM of
¹¹⁸ sequences classified as Chlorophyta was 8.2 %, while in June 2018 it reached 13.6 %.

¹¹⁹ To evaluate whether the procedure for protein extraction is suitable for metaproteomic analysis,
¹²⁰ obtained proteins were trypsin digested and sequenced using a mass spectrometer. Obtained
¹²¹ MS/MS spectra were searched against a protein database from sequenced metagenomes. From
¹²² 14,219 to 16,449 proteins were identified in isolated protein samples (Fig. 5b). In addition,
¹²³ successful identification of proteins allowed for an assessment of the relative contribution of
¹²⁴ each COG functional category to the total number of identified proteins (Fig. 5b). Functional
¹²⁵ categories containing the highest number of identified proteins were C (Energy production and
¹²⁶ conversion), G (Carbohydrate transport and metabolism), P (Inorganic ion transport and metabolism),
¹²⁷ O (Posttranslational modification, protein turnover, chaperones) and E (Amino acid transport and
¹²⁸ metabolism). Isolated proteins could originate from epiphytic organisms inhabiting the macrophyte
¹²⁹ surface and/or from macrophyte tissue underlying them. The contribution of proteins originating
¹³⁰ from host tissues was evaluated by identifying all the proteins predicted to belong to any taxonomic
¹³¹ group within the phyla Streptophyta and Chlorophyta and by calculating their contribution to the
¹³² number and abundance (NAAF – Normalized Abundance Area Factor) of all identified proteins.
¹³³ On average, proteins isolated from the surface of *C. nodosa* contained 1.8 ± 0.06 % of proteins
¹³⁴ associated with Streptophyta, contributing to 2.2 ± 0.8 % of total proteins. Similar to metagenomes,
¹³⁵ proteins associated with Chlorophyta contributed more to *C. cylindracea* than proteins associated
¹³⁶ with Streptophyta to *C. nodosa*. Chlorophyta associated proteins comprised 5.2 ± 0.06 % of all
¹³⁷ identified proteins in *C. cylindracea*, contributing 19.2 ± 1.5 % to the total protein abundance.

138 **Discussion**

139 To test whether the developed DNA and protein isolation protocols efficiently detach microbes
140 from the macrophyte surface we selected *C. nodosa* and *C. cylindracea* as representatives of seagrass
141 and macroalgal species. These species differ morphologically. While *C. nodosa* leaves are flat, *C.*
142 *cylindracea* thalli are characterized by an uneven surface (Kuo and den Hartog, 2001; Verlaque *et*
143 *al.*, 2003). The developed protocol led to an almost complete removal of epiphytic cells from the
144 surfaces of both species comparable to the result of Burke *et al.* (2009), indicating that structural
145 differences do not impact the removal efficiency. In addition, isolation protocols were tested in two
146 contrasting seasons, as it is known that macrophytes are harbouring more algal epiphytes during
147 autumn and winter (Reyes and Sansón, 2001). No differences in the removal efficiency was observed
148 between seasons suggesting that these protocols can be used on macrophyte samples retrieved
149 throughout the year. Also, no removal differences were observed on samples derived from the same
150 host but from different locations.

151 Successful amplification and sequencing of the V4 region of the 16S rRNA gene proved that the
152 isolated DNA can be used to estimate the microbial epiphytic diversity. Taxonomic groups detected
153 in this step can also be often found in epiphytic communities associated with other macrophytes
154 (Burke and Thomas *et al.*, 2011; Morrissey *et al.*, 2019). A problem often encountered in studies
155 focusing on epiphytic communities is the presence of large proportions of chloroplast 16S rRNA
156 sequences in the pool of amplified molecules, especially if the epiphytic DNA was isolated without
157 prior selection (Staufenberger *et al.*, 2008). These sequences can derive from host chloroplasts or
158 from eukaryotic epiphytic chloroplast DNA. Although the proportion of obtained chloroplast 16S
159 rRNA sequences in our samples was substantial, they derived almost exclusively from eukaryotic
160 epiphytes. High proportion of chloroplast 16S rRNA sequences in studies applying selective
161 procedures that include direct cellular lysis on host surfaces were observed before (Michelou *et*
162 *al.*, 2013). It is possible that chloroplast-specific sequences even in these studies originated from
163 eukaryotic epiphytic cells and not from host chloroplasts. Indeed, it is common during 16S rRNA

164 profiling of pelagic microbial communities to observe high proportions of chloroplast sequences
165 (Gilbert *et al.*, 2009; Korlević *et al.*, 2016). In addition, a very low proportion of chloroplast 16S
166 rRNA sequences in samples originating from *C. cylindracea* in comparison to *C. nodosa* could be
167 explained by the presence of three introns in the gene for 16S rRNA in some members of the genus
168 *Caulerpa* that could hamper the amplification process (Lam and Lopez-Bautista, 2016).

169 High quality DNA is also needed for metagenomics. The obtained number of metagenomic
170 sequences and assembly statistics were comparable to metagenomes and metatranscriptomes derived
171 from similar surface associated communities (Crump *et al.*, 2018; Cúcio *et al.*, 2018). In addition,
172 functional annotation of predicted coding sequences to COG functional categories showed that
173 the obtained metagenomes can be used to determine the metabolic capacity of surface associated
174 communities (Leary *et al.*, 2014; Cúcio *et al.*, 2018). The proportion of coding sequences, including
175 their RPKM, originating from *C. nodosa* metagenomes and classified as Streptophyta was low
176 indicating that the isolation procedure was specific for epiphytic cells. DNA samples isolated from
177 the surface of *C. cylindracea* exhibited a low proportion of Chlorophyta coding sequences, however,
178 their RPKM was higher than in the samples originating from *C. nodosa*. One of the reasons for this
179 elevated RPKM of Chlorophyta sequences in *C. cylindracea* could be the differences in the tissue
180 structure between these two host species. While *C. nodosa* leaves are composed of individual cells,
181 the thallus of *C. cylindracea* is, like in other siphonous algal species, composed of a single large
182 multinucleate cell (Coneva and Chitwood, 2015). The absence of individual cells in *C. cylindracea*
183 could cause a leakage of genetic material into the extraction buffer causing an elevated presence of
184 host sequences in the samples for metagenome analyses.

185 To obtain insight into the metabolic status of uncultivated prokaryotes, a metaproteomic
186 approach is required (Saito *et al.*, 2019). The applied protocol for epiphytic protein isolation
187 followed by a metaproteomic analysis identified between 14,219 and 16,449 proteins, which is
188 higher than previously reported for e.g. soils (Chourey *et al.*, 2010; Hultman *et al.*, 2015), seawater
189 (Williams *et al.*, 2012) and biofilms (Leary *et al.*, 2014). The functional annotation of identified

190 proteins into COG functional categories showed that the protein isolation protocol can be used to
191 assess the metabolic status of the epiphytic community (Leary *et al.*, 2014). Similar to the results of
192 the metagenomic analysis, the number and abundance of identified proteins affiliated to Streptophyta
193 in *C. nodosa* samples were low, indicating that the procedure is selective for epiphytic cell proteins.
194 In addition, a higher number and abundance of identified proteins associated with Chlorophyta were
195 observed in *C. cylindracea* samples. The cause of this elevated presence of Chlorophyta-associated
196 proteins can be, similar to the DNA isolation protocol, explained by the absence of individual cells
197 in this siphonous alga (Coneva and Chitwood, 2015).

198 In conclusion, the developed protocols for DNA and protein isolation from macrophyte surfaces
199 almost completely remove the epiphytic community from both, *C. nodosa* and *C. cylindracea*, in
200 different seasons. Also, the obtained DNA and proteins are suitable for 16S rRNA sequencing,
201 metagenomics and metaproteomics analyses while the obtained material contains low quantities of
202 host DNA and proteins making the protocols specific for epiphytes. Furthermore, the protocols are
203 based on universally available laboratory chemicals hence, making them widely applicable.

204 **Experimental procedures**

205 **Sampling**

206 Leaves of *C. nodosa* were sampled in a *C. nodosa* meadow in the Bay of Saline, northern
207 Adriatic Sea ($45^{\circ}7'5''$ N, $13^{\circ}37'20''$ E) and in a *C. nodosa* meadow invaded by *C. cylindracea* in
208 the Bay of Funtana, northern Adriatic Sea ($45^{\circ}10'39''$ N, $13^{\circ}35'42''$ E). Thalli of *C. cylindracea*
209 were sampled in the same *C. nodosa* invaded meadow in the Bay of Funtana and at a locality of
210 only *C. cylindracea* located in the proximity of the invaded meadow. Leaves and thalli for 16S
211 rRNA analysis, metagenomics and metaproteomics were collected in two contrasting seasons, on 4
212 December 2017 (16S rRNA analysis and metaproteomics), 14 December 2017 (metagenomics) and
213 18 June 2018 (16S rRNA analysis, metagenomics and metaproteomics). During spring 2018, the *C.*
214 *nodosa* meadow in the Bay of Saline decayed to an extent that no leaves could be retrieved (Najdek
215 *et al.*, 2020). In addition, as not enough DNA for both metagenomic and 16S RNA analysis were
216 obtained during the sampling on 4 December 2017, an additional sampling on 14 December 2017
217 was carried out in the Bay of Funtana. Leaves and thalli were collected by diving and transported to
218 the laboratory in containers placed on ice and filled with seawater from this site. Upon arrival to the
219 laboratory, *C. nodosa* leaves were cut into sections of 1 – 2 cm, while *C. cylindracea* thalli were
220 cut into 5 – 8 cm long sections. Leaves and thalli were washed three times with sterile artificial
221 seawater (ASW) to remove loosely attached microbial cells.

222 **DNA isolation**

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

227 duplicate isolations were performed. Lysozyme was added (final concentration 1 mg ml⁻¹) and the
228 mixture was incubated at 37 °C for 30 min. Subsequently, proteinase K (final concentration 0.5
229 mg ml⁻¹) and SDS (final concentration 1 %) were added and the samples were incubated at 55 °C
230 for 2 h. Following the incubation, tubes were vortexed for 10 min and the mixture containing lysed
231 epiphytic cells was separated from host leaves or thalli by transferring the solution into a clean tube.
232 The lysate was extracted twice with a mixture of phenol:chloroform:isoamyl alcohol (25:24:1; pH 8)
233 and once with chloroform:isoamyl alcohol (24:1). After each addition of an organic solvent mixture,
234 tubes were slightly vortexed and centrifuged at 4,500 × g for 10 min. Following each centrifugation
235 the aqueous phases were retrieved. After the final extraction 1/10 of chilled 3 M sodium acetate
236 (pH 5.2) was added. DNA was precipitated by adding 1 volume of chilled isopropanol, incubating
237 the mixtures overnight at -20 °C and centrifuging at 16,000 × g and 4 °C for 20 min. The pellet
238 was washed twice with 1 ml of chilled 70 % ethanol and centrifuged after each washing step at
239 20,000 × g and 4 °C for 10 min. After the first washing step duplicate pellets from the same sample
240 were pooled and transferred to a clean 1.5 ml tube. The dried pellet was re-suspended in 100 µl of
241 deionized water.

242 Illumina 16S rRNA sequencing

243 An aliquot of isolated DNA was treated with RNase A (final concentration 200 µg ml⁻¹)
244 for 2 h at 37 °C. The DNA concentration was determined using the Quant-iT PicoGreen
245 dsDNA Assay Kit (Invitrogen, USA) according to the manufacturer's instructions and
246 diluted to 1 ng µl⁻¹. The V4 region of the 16S rRNA gene was amplified using a two-step
247 PCR procedure. In the first PCR the 515F (5'-GTGYCAGCMGCCGCGGTAA-3') and
248 806R (5'-GGACTACNVGGGTWTCTAAT-3') primers from the Earth Microbiome Project
249 (<http://press.igsb.anl.gov/earthmicrobiome/protocols-and-standards/16s/>) were used to amplify
250 the target region (Caporaso *et al.*, 2012; Apprill *et al.*, 2015; Parada *et al.*, 2016). These primers
251 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⁻¹ 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 analysed 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) (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 mothur and the RDP (<http://rdp.cme.msu.edu/>) training set (version 16) reference files adapted for

279 mothur (Cole *et al.*, 2014). In comparison to SILVA, RDP allows a more detailed classification of
280 chloroplast sequences. Based on the ATCC MSA-1002 mock community included in the analysis a
281 sequencing error rate of 0.009 % was determined, which is in line with previously reported values
282 for next-generation sequencing data (Kozich *et al.*, 2013; Schloss *et al.*, 2016). In addition, the
283 negative control processed together with the samples yielded only 2 sequences after sequence
284 quality curation.

285 **Metagenomics**

286 Four DNA samples were selected and sent on dry ice to IMGM Laboratories, Martinsried,
287 Germany for metagenomic sequencing. DNA was purified using AMPure XP Beads (Beckman
288 Coulter, USA) applying a bead:DNA ratio of 1:1 (v/v), quantified with a Qubit dsDNA HS Assay
289 Kit (Thermo Fisher Scientific, USA) and check for integrity on a 1 % agarose gel. Metagenomic
290 sequencing libraries were generated from 300 ng of input DNA using a NEBNext Ultra II FS
291 DNA Library Prep Kit for Illumina (New England Biolabs, USA) according to the manufacturer's
292 instructions. Fragments were selected (500 – 700 bp) using AMPure XP Beads, PCR enriched for 3
293 – 5 cycles and quality controlled. Libraries generated from different DNA samples were pooled and
294 sequenced on an Illumina NovaSeq 6000 sequencing system (2 × 250 bp).

295 Obtained sequences were analysed on the Life Science Compute Cluster (LiSC) (CUBE –
296 Computational Systems Biology, University of Vienna). Individual sequences were assembled using
297 MEGAHIT (version 1.1.2) (Li *et al.*, 2015) under default settings. Putative genes were predicted
298 from contigs longer than 200 bp using Prodigal (version 2.6.3) (Hyatt *et al.*, 2010) in metagenome
299 mode (-p meta). Abundances of predicted genes were expressed as Reads Per Kilobase Million
300 (RPKM) and calculated using the BWA algorithm (version 0.7.16a) (Li and Durbin, 2009). All
301 predicted genes were functionally annotated using the eggNOG-mapper (Huerta-Cepas *et al.*, 2017)
302 and eggNOG database (version 5.0) (Huerta-Cepas *et al.*, 2019). Sequence taxonomy classification
303 was determined using the lowest common ancestor algorithm adapted from DIAMOND (version

304 0.8.36) (Buchfink *et al.*, 2015) and by searching against the NCBI non-redundant database (NR). To
305 determine the phylogeny, the top 10 % hits with an e-value $< 1 \times 10^{-5}$ were used (--top 10). Sequence
306 renaming, coverage information computing and metagenomic statistics calculations were performed
307 using software tools from BBTools (<https://jgi.doe.gov/data-and-tools/bbtools>). Metagenomic
308 sequences obtained in this study have been deposited in the European Nucleotide Archive (ENA) at
309 EMBL-EBI under accession numbers SAMEA6648795, SAMEA6648797, SAMEA6648809 and
310 SAMEA6648811.

311 Protein isolation

312 Proteins were isolated according to the protocol for protein isolation from soil described in
313 Chourey *et al.* (2010) and modified by Hultman *et al.* (2015). This protocol was further modified
314 and adapted for microbial protein isolation from macrophyte surfaces as described below. Twenty
315 (20) ml of protein extraction buffer (4 % SDS, 100 mm Tris-HCl [pH 8.0]) was added to 5 g wet
316 weight of leaves or 10 g wet weight of thalli. The mixture was incubated in boiling water for 5
317 min, vortexed for 10 min and incubated again in boiling water for 5 min. After a brief vortex, the
318 lysate was transferred to a clean tube separating the host leaves or thalli from the mixture containing
319 lysed epiphytic cells. Dithiothreitol (DTT; final concentration 24 mM) was added and proteins were
320 precipitated with chilled 100 % trichloroacetic acid (TCA; final concentration 20 %) overnight at
321 -20°C . Precipitated proteins were centrifuged at $10,000 \times g$ and 4°C for 40 min. The obtained
322 protein pellet was washed three times with chilled acetone. During the first washing step the pellet
323 was transferred to a clean 1.5 ml tube. After each washing step, samples were centrifuged at 20,000
324 $\times g$ and 4°C for 5 min. Dried pellets were stored at -80°C until further analysis.

325 **Metaproteomics**

326 Isolated proteins were whole trypsin digested using the FASP (filter-aided sample preparation)
327 Protein Digestion Kit (Expedeon, UK) according to the manufacturer's instructions with small
328 modifications (Wiśniewski *et al.*, 2009). Prior to loading the solution onto the column, protein
329 pellets were solubilized in a urea sample buffer included in the kit amended with DTT (final
330 concentration 100 mM) for 45 min at room temperature and centrifuged at 20,000 × g for 2 – 5
331 min at room temperature to remove larger particles. The first washing step after protein solution
332 loading was repeated twice. In addition, the centrifugation steps were prolonged if the column
333 was clogged. Trypsin digestion was performed on column filters at 37 °C overnight for 18 h.
334 The final filtrate containing peptides was acidified with 1 % (final concentration) trifluoroacetic
335 acid (TFA), freezed at –80 °C, lyophilized and sent to VIME – Vienna Metabolomics Center
336 (University of Vienna) for metaproteomic analysis. Peptides were re-suspended in 1 % (final
337 concentration) TFA, desalting using the Pierce C18 Tips (Thermo Fisher Scientific, USA) according
338 to the manufacturer's instructions and sequenced on a Q Exactive Hybrid Quadrupole-Orbitrap
339 Mass Spectrometer (Thermo Fisher Scientific, USA). Obtained MS/MS spectra were searched
340 against a protein database composed of combined sequenced metagenomes using SEQUEST-HT
341 engines and validated with Percolator in Proteome Discoverer 2.1 (Thermo Fisher Scientific, USA).
342 The target-decoy approach was used to reduce the probability of false peptide identification. Results
343 whose false discovery rate at the peptide level was < 1 % were kept. For protein identification
344 a minimum of two peptides and one unique peptide were required. For protein quantification, a
345 chromatographic peak area-based free quantitative method was applied.

346 **Data processing and visualization**

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

³⁴⁹ *al.*, 2019) and multiple other packages (Xie, 2014, 2015, 2019, 2020; Neuwirth, 2014; Xie *et*
³⁵⁰ *al.*, 2018; Wilke, 2018; Allaire *et al.*, 2019; Zhu, 2019; Bengtsson, 2020). The detailed analysis
³⁵¹ procedure including the R Markdown file for this paper are available as a GitHub repository
³⁵² (https://github.com/MicrobesRovinj/Korlevic_SelectiveRemoval_EnvironMicrobiol_2020).

³⁵³ **Confocal microscopy**

³⁴⁴ Host leaves and thalli from DNA and protein isolation steps were washed seven times in
³⁴⁵ deionized water and fixed with formaldehyde (final concentration ~3 %). In addition, non-treated
³⁴⁶ leaves and thalli, washed three times in ASW to remove loosely attached microbial cells, were fixed
³⁴⁷ in the same concentration of formaldehyde and used as a positive control. For long term storage,
³⁴⁸ fixed leaves and thalli were immersed in a mixture of phosphate-buffered saline (PBS) and ethanol
³⁴⁹ (1:1) and stored at -20 °C. Treated and untreated segments of leaves and thalli were stained in a 2
³⁵⁰ × solution of SYBR Green I and examined under a Leica TCS SP8 X FLIM confocal microscope
³⁵¹ (Leica Microsystems, Germany).

³⁶² **Acknowledgements**

³⁶³ This work was funded by the Croatian Science Foundation through the MICRO-SEAGRASS
³⁶⁴ project (project number IP-2016-06-7118). ZZ and GJH were supported by the Austrian Science
³⁶⁵ Fund (FWF) project ARTEMIS (project number P28781-B21). We would like to thank Margareta
³⁶⁶ Buterer for technical support, Paolo Paliaga for help during sampling, Lucija Horvat for technical
³⁶⁷ support in confocal microscopy, Dušica Vujaklija for help in accessing the confocal microscope
³⁶⁸ and Sonja Tischler for technical support in mass spectrometry. In addition, we would like to thank
³⁶⁹ the University Computing Center of the University of Zagreb for access to the computer cluster
³⁷⁰ Isabella and CUBE – Computational Systems Biology of the University of Vienna for access to the
³⁷¹ Life Science Compute Cluster (LiSC).

372 **References**

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

- 394 Cai, X., Gao, G., Yang, J., Tang, X., Dai, J., Chen, D., and Song, Y. (2014) An ultrasonic
395 method for separation of epiphytic microbes from freshwater submerged macrophytes. *J Basic*
396 *Microbiol* **54**: 758–761.
- 397 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., et al.
398 (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq
399 platforms. *ISME J* **6**: 1621–1624.
- 400 Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., et al.
401 (2010) Direct cellular lysis/protein extraction protocol for soil metaproteomics. *J Proteome Res* **9**:
402 6615–6622.
- 403 Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., et al. (2014) Ribosomal
404 Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res* **42**:
405 D633–D642.
- 406 Coneva, V. and Chitwood, D.H. (2015) Plant architecture without multicellularity: Quandaries
407 over patterning and the soma-germline divide in siphonous algae. *Front Plant Sci* **6**: 287.
- 408 Crump, B.C., Wojahn, J.M., Tomas, F., and Mueller, R.S. (2018) Metatranscriptomics and
409 amplicon sequencing reveal mutualisms in seagrass microbiomes. *Front Microbiol* **9**: 388.
- 410 Cúcio, C., Overmars, L., Engelen, A.H., and Muyzer, G. (2018) Metagenomic analysis shows
411 the presence of bacteria related to free-living forms of sulfur-oxidizing chemolithoautotrophic
412 symbionts in the rhizosphere of the seagrass *Zostera marina*. *Front Mar Sci* **5**: 171.
- 413 Egan, S., Harder, T., Burke, C., Steinberg, P., Kjelleberg, S., and Thomas, T. (2013) The
414 seaweed holobiont: Understanding seaweed-bacteria interactions. *FEMS Microbiol Rev* **37**:
415 462–476.

416 Gilbert, J.A., Field, D., Swift, P., Newbold, L., Oliver, A., Smyth, T., et al. (2009) The
417 seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol* **11**:
418 3132–3139.

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

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

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

427 Hultman, J., Waldrop, M.P., Mackelprang, R., David, M.M., McFarland, J., Blazewicz, S.J., et
428 al. (2015) Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature*
429 **521**: 208–212.

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

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

436 Korlević, M., Šupraha, L., Ljubešić, Z., Henderiks, J., Ciglenečki, I., Dautović, J., and Orlić, S.
437 (2016) Bacterial diversity across a highly stratified ecosystem: A salt-wedge Mediterranean estuary.
438 *Syst Appl Microbiol* **39**: 398–408.

- 439 Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013)
- 440 Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon
- 441 sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* **79**: 5112–5120.
- 442 Kuo, J. and den Hartog, C. (2001) Seagrass taxonomy and identification key. In *Global*
- 443 *Seagrass Research Methods*. Short, F.T. and Coles, R.G. (eds). Amsterdam: Elsevier Science B.V.,
- 444 pp. 31–58.
- 445 Lam, D.W. and Lopez-Bautista, J.M. (2016) Complete chloroplast genome for *Caulerpa*
- 446 *racemosa* (Bryopsidales, Chlorophyta) and comparative analyses of siphonous green seaweed
- 447 plastomes. *Cymbella* **2**: 23–32.
- 448 Leary, D.H., Li, R.W., Hamdan, L.J., IV, W.J.H., Lebedev, N., Wang, Z., et al. (2014)
- 449 Integrated metagenomic and metaproteomic analyses of marine biofilm communities. *Biofouling*
- 450 **30**: 1211–1223.
- 451 Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015) MEGAHIT: An ultra-fast
- 452 single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph.
- 453 *Bioinformatics* **31**: 1674–1676.
- 454 Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows–Wheeler
- 455 transform. *Bioinformatics* **25**: 1754–1760.
- 456 Longford, S., Tujula, N., Crocetti, G., Holmes, A., Holmström, C., Kjelleberg, S., et al. (2007)
- 457 Comparisons of diversity of bacterial communities associated with three sessile marine eukaryotes.
- 458 *Aquat Microb Ecol* **48**: 217–229.
- 459 Massana, R., Murray, A.E., Preston, C.M., and DeLong, E.F. (1997) Vertical distribution and
- 460 phylogenetic characterization of marine planktonic *Archaea* in the Santa Barbara Channel. *Appl*
- 461 *Environ Microbiol* **63**: 50–56.

- 462 Michelou, V.K., Caporaso, J.G., Knight, R., and Palumbi, S.R. (2013) The ecology of microbial
463 communities associated with *Macrocystis pyrifera*. *PLoS One* **8**: e67480.
- 464 Morrissey, K.L., Çavas, L., Willems, A., and De Clerck, O. (2019) Disentangling the influence
465 of environment, host specificity and thallus differentiation on bacterial communities in siphonous
466 green seaweeds. *Front Microbiol* **10**: 717.
- 467 Najdek, M., Korlević, M., Paliaga, P., Markovski, M., Ivančić, I., Iveša, L., et al. (2020)
468 Dynamics of environmental conditions during the decline of a *Cymodocea nodosa* meadow.
469 *Biogeosciences* **17**: 3299–3315.
- 470 Neuwirth, E. (2014) RColorBrewer: ColorBrewer palettes.
- 471 Nõges, T., Luup, H., and Feldmann, T. (2010) Primary production of aquatic macrophytes and
472 their epiphytes in two shallow lakes (Peipsi and Võrtsjärv) in Estonia. *Aquat Ecol* **44**: 83–92.
- 473 Parada, A.E., Needham, D.M., and Fuhrman, J.A. (2016) Every base matters: Assessing small
474 subunit rRNA primers for marine microbiomes with mock communities, time series and global field
475 samples. *Environ Microbiol* **18**: 1403–1414.
- 476 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA
477 ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic
478 Acids Res* **41**: D590–D596.
- 479 R Core Team (2019) R: A language and environment for statistical computing, Vienna, Austria:
480 R Foundation for Statistical Computing.
- 481 Reyes, J. and Sansón, M. (2001) Biomass and production of the epiphytes on the leaves of
482 *Cymodocea nodosa* in the Canary Islands. *Bot Mar* **44**: 307–313.

- 483 Richter-Heitmann, T., Eickhorst, T., Knauth, S., Friedrich, M.W., and Schmidt, H. (2016)
- 484 Evaluation of strategies to separate root-associated microbial communities: A crucial choice in
- 485 rhizobiome research. *Front Microbiol* **7**: 773.
- 486 Saito, M.A., Bertrand, E.M., Duffy, M.E., Gaylord, D.A., Held, N.A., Hervey, W.J., et al.
- 487 (2019) Progress and challenges in ocean metaproteomics and proposed best practices for data
- 488 sharing. *J Proteome Res* **18**: 1461–1476.
- 489 Schloss, P.D., Jenior, M.L., Koumpouras, C.C., Westcott, S.L., and Highlander, S.K. (2016)
- 490 Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. *PeerJ* **4**:
- 491 e1869.
- 492 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al.
- 493 (2009) Introducing mothur: Open-source, platform-independent, community-supported software for
- 494 describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.
- 495 Staufenberger, T., Thiel, V., Wiese, J., and Imhoff, J.F. (2008) Phylogenetic analysis of bacteria
- 496 associated with *Laminaria saccharina*. *FEMS Microbiol Ecol* **64**: 65–77.
- 497 Su, C., Lei, L., Duan, Y., Zhang, K.-Q., and Yang, J. (2012) Culture-independent methods for
- 498 studying environmental microorganisms: Methods, application, and perspective. *Appl Microbiol*
- 499 *Biotechnol* **93**: 993–1003.
- 500 Uku, J., Björk, M., Bergman, B., and Díez, B. (2007) Characterization and comparison of
- 501 prokaryotic epiphytes associated with three East African seagrasses. *J Phycol* **43**: 768–779.
- 502 Verlaque, M., Durand, C., Huisman, J.M., Boudouresque, C.-F., and Le Parco, Y. (2003) On the
- 503 identity and origin of the Mediterranean invasive *Caulerpa racemosa* (Caulerpales, Chlorophyta).
- 504 *Eur J Phycol* **38**: 325–339.

- 505 Weidner, S., Arnold, W., and Puhler, A. (1996) Diversity of uncultured microorganisms
506 associated with the seagrass *Halophila stipulacea* estimated by restriction fragment length
507 polymorphism analysis of PCR-amplified 16S rRNA genes. *Appl Environ Microbiol* **62**: 766–771.
- 508 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., et al. (2019)
509 Welcome to the tidyverse. *J Open Source Softw* **4**: 1686.
- 510 Wilke, C.O. (2018) cowplot: Streamlined plot theme and plot annotations for 'ggplot2'.
- 511 Williams, T.J., Long, E., Evans, F., DeMaere, M.Z., Lauro, F.M., Raftery, M.J., et al. (2012) A
512 metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal
513 surface waters. *ISME J* **6**: 1883–1900.
- 514 Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample
515 preparation method for proteome analysis. *Nat Methods* **6**: 359–362.
- 516 Xie, Y. (2015) Dynamic Documents with R and knitr, 2nd ed. Boca Raton, Florida: Chapman
517 and Hall/CRC.
- 518 Xie, Y. (2014) knitr: A comprehensive tool for reproducible research in R. In *Implementing
519 Reproducible Computational Research*. Stodden, V., Leisch, F., and Peng, R.D. (eds). New York:
520 Chapman and Hall/CRC, pp. 3–32.
- 521 Xie, Y. (2019) TinyTeX: A lightweight, cross-platform, and easy-to-maintain LaTeX
522 distribution based on TeX Live. *TUGboat* **40**: 30–32.
- 523 Xie, Y. (2020) tinytex: Helper functions to install and maintain 'TeX Live', and compile
524 'LaTeX' documents.
- 525 Xie, Y., Allaire, J.J., and Grolemund, G. (2018) R Markdown: The Definitive Guide, 1st ed.
526 Boca Raton, Florida: Chapman and Hall/CRC.

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

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

531 **Figure captions**

532 **Fig. 1.** Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of
533 Saline and the Bay of Funtana (mixed and monospecific settlements) sampled on 4 December 2017
534 and stained with SYBR Green I. Scale bar at all images is 60 µm.

535 **Fig. 2.** Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of
536 Funtana (mixed and monospecific settlements) sampled on 19 June 2018 and stained with SYBR
537 Green I. Scale bar at all images is 60 µm.

538 **Fig. 3.** Taxonomic classification and relative contribution of the most abundant ($\geq 1\%$) bacterial
539 and archaeal sequences from surfaces of two marine macrophytes (*C. nodosa* and *C. cylindracea*)
540 sampled in the Bay of Saline and the Bay of Funtana (mixed and monospecific settlements) and in
541 two contrasting seasons (4 December 2017 and 19 June 2018).

542 **Fig. 4.** Taxonomic classification and relative contribution of chloroplast sequences from surfaces of
543 two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline and the Bay
544 of Funtana (mixed and monospecific settlements) and in two contrasting seasons (4 December 2017
545 and 19 June 2018).

546 **Fig. 5.** Relative contribution of each COG category to the total number of annotated coding
547 sequences (A) from metagenomes and identified proteins (B) from metaproteomes associated with
548 surfaces of two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline
549 and the Bay of Funtana (mixed and monospecific settlements) and in two contrasting seasons (4/14
550 December 2017 and 19 June 2018). Total number of annotated coding sequences and identified
551 proteins is given above the corresponding bar.

552 **Figures**

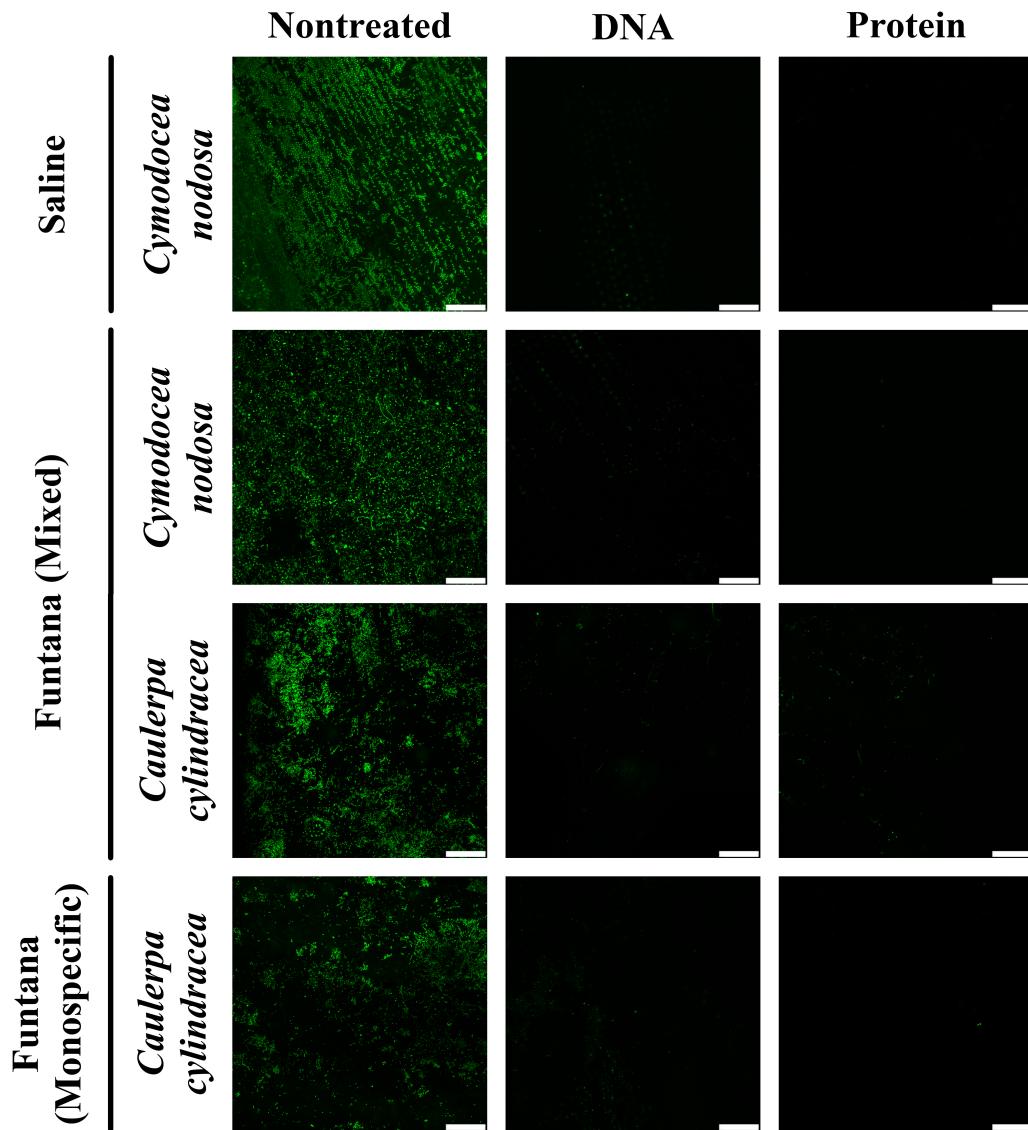


Fig. 1. Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of Saline and the Bay of Funtana (mixed and monospecific settlements) sampled on 4 December 2017 and stained with SYBR Green I. Scale bar at all images is 60 µm.

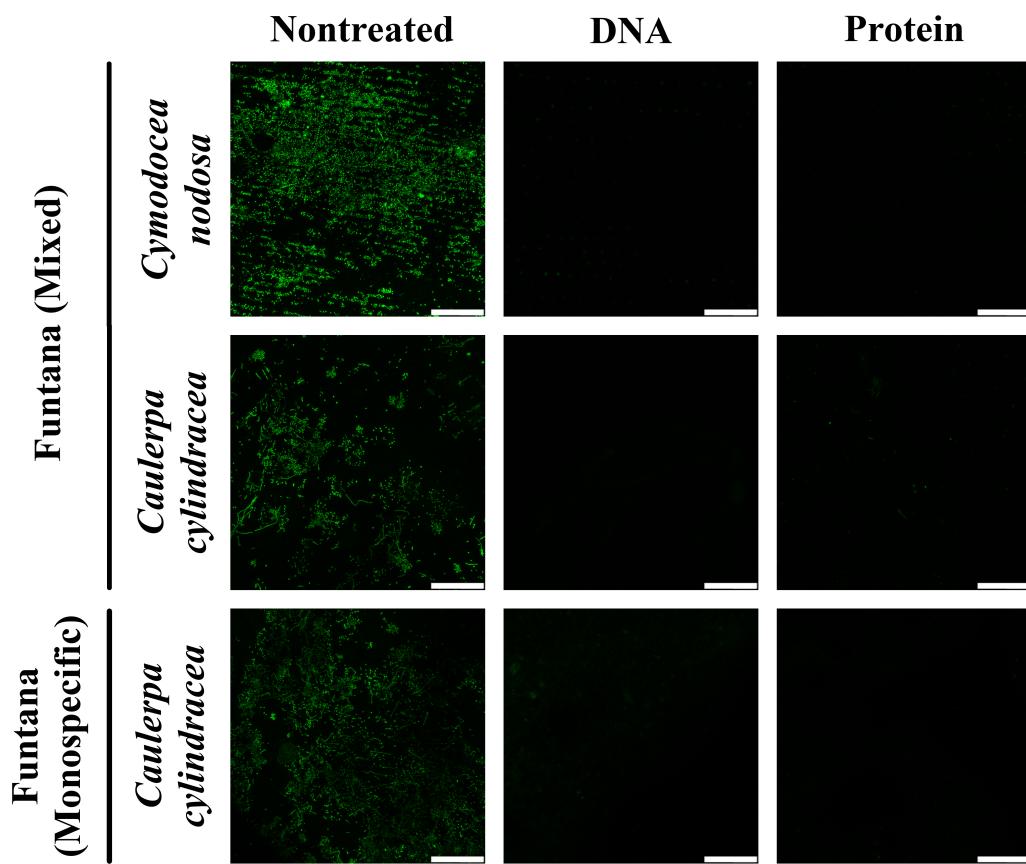


Fig. 2. Confocal microscope images of *C. nodosa* and *C. cylindracea* surfaces from the Bay of Funtana (mixed and monospecific settlements) sampled on 19 June 2018 and stained with SYBR Green I. Scale bar at all images is 60 μ m.

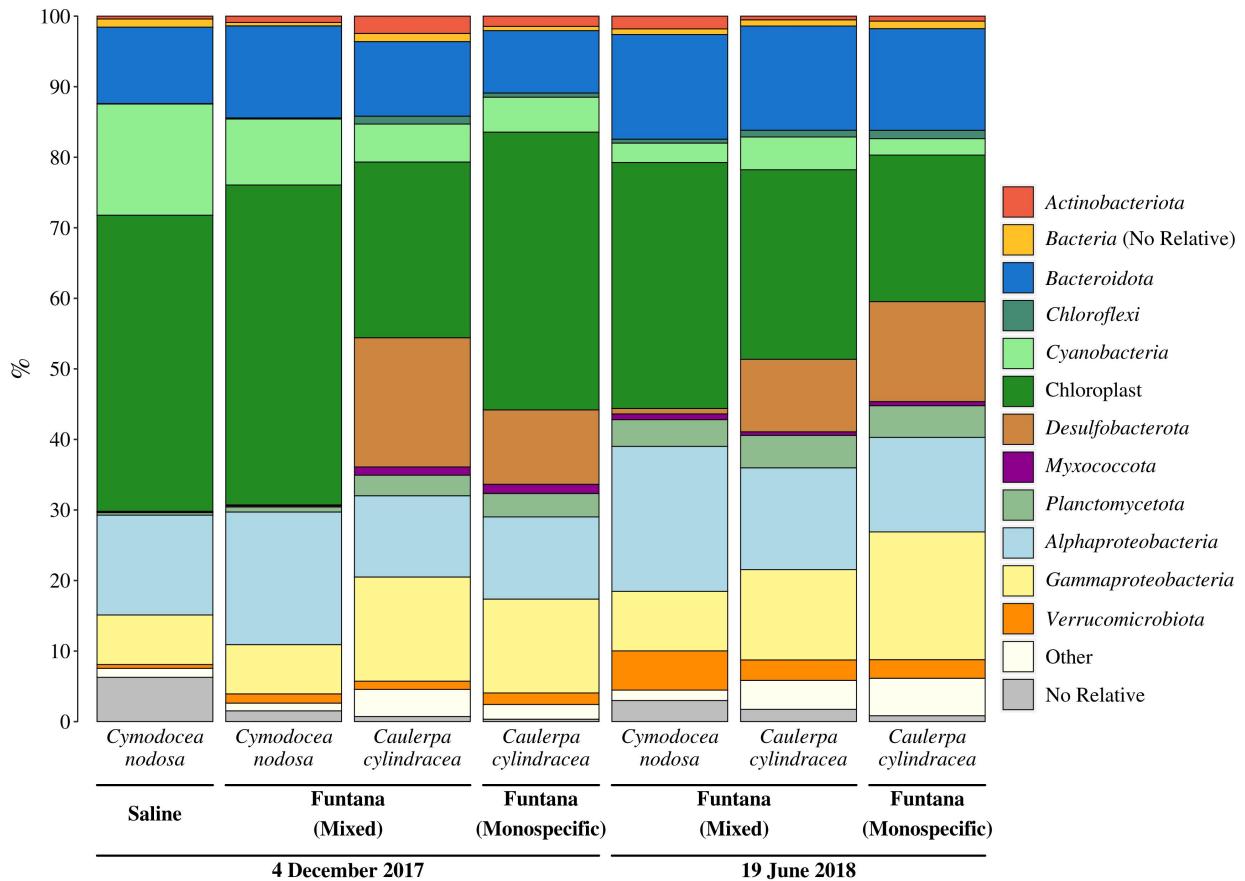


Fig. 3. Taxonomic classification and relative contribution of the most abundant ($\geq 1\%$) bacterial and archaeal sequences from surfaces of two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline and the Bay of Funtana (mixed and monospecific settlements) and in two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

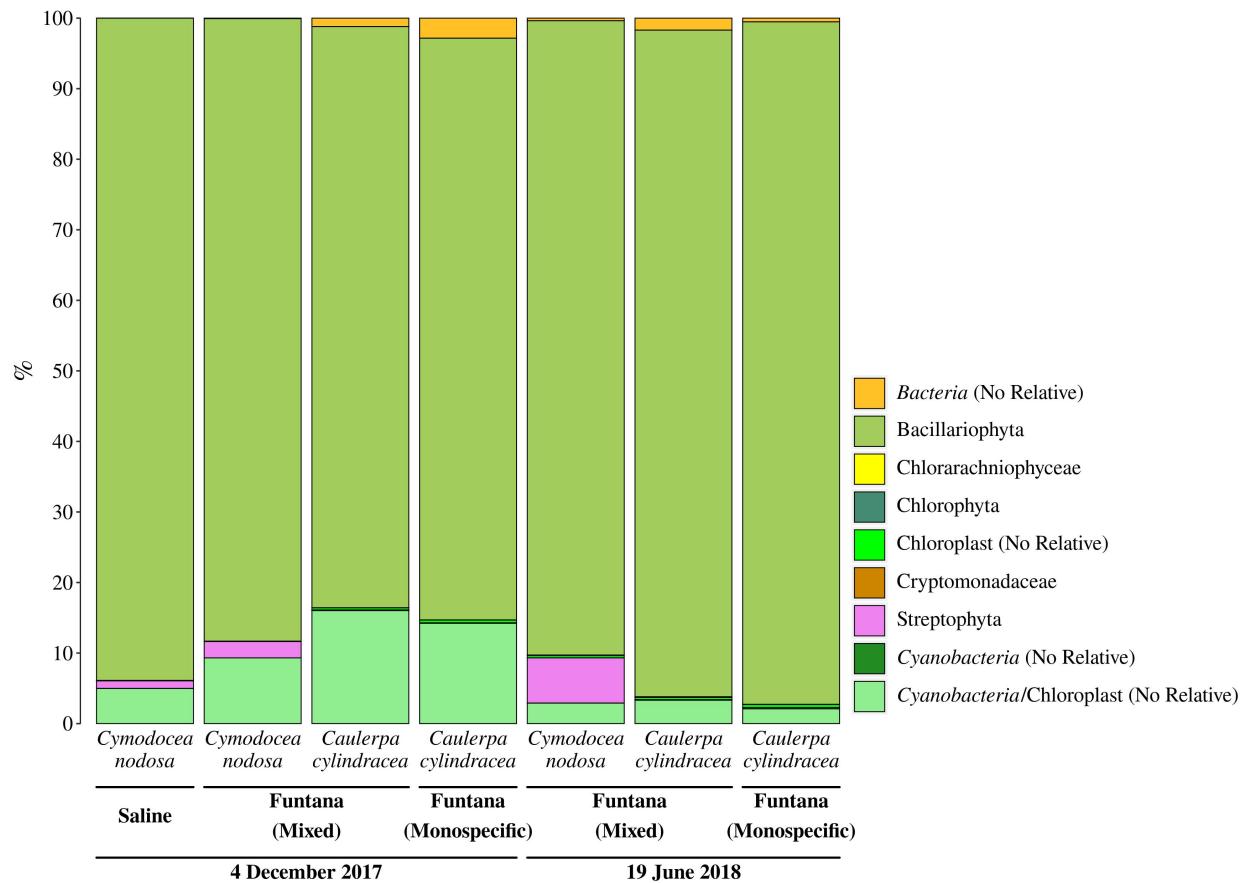


Fig. 4. Taxonomic classification and relative contribution of chloroplast sequences from surfaces of two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline and the Bay of Funtana (mixed and monospecific settlements) and in two contrasting seasons (4 December 2017 and 19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

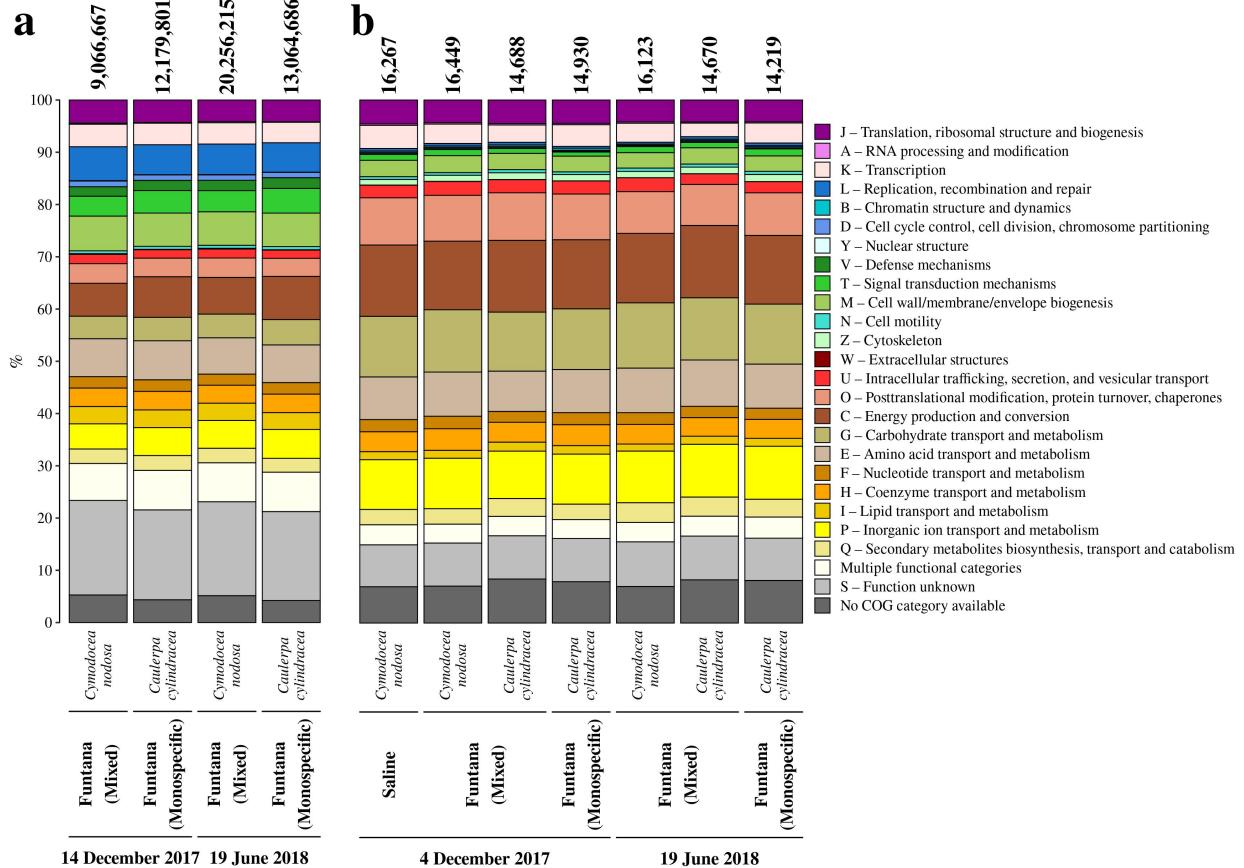


Fig. 5. Relative contribution of each COG category to the total number of annotated coding sequences (A) from metagenomes and identified proteins (B) from metaproteomes associated with surfaces of two marine macrophytes (*C. nodosa* and *C. cylindracea*) sampled in the Bay of Saline and the Bay of Funtana (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.