Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site

Journal of Phycology
April 2012, Volume 48, Issue 2, pages 483–490
http://dx.doi.org/10.1111/j.1529-8817.2012.01124.x
© 2012 Phycological Society of America

Archimer
http://archimer.ifremer.fr

The definitive version is available at http://onlinelibrary.wiley.com/

Selective elimination of chloroplastidial DNA for metagenomics of bacteria associated to green alga (*Caulerpa taxifolia, Bryopsidophyceae)*

Tânia Aires[1], Núria Marbà[2], Ester A. Serrao[1], Carlos M. Duarte[3], Sophie Arnaud-Haond[4,*]

[1] CCMAR –Center for Marine Sciences, CIMAR, FCT, University of Algarve, Gambelas, P-8005–139, Faro, Portugal
[2] Department of Global Change Research, IMEDEA (CSIC-UIB) Institut Mediterrani d'Estudis Avançats, Miquel Marques 21, 07190 Esporles, Mallorca, Spain
[3] Department of Global Change Research. IMEDEA (CSIC-UIB) Institut Mediterrani d'Estudis Avançats, Miquel Marques 21, 07190 Esporles, Mallorca, Spain The UWA Oceans Institute, University of Western Australia, 35 Stirling Highway, Crawley 6009, Australia
[4] CCMAR – Center for Marine Sciences, CIMAR, FCT, University of Algarve, Gambelas, P-8005–139, Faro, Portugal IFREMER- Technopole de Brest-Iroise BP 70 29280 Plouzané, France

*: Corresponding author : Sophie Arnaud-Haond, email address : sophie.arnaud@ifremer.fr.

Abstract :

Molecular analyses of bacteria associated with photosynthetic organisms are often confounded by coamplification of the chloroplastidial 16S rDNA with the targeted bacterial 16S rDNA. This major problem has hampered progress in the characterization of bacterial communities associated to photosynthetic organisms and has limited the full realization of the potential offered by the last generation of metagenomics approaches. A simple and inexpensive method is presented, based on ethanol and bleach treatments prior to extraction, to efficiently discard a great part of chloroplastidial DNA without affecting the characterization of bacterial communities through pyrosequencing. Its effectiveness for the description of bacterial lineages associated to the green alga *Caulerpa taxifolia* (M. Vahl) C. Agardh was much higher than that of the preexisting enrichment protocols proposed for plants. Furthermore, this new technique requires a very small amount of biological material compared to the other current protocols, making it more realistic for systematic use in ecological and phylogenetic studies and opening promising prospects for metagenomics of green algae, as shown by our data.

Keywords : 16S ; Chloroplasts ; Elimination ; green algae ; metagenomics

List of Abbreviations:

µL- microliters

µm- micrometers

16S rRNA- 16S ribosomal RNA

BE- Bleach-Ethanol pre-treatment

CF- Chloroplast Filtration protocol

CTAB- Cetyl trimethylammonium bromide

DAPI- 49.6-di amidino-2-phenylindole

DGGE- Denaturing Gradient Gel Electrophoresis

DIC- Differential interference contrast microscopy

mg- milligrams

min- minutes

mL- milliliters

Ng- nanograms

OUT- Operational Taxonomic Unit

RDP- Ribosomal Database Project

sec- seconds

SK- Spin Kit for soil

TGGE- Temperature gradient gel electrophoresis

63 TRFLP- Terminal Restriction Fragment Length Polymorphism

Introduction

The exploration of microbial diversity has been hindered until recently by the fact that about ninety-nine per cent of all microorganisms in almost every environment cannot be cultured, particularly symbiotic bacteria due to highly specialized relationships with their hosts (Amann et al. 1990, Handelsman 2004). Metagenomics is expected to become the main technique used to overcome the culture bottleneck and to unveil the diversity of bacterial communities associated with a wide range of environmental niches (Venter et al. 2004, Edwards et al. 2006, Daniel 2005, Martín-Cuadrado et al. 2007, Wang et al. 2009). The exploration of prokaryote diversity living inside other eukaryotes has not, however, progressed at the same pace as its exploration in other, non-living habitats.

Bacterial communities living associated to plants and algae have been reported across a wide range of aquatic and terrestrial niches, such as root nodules (Sun et al. 2008), galls (Ashen and Goff 1996, Ashen and Goff 1998), or inside other plant tissues (Meusnier et al. 2001, Koopman et al. 2010). Their identity, role and function are essential information for the understanding of the metabolism and dynamics of macroscopic primary producers and the ecosystems they support. The study of bacterial communities in association with macroalgae progressed significantly during the last decade, due to the use of several culture-independent molecular approaches such as community fingerprinting (DGGE, TGGE and TRFLP) and clone libraries, allowing an increase in the number and complexity of studies (Goecke et al. 2010). Several recent studies on 36 macroalgal species, brought to light 56 new bacterial species (Goecke et al. 2010) and others also demonstrated a highly specific association

of bacterial communities with marine macroalgae (Staufenberger et al. 2008, Lachnit et al. 2009, Wiese et al. 2009). A major stumbling block for the application of metagenomic analyses to investigate plant microbiota is the cyanobacterial origin of all chloroplast lineages, whether derived from primary or secondary endosymbiosis (Whatley et al. 1979), which will interfere with the 16S rRNA characterization of bacteria through massive co-amplification of the host's chloroplastidial DNA with the conserved 16S rRNA primers typically used (Chelius and Triplett 2001, Sakai et al. 2004, Green and Minz 2005, Wang et al. 2008). The construction of metagenomic libraries for bacteria associated with plants or algae is therefore technically challenging, as in the absence of a previous elimination of chloroplastidial DNA from extraction products a high amount of plant/algae-derived DNA is likely to mask the microbial contribution. Enrichment procedures based on enzymatic hydrolysis and/or subsequent differential centrifugation have recently been proposed to solve the problem of chloroplastidial contamination (Jiao et al. 2006, Wang et al. 2008) when characterizing bacterial communities associated with seeds and leaves. Yet this requires a huge amount of tissue ["in kilogram quantities"; (Wang et al. 2008)], which is not feasible or manageable for a variety of plants or algae or for specific purposes where single individuals are required. The alternative method (Jiao et al. 2006) involves a long enzymatic step at 28ºC, which may bias the bacterial community structure owing to microbial propagation (Wang et al. 2008).

Here we examined different approaches to remove "contamination" by plastidial DNA in studies aiming at characterizing bacteria associated to photosynthetic organisms. The methods under comparison were enzymatic hydrolysis and differential centrifugation (Jiao et al. 2006), and a simpler/inexpensive method based on

111 preliminary disinfection, previously described but aimed only at surface sterilizing
112 wheat roots (Coombs and Franco, 2003). We conducted these tests on the siphonous
113 green alga *Caulerpa taxifolia* (Vahl) C. Agardh, within which a diverse community of
114 endophytic bacteria has been described (Chisholm et al. 1996, Meusnier et al. 2001,
115 Delbridge et al. 2004). Comparisons across methods were based on pyrosequencing
116 estimates of ratios of chloroplastidial *versus* bacterial DNA sequences to select the
117 method revealing the most complete diversity of the associated bacterial
118 communities. Finally, the best method selected for the green algal samples was
119 applied to leaves and roots of a seagrass (*Posidonia oceanica*) in order to test for its
120 usefulness for plants.

Material & Methods

Sampling

Samples of *C. taxifolia* were collected in Villefranche-sur-mer (France) (43°42'2.93"N / 7°19'0.32"E) in March 2009 and were kept at -80ºC until DNA extraction. Sample size consisted of two individuals per treatment, each individual containing 100 mg of thalli including fronds, stolons and rhizoids. Sediment or macroscopic living organisms encrusted on the thallus surface were removed with sterile seawater and a toothbrush, identically for all control and treatment samples.

In order to test on plants the usefulness of the protocol here estimated as being the best for an alga, 2 individuals of the seagrass *Posidonia oceanica* were collected in Marseille, France (43°16'47.44"N 5°20'56.27"E). These were subdivided into samples of leaves and of the corresponding roots and were later processed using the treatment defined as optimal for *C. taxifolia* (see results).

Removal of chloroplasts and external bacteria by a bleach-ethanol pre-treatment

To remove chloroplasts and external bacteria, in order to get 16S DNA from just the endophytic bacteria, a “bleach-ethanol pre-treatment” developed by Coombs and Franco (2003) as a surface-sterilization method, was applied prior to DNA extraction. This consisted of three simple steps with modifications developed to avoid damaging the algal material: 1) the material was placed in a bath of 99% ethanol during 1 min; 2) the sample was subsequently transferred for 5 min to a 3% bleach (diluted in

seawater) solution; 3) the sample was immersed for 30 sec in a 99% ethanol solution; and finally the sample was rinsed with sterile seawater.

To attest the efficiency of this method in removing epiphytic (i.e., external) bacteria, samples were stained for 5 min with 0.1 mg/mL 49.6-di amidino-2-phenylindole (DAPI) and subsequently viewed under a confocal and fluorescence microscope with an ApoTome.2 (Zeiss, Germany) to determine whether the outer surface bacteria were effectively eliminated by the sterilization protocol applied. Differential interference contrast microscopy (DIC) (Axiovert 200 MAT Microscope (Zeiss, Ontario, NY, USA) was also performed.

This treatment was compared with a previously described chloroplast removal method, based on a filtration step, the method II of Jiao et al. (2006).

Bacterial DNA extraction

For DNA extraction two types of methods were tested, the FastDNA Spin Kit for Soil (MP biomedicals LLC, Solon, OH, USA, Catalog # 6540-600) hereafter named SK, and the CTAB plant DNA extraction procedure of Doyle and Doyle (1987).

Experimental treatments

The bleach-ethanol (BE) pre-treatment combined with/without chloroplast filtration (CF) and with two bacterial DNA extraction methods (SK and CTAB), plus a non-pre-treated control resulted in five treatments (Table 1): A) - no BE pre-treatment,

chloroplast removal by filtration (CF), SK DNA extraction. This was the control for the performance of the "bleach-ethanol pre-treatment (BE)"; B) - BE pre-treatment, CF, SK DNA extraction; C) - BE pre-treatment, no CF, SK DNA extraction; D)- BE pre-treatment, CF, CTAB DNA extraction; E)- BE pre-treatment, no CF, CTAB DNA extraction.

DNA sequencing

The DNA concentration in each sample was measured on a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA), to verify that samples had concentrations in excess of 50 ng μL^{-1}, the value required for metagenomic analyses.

DNA was submitted to BIOCANT (Cantanhede, Portugal) for tag-pyrosequencing. Samples were amplified with modified primers for region V4 of 16S rRNA (Wang et al. 2007) and the amplicons were sequenced using 454 GS FLX (Roche-454 Life Sciences, 454 Life Sciences, Branford, CT, USA) with Titanium chemistry. The quality criteria imposed to consider sequences in the analysis were: i) sequences with less than 100 nucleotides were discarded; ii) sequences with 2 or more undetermined nucleotides were discarded. Sequences with differences below 3% were grouped and considered the same Operational Taxonomic Unit (OTU).

Statistical analyses

The RDP (Ribosomal Database Project) pyrosequencing pipeline (Cole et al. 2009) was used to align the 454 sequence data and assign identities with 80% confidence. The software MOTHUR version 1.4.1 (Schloss et al. 2009) was used to construct distance matrices (dist.seq function), assign sequences to OTUs (97% similarity), calculate OTU

richness using the nonparametric estimator Chao 1 (Chao 1987), calculate coverage (%) and construct rarefaction curves. The Chao parameter predicts total diversity (OTU richness) of a community, providing an indicator of the coverage of the bacterial community allowed by the sequencing effort (Handelsman 2004). Chao estimation was based on (Chao 1984):

$$S_{Chao1} = S_{obs} + \frac{n_1(n_1 - 1)}{n_2(n_2 - 1)}$$

S_{obs}- Number of observed OTUs

n_1- Number of OTUs observed once (singletons)

n_2- Number of OTUs observed twice (doubletons)

Results

The bleach-ethanol sterilization was efficient in reducing external bacteria to very few that could only be found after screening several images, contrary to the control treatment, as revealed by using DAPI-stained bacterial DNA fluorescence on treated and control samples ~~(Fig.1 a and b)~~.

In *C. taxifolia*, most organisms present in each sample were assigned to, at least, the Class level (Table 2). About 98% of all sequences passed the quality control and were used in further statistical analyses (Table 1). There were major, four-fold differences, in the number of sequences derived using different protocols (Table 1). Samples prepared without the "bleach-ethanol pre-treatment" (control treatment A) yielded the lowest results in all statistical parameters measured (Table 1). Treatment D yielded the highest number of OTUs (505 upon 1800 sequences, 455 OTUs when standardized for the lowest number of sequences, 1328) and treatment A (used as control) yielded

207 the lowest (158 upon 1328 sequences) (Table 1 and Fig.2). The number of OTUs varied
208 among treatments, but all seemed close to reach the plateau along the rarefaction
209 curve (Fig.2). Although the number of OTUs was beneath the expected Chao 1 index in
210 all treatments, coverage was satisfactory for most samples (about 80% to 90%) except
211 for the non-disinfected one (A, with ca. 60%) due to dominance of chloroplastidial
212 sequences.

213 Comparison of the no bleach-ethanol treatment (control A) with its equivalent pre-
214 treated (B) reveals that the sequences of the former are dominated by the Kingdom
215 Viridiplantae, *i.e.* chloroplastial sequences (77.64%; Fig.3), whereas treatment B
216 resulted in only 0.54% of chloroplasts sequences (Table 1). Comparing pairwise
217 treatments that differ only in the use or not of the chloroplast filtration step (B *vs* C
218 and D *vs* E), showed no consistent differences in the number of chloroplast sequences,
219 which was always negligible (i.e respectively 0.54 *vs* 0.26 % and 0.22 *vs* 0.55 %; Table
220 1) in all these cases pre-treated with bleach-ethanol.

221 As for the diversity of OTUs, the no ethanol-bleach treatment A showed much fewer
222 bacterial OTUs compared to all others, mostly due to dominance by chloroplastidial
223 sequences (Fig.3). The difference in diversity between filtered *vs* non-filtered is
224 negligible for the treatments with the Soil Kit extraction (B and C) (392 and 378 OTUs,
225 respectively, for 1328 sequences; Table 1, Fig.2). For CTAB extractions, a larger
226 difference is observed between the filtered (D) and non-filtered (E) treatments (1800
227 *vs* 1466 for the number of sequences and 505 *vs* 235 for the number of OTUs
228 identified; 455 *vs* 227 OTUs once standardized for the lowest amount of sequences
229 that passed the quality control, 1328 in treatment A; Table 1, Fig.2). In summary, our

230 results showed that adding the filtration step to the bleach-ethanol pre-treatment did
231 not result in dramatic diversity changes with any of the two extraction methods used
232 (CTAB and the Soil Kit), and that the Soil Kit extraction seems to deliver more diversity
233 and more consistent results (coverage 90/91% for Soil Kit against 76/85% for CTAB;
234 Table 1).

235 The Delta, Gamma and Alfa subdivisions of *Proteobacteria* were the most abundant
236 OTUs in all treatments (Fig.4, Table2), independently of the chloroplast removal or
237 DNA extraction protocol. The relative proportions of the main bacteria classes were
238 similar among samples, except for sample A where the number of classes identified
239 was much smaller and *Clostridia* (one of the most abundant in other treatments) was
240 not present (Fig.4, Table 2).

241 Detailed results on sequences obtained in these treatments are not presented since
242 the aim of this study was not to fully describe the bacterial community associated with
243 *C. taxifolia*.

244 Application to leaves and roots of a plant host (*P. oceanica*) of the treatment that
245 resulted in the best bacterial diversity description for the alga (bleach-ethanol followed
246 by soil kit DNA extraction) was not effective in removing chloroplasts. This resulted in
247 more than 90% of chloroplastidial sequences in all samples except the roots of
248 individual 1 with 80% (Table 1). In both root samples (individuals 1 and 2) many
249 sequences (17.70% and 9.38% respectively) were unassigned (RDP Classifier
250 automatically bins unclassified data <80% confidence) and taxonomic assignment of
251 bacteria was low, respectively 0.11%, 0.15%, 0.77% and 1.04% for leaves and roots of
252 individuals 1 and 2. Coverage was below 50% in almost all cases.

253 Discussion

254 This study discovered an effective way of obtaining representative bacterial 16S DNA
255 sequences associated to a green alga without interference from the dominant
256 chloroplast 16S DNA. This is a significant methodological step forward towards
257 progress in the large-scale characterization of microbial communities associated to
258 primary producers. A large amount of information and data on microbial ecology have
259 been obtained during the last five years through metagenomics approaches (Galperin
260 2004, Edwards and Rohwer 2005, Poinar et al. 2006, Wilhelm et al. 2007, Thurber et al.
261 2008, Lazarevic et al. 2009, Kielak et al. 2010). Although some cryptic habitats have
262 already been explored and their hidden microbial life unveiled (Venter et al. 2004,
263 Edwards et al. 2006), the study of microbes associated with plants and algae remains a
264 major challenge for metagenomics. Predominant cross-amplification of chloroplasts
265 masking bacterial 16S rDNA is likely the major factor explaining the paucity of studies
266 taking advantage of the last metagenomics advances to characterize bacteria
267 associated with algae and plants.

268 Another important result was the low performance, on milligrams of *C. taxifolia* tissue,
269 of the first protocol proposed to discard chloroplast sequences (chloroplast filtration),
270 which had been demonstrated on grams to kilograms of leaves and/or seeds of tree
271 species (Jiao et al. 2006). Only about 23% of the DNA sequences obtained using
272 chloroplast filtration alone did correspond to bacterial strains compared to about 90 to
273 97% reported by Jiao et al (2006). The differential centrifugations of this procedure
274 were previously shown as not efficient in eliminating chloroplasts (Wang et al. 2008).
275 Two main hypotheses might explain this unexpected difference. Either the structural
276 difference in algal *versus* trees tissues resulted in a different success of the filtration

277 protocol, or the enormous difference in the amount of material used, miligrams for *C.*
278 *taxifolia versus* kilograms in the previous studies, influenced the success of filtration.
279 The alternative method used here to remove chloroplasts, based on a “bleach-ethanol
280 pre-treatment” that was previously described for surface disinfection (Coombs and
281 Franco, 2003), proved to be a simple method to discard a great part of the chloroplasts
282 for optimization of metagenomics analyses of endophytic bacteria. Less than 1% of
283 chloroplast sequences remained in the samples pre-treated with bleach-ethanol,
284 independently of whether the chloroplast filtration step was added or not, revealing
285 that this step caused no improvement in the ratio of bacterial *versus* chloroplastidial
286 sequences. An overwhelming difference between the no bleach-ethanol pre-treatment
287 and the corresponding pre-treated one shows the efficiency of this pre-treatment in
288 eliminating chloroplasts.
289 This effectiveness of the bleach-ethanol pre-treatment may be expected on a wider
290 range of photosynthetic organisms and in particular on a diversity of green algae on
291 which elimination of chloroplasts with bleach is consistently used in taxonomy and
292 histology (Habib Langar, pers. com.). However, Hollants and colleagues (2010)
293 developed a surface disinfection method for the green coenocytic alga *Bryopsis sp.*
294 based on a combination of CTAB buffer, proteinase K and the bactericidal cleanser
295 Umonium Master and compared it to ethanol and bleach surface disinfection. Ethanol
296 cleaning was ineffective both in eliminating epiphytes and for bacterial 16S rDNA
297 amplification. Instant bleach in this study caused an apparent elimination of the
298 endosymbionts, suggesting that *Bryopsis* cell structures may be more fragile than
299 those of *Caulerpa* causing higher bleaching susceptibility.

As the first step in the disinfection protocol is an ethanol bath for 1 minute, this alcohol may act as a detergent/solvent, breaking down the phospholipid bilayer and opening holes in the membrane, making it permeable and, in this case, promoting the invasion of the bleach (Baker et al. 1979) that further damages and apparently eliminates the chloroplasts. One concern might be that bleach would also selectively damage and eliminate some bacterial lineages inside the alga, even though bacteria have more complex and robust membranes and cell walls compared to chloroplasts. Yet, despite the much more limited number of bacterial sequences obtained for the non-disinfected sample, the bleach-ethanol disinfection did not seem to lead to any significant loss of strains according to overall list of OTUs recognized in each sample (i.e. no major strains appeared in the non-disinfected treatment A that were not in the other treatments). Considering Class the taxonomic level of our analysis, the only strain that appears exclusively in the non-disinfected (not discarding the possibility of being present in the other samples, although only classified to Phylum) sample, was Flavobacteria, namely from the genus *Aquimarina* and *Tenacibaculum*. These strains are known to be present as free-living in marine waters (Nedashkovskaya et al. 2006) in the first case, and to be fixed to the surface of marine organisms (Suzuki et al. 2001) in the second. *Tenacibaculum* genus is thought to be one of the strains responsible for inducing morphogenesis in algae (Hanzawa et al. 1998, Matsuo et al. 2003). Their exclusive presence in the non-disinfected protocol may therefore be explained by an external occurrence, explaining their absence on the surface of samples disinfected precisely to improve the external cleaning, i.e., removing external bacteria. Furthermore, coverage was relatively high and satisfying in all disinfected samples, particularly those obtained from the FastDNA® SPIN Kit for Soil (about 90% with or

without filtration), whereas the non-disinfected sample reached only 64% coverage. It is therefore likely that dominant OTUs are well characterized, although some poorly represented OTUs could still escape detection.

Contrary to previous methods proposed to eliminate chloroplasts for metagenomics purposes, the “bleach-ethanol pre-treatment” surface-disinfection is fast and inexpensive, requiring a limited amount of material and solving the serious problem of excess chloroplast DNA when estimating bacterial diversity, while at the same time selectively removing epiphytic bacteria. It seems therefore more accurate for ecological or phylogenetic studies of associated bacteria that may require a protocol adapted to a lower amount of plant or algal tissues, and allowing a higher throughput (i.e. a much larger number of specimens to be analyzed). Besides, it showed superior in performance in these conditions for *C. taxifolia.* The results presented here also suggest a slightly better performance of the FastDNA® SPIN Kit for Soil over the classical CTAB method for the *C. taxifolia* samples studied here.

The additional tests made with *P. oceanica* suggest that this disinfection protocol is not as effective for plants, as the chloroplast sequences remained on the order of 90%. Among the 10 to 20% remaining OTUs the low rate of identified bacterial and high number of unassigned (<80% confidence) sequences did not give very encouraging results. Besides, the RDP classifier is a database for bacteria and, despite effectively identifying chloroplast-derived sequences, there is evidence that it is not that effective in identifying plant mitochondrial sequences (Kretzer et al. 2009). Unassigned sequences may therefore be either cryptic bacterial lineages absent from the database, or mitochondrial sequences from *P. oceanica*.

347 The results presented in this study open promising perspectives to unlock the
348 characterization of diversity of bacterial communities associated to algae, although
349 other methods still have to be explored for plants.

350

351 Acknowledgments:

352 We would like to thank to Frédéric Zuberer for his invaluable help in sampling *C.*
353 *taxifolia* in Villefranche, BIOCANT center for metagenomics analysis and support in
354 bioinformatic issues, Ronan Becheler and Neus Garcias for the technical assistance in
355 the lab, and Habib Langar (from the INSTM of Salâmbo, Tunis) for useful discussions
356 and comments about this method. This work was supported by the project IBISA (PTDC
357 / MAR / 64749 / 2006) funded by the Portuguese Foundation for Science and
358 Technology (FCT) and FEDER and a PhD fellowship (TA) from FCT and FSE.

359

360

361

362

363

364

365

366

References:

Amann, R.I., Binder, B.L., Chisholm, S.W., Devereux, R. & Stahl, D.A. 1990. Combination of 16S rRNA targeted oligonucleotide probes with Flow-cytometry for analysing mixed microbial populations. *Appl Environ Microbiol*. 56:, 1919–1925.

Ashen, J.B. & Goff, L.J. 1996. Molecular identification of a bacterium associated with gall formation in the marine red alga Prionitis lanceolata. *J Phycol*. 32: 286-297.

Ashen, J.B. & Goff, L.J. 1998. Galls on the marine red alga Prionitis lanceolata (Halymeniaceae): Specific induction and subsequent development of an algal-bacterial symbiosis. *Am J Bot.* 85: 1710-1721.

Baker, E.C., Mustakas, G.C. & Warner, K.A. 1979. Extraction of Defatted Soybean Flours and Flakes with Aqueous Alcohols - Evaluation of Flavor and Selected Properties. *J Agric Food Chem*. 27: 969-973.

Chao, A. 1984. Non-parametric estimation of the number of classes in a population. *Scand J Stat*. 11: 265-270.

Chao, A. 1987. Estimating the Population-Size for Capture Recapture Data with Unequal Catchability. *Biometrics*. 43: 783-791.

Chelius, M.K. & Triplett, E.W. (2001). The diversity or Archaea and Bacteria in association with the roots of Zea mays L. *Microb Ecol.* 41:252–263.

Chisholm, J.R.M., Dauga, C., Ageron, E. & Grimont, P.A.D. 1996. Roots in mixotrophic algae. *Nature*. 381: 382.

Coombs, J.T. & Franco, C.M.M. 2003. Isolation and identification of actinobacteria from surface-sterilized wheat roots. *Appl Environ Microbiol*. 69: 5603-5608.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Marsh, T., Garrity, G.M. & Tiedje, J.M. 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nuc. Acids Res.* 37:D141-5.

Daniel, R. 2005. The metagenomics of soil. *Nat Rev Microbiol*. 3: 470-478.

Delbridge, L., Coulburn, J., Fagerber, W. & Tisa, L.S. 2004. Community profiles of bacterial endosymbionts in four species of *Caulerpa. Symbiosis*. 37: 335–344.

Doyle, J.J. & Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19: 11-15.

412 Edwards, R.A., Rodriguez-Brito, B., Wegley, L., Haynes, M., Breitbart, M., Peterson,
413 D.M., Saar, M.O., Alexander, S., Alexander, E.C. & Rohwer, F. 2006. Using
414 pyrosequencing to shed light on deep mine microbial ecology. *Bmc Genomics*. 7: 57.
415
416 Edwards, R.A. & Rohwer, F. 2005. Viral metagenomics. *Nat Rev Microbiol*. 3: 504-510.
417
418
419 Galperin, M.Y. 2004. Metagenomics: from acid mine to shining sea. *Environ Microbiol*.
420 6: 543-545.
421
422
423 Green, S.J. & Minz, D. (2005). Suicide polymerase endonuclease restriction, a novel
424 technique for enhancing PCR amplification of minor DNA templates. *Appl Environ*
425 *Microbiol.* 71:4721–4727.
426
427 Goecke, F., Labes, A., Wiese, J. & Imhoff, J.F. (2010). Chemical interactions between
428 marine macroalgae and bacteria*. Mar. Ecol. Prog. Ser.* 409: 267-300.
429
430 Handelsman, J. 2004. Metagenomics: applications of genomics to uncultured micro-
431 organisms. *Microbiol Mol Biol Rev.* 68:669–685.
432
433 Hanzawa, N., Nakanishi, K., Nishijima, M. & Saga, N. 1998. 16S rDNA-based
434 phylogenetic analysis of marine flavobacteria that induce algal morphogenesis. *J Mar*
435 *Biotechnol*. 6: 80-82.

436

437 Hollants, J., Leliaert, F., De Clerck, O. & Willems, A. 2010. How endo- is endo-? Surface
438 sterilization of delicate samples: a Bryopsis (Bryopsidales, Chlorophyta) case study.
439 *Symbiosis*. 51: 131-138.

440

441 Jiao, J.Y., Wang, H.X., Zeng, Y. & Shen, Y.M. 2006. Enrichment for microbes living in
442 association with plant tissues. *J Appl Microbiol*. 100: 830-837.

443

444 Kielak, A., Rodrigues, J. L. M., Kuramae, E.E., Chain, P.S.G., van Veen, J.A. & Kowalchuk,
445 G.A. 2010. Phylogenetic and metagenomic analysis of Verrucomicrobia in former
446 agricultural grassland soil. *FEMS Microbiol Ecol*. 71: 23-33.

447

448 Koopman, M. M., Fuselier, D.M., Hird,S. & Carstens, B.C. 2010. The Carnivorous Pale
449 Pitcher Plant Harbors Diverse, Distinct, and Time-Dependent Bacterial Communities.
450 *Appl Environ Microbiol.*76: 1851-1860.

451

452 Kretzer, A.M., King, Z.R. & Bai S. 2009. Bacterial communities associated with
453 tuberculate ectomycorrhizae of Rhizopogon spp. *Mycorrhiza*. 19:277–282.

454

455 Lachni, T., Blümel, M., Imhoff, J.F. & Wahl, M. 2009. Specific epibacterial communities
456 on macroalgae: phylogeny matters more than habitat. *Aquat Biol* 5: 181–186.

457

458 Lazarevic, V., Whiteson, K., Huse, S., Hernandez, D., Farinelli, L., Osteras, M., Schrenzel,
459 J. & Francois, P. 2009. Metagenomic study of the oral microbiota by Illumina high-
460 throughput sequencing. *J Microbiol Methods*. 79: 266-271.
461
462 Martín-Cuadrado, A.B., López-García, P., Alba, J.C., Moreira, D., Monticelli, L.,
463 Strittmatter, A., Gottschalk, G. & Rodríguez-Valera, F. 2007. Metagenomics of the Deep
464 Mediterranean, a Warm Bathypelagic Habitat. *Plos One*. 2, e914.
465
466 Matsuo, Y., Suzuki, M., Kasai, H., Shizuri, Y. & Harayama, S. 2003. Isolation and
467 phylogenetic characterization of bacteria capable of inducing differentiation in the
468 green alga Monostroma oxyspermum. *Environ Microbiol*. 5: 25-35.
469
470 Meusnier, I., Olsen, J.L., Stam, W.T., Destombe, C. & Valero, M. 2001. Phylogenetic
471 analyses of *Caulerpa taxifolia* (Chlorophyta) and of its associated bacterial microflora
472 provide clues to the origin of the Mediterranean introduction. *Mol Ecol.* 10: 931–946.
473
474 Nedashkovskaya, O.I., Vancanneyt, M., Christiaens, L., Kalinovskaya, N.I., Mikhailov,
475 V.V. & Swings, J. 2006. Aquimarina intermedia sp nov., reclassification of Stanierella
476 latercula (Lewin 1969) as Aquimarina latercula comb. nov and Gaetbulimicrobium
477 brevivitae Yoon et al. 2006 as Aquimarina brevivitae comb. nov and emended
478 description of the genus Aquimarina*. Int J Syst Evol Microbiol*. 56: 2037-2041.
479

480 Poinar, H., Schwarz, C., MacPhee, R., Miller, W. & Schuster, S. 2006. Mammoth
481 metagenomics: New technologies, new insights, new possibilities. *J Vertebr Paleontol*.
482 26: 110A-110A.
483
484
485 Sakai, M., Matsuka, A., Komura, T., Kanazawa, S. (2004). Application of a new PCR
486 primer for terminal restriction fragment length polymorphism analysis of the bacterial
487 communities in plant roots. *J Microbiol Methods.* 59:81–89.
488
489 Schloss, Patrick D; Westcott, Sarah L; Ryabin, Thomas; Hall, Justine R; Hartmann,
490 Martin; Hollister, Emily B; Lesniewski, Ryan A; Oakley, Brian B; Parks, Donovan H;
491 Robinson, Courtney J; Sahl, Jason W; Stres, Blaz; Thallinger, Gerhard G; Van Horn,
492 David J; Weber & Carolyn F. 2009. Introducing mothur: Open-source, platform-
493 independent, community-supported software for describing and comparing microbial
494 communities. *Appl Environ Microbiol*. 75(23): 7537-41
495
496 Staufenberger, T., Thiel, V., Wiese, J. & Imhoff JF. 2008. Phylogenetic analysis of
497 bacteria associated with *Laminaria saccharina. FEMS Microbiol Ecol*. 64: 65–77.
498
499 Sun, L., Qiu, F.B., Zhang, X.X., Dai, X., Dong, X.Z. & Song, W. 2008. Endophytic bacterial
500 diversity in rice (Oryza sativa L.) roots estimated by 16S rDNA sequence analysis.
501 *Microb Ecol*. 55: 415-424.
502

503 Suzuki, M., Nakagawa, Y., Harayama, S. & Yamamoto, S. 2001. Phylogenetic analysis
504 and taxonomic study of marine Cytophaga-like bacteria: proposal for Tenacibaculum
505 gen. nov with Tenacibaculum maritimum comb. nov and Tenacibaculum ovolyticum
506 comb. nov., and description of Tenacibaculum mesophilum sp nov and Tenacibaculum
507 amylolyticum sp nov. *Int J Syst Evol Microbiol*. 51: 1639-1652.
508
509 Thurber, R.L.V., Barott, K.L., Hall, D., Liu, H., Rodriguez-Mueller, B., Desnues, C.,
510 Edwards, R.A., Haynes, M., Angly, F.E., Wegley, L. & Rohwer, F.L. 2008. Metagenomic
511 analysis indicates that stressors induce production of herpes-like viruses in the coral
512 Porites compressa. *Proc Natl Acad Sci USA*. 105: 18413-18418.
513
514 Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu,
515 D.Y., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas,
516 M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H.,
517 Pfannkoch, C., Rogers, Y.H. & Smith, H.O. 2004. Environmental genome shotgun
518 sequencing of the Sargasso Sea. *Science*. 304: 66-74.
519
520 Wang, Q., Garrity, G.M., Tiedje, J.M. & Cole, J.R. 2007. Naive Bayesian classifier for
521 rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ*
522 *Microbiol*. 73: 5261-5267.
523
524 Wang, H.X., Geng, Z.L., Zeng, Y. & Shen, Y.M. 2008. Enriching plant microbiota for a
525 metagenomic library construction. *Environ Microbiol*. 10: 2684-2691.
526

527 Wang, S.F., Xiao, X., Jiang, L.J., Peng, X.T., Zhou, H.Y., Meng, J.& Wang, F.P. 2009.
528 Diversity and Abundance of Ammonia-Oxidizing Archaea in Hydrothermal Vent
529 Chimneys of the Juan de Fuca Ridge. *Appl Environ Microbiol*. 75: 4216-4220.
530
531 Whatley, J.M., John, P. & Whatley, F.R. 1979. From Extracellular to Intracellular -
532 Establishment of Mitochondria and Chloroplasts. *Proc. R. Soc. London, Ser. B*. 204: 165-
533 187.
534
535 Wiese, J., Thiel, V., Nagel, K., Staufenberger, T. & Imhoff, J.F. 2009. Diversity of
536 antibiotic active bacteria associated with the brown alga *Laminaria saccharina* from
537 the Baltic Sea. *Mar Biotechnol*. 11: 287–300.
538
539 Wilhelm, L.J., Tripp, H.J., Givan, S.A., Smith, D.P. & Giovannoni, S.J. 2007. Natural
540 variation in SARII marine bacterioplankton genomes inferred from metagenomic data.
541 *Biol Direct*. **2**:27.

Table 1- Treatments applied to obtain bacterial DNA associated to *Caulerpa taxifolia* (A-E) and *Posidonia oceanica* (**Individual1/individual2**), and summary of sequencing results and statistical parameters calculated for each protocol. The % of coverage represents the number of OTUs in each sample divided by Chao factor.

Experimental treatments	BE	CF	DNA extract.	N of sequences obtained	N of sequences after quality control[1]	OTU	Chao1	Coverage (%)	N of OTUs standardized for the lowest*	% of chloroplastidial sequences
A (control)	-	+	SK	1370	1328*	158	248.621	63.55	158	77.64
B	+	+	SK	1522	1496	403	439.457	91.41	392	0.54
C	+	-	SK	1555	1533	391	435.536	89.77	378	0.26
D	+	+	CTAB	1839	1800	505	597.764	84.48	455	0.22
E	+	-	CTAB	1488	1466	235	308.244	76.23	227	0.55
***P. oceanica* Leaves**	+	-	SK	6361/8078	6361/8078	349/466	768.45/906.917	45,41/51.38	-	97.30/94.41
***P. oceanica* Roots**	+	-	SK	9512/12073	9462/12073	942/578	2654.349/1334.276	35.48/43.32	-	80.44/90.01

BE- "Bleach-Ethanol pre-treatment"; CF - Chloroplast filtration step (Jiao et al. 2006); SK- FastDNA® SPIN Kit for Soil DNA extraction, CTAB – DNA extraction following Doyle & Doyle 1987. [1] – Removal of chimeras and sequences with less than 100 bp or more than 2 undetermined nucleotides.

Table 2- Distribution (in %) of each OTU Class by the different Treatments

OTUs Class	Treatment A	Treatment B	Treatment C	Treatment D	Treatment E
Deltaproteobacteria	11.11	44.35	30.17	39.65	31.94
Gammaproteobacteria	31.50	29.03	31.03	28.40	33.33
Alphaproteobacteria	44.44	10.49	11.21	8.88	16.67
Clostridia	0.00	2.42	8.62	7.10	5.56
Betaproteobacteria	1.85	3.23	3.45	5.33	0.00
Cyanophyceae	3.70	4.83	5.17	2.96	8.33
Bacilli	1.85	2.42	6.04	2.37	1.39
Deferribacter	0.00	1.61	0.00	1.18	0.00
Bacteroidea	0.00	0.81	1.72	1.18	0.00
Verrucomicrobiae	0.00	0.00	0.00	0.59	0.00
Nitrospira	0.00	0.00	0.00	0.59	0.00
Mollicutes	0.00	0.00	0.00	0.59	0.00
Chlamydiae	0.00	0.00	0.00	0.59	0.00
"Chlorobi group"	0.00	0.00	0.00	0.59	0.00
Spirochaetes	0.00	0.00	2.59	0.00	0.00
Flavobacteria	3.70	0.00	0.00	0.00	0.00
Deinoccoci	0.00	0.00	0.00	0.00	1.39
Sphingobacteria	1.85	0.81	0.00	0.00	1.39

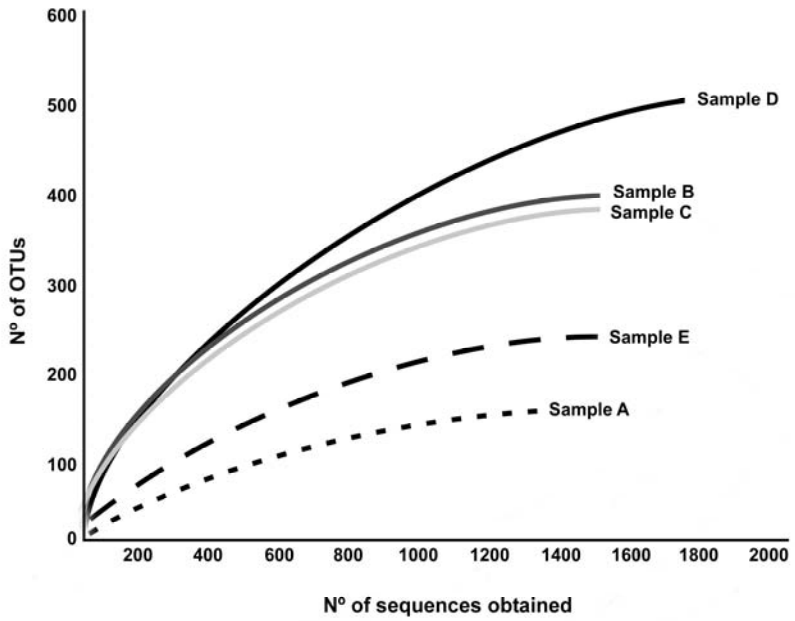

Fig.1. Rarefaction curves derived for samples processed using the different protocols (A-E, cf. Table 1).

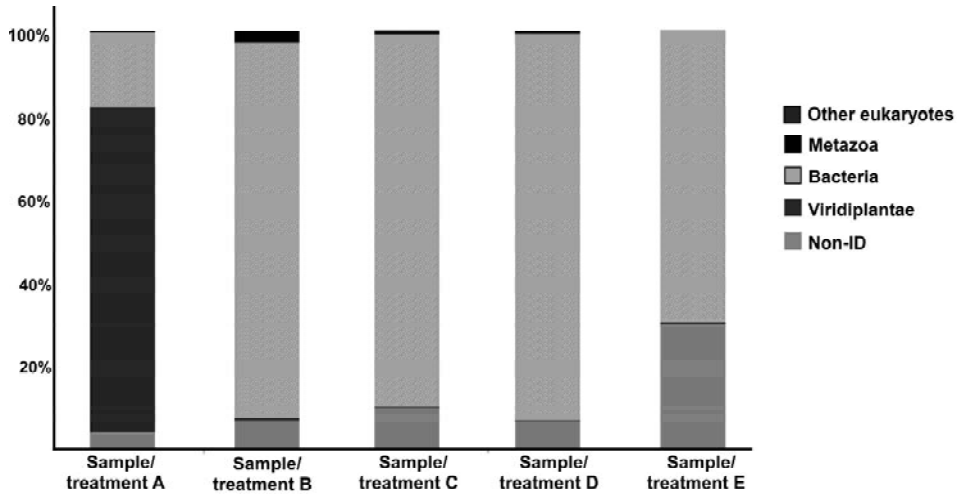

Fig.2. Distribution of OTUs per Kingdom in samples derived using the different treatments (A-E, cf. Table 1). Non-ID sequences were those that had more than 3% difference from the best BLAST hit (and then rejected) or sequences which BLAST result was “uncultured”.

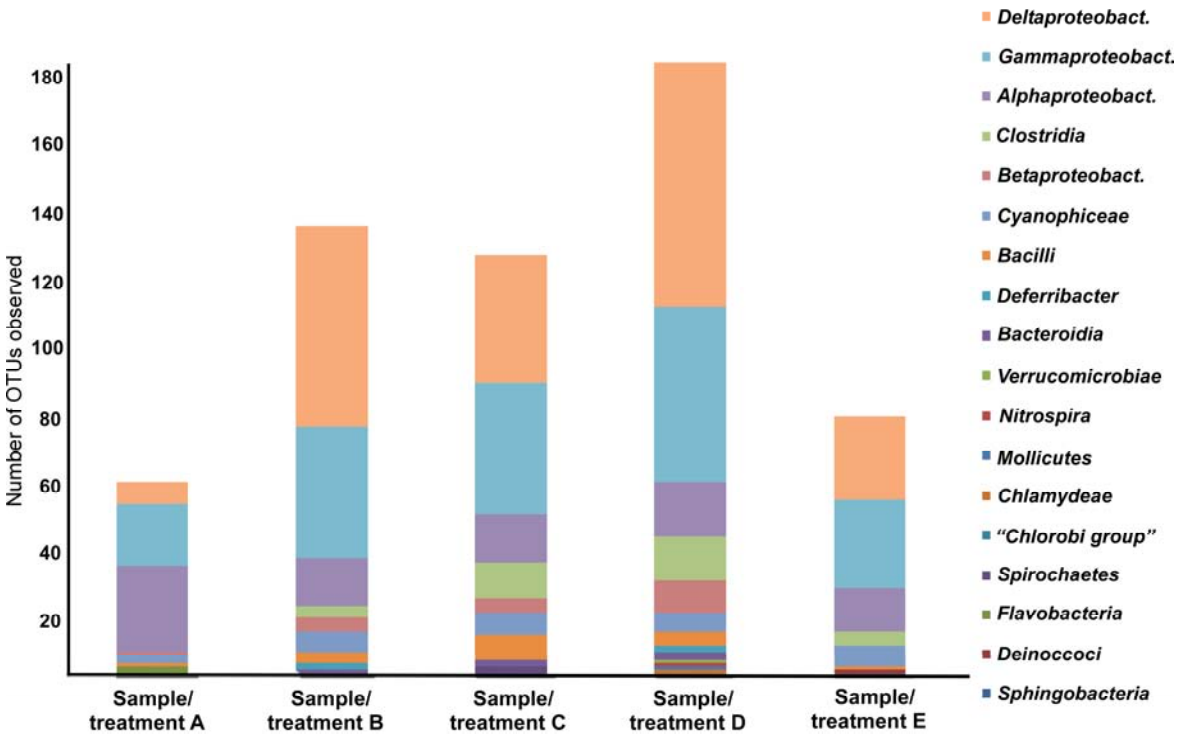

Fig.3. Distribution of OTUs into bacteria classes in samples derived using the different treatments (A-E, cf. Table 1).