Selective DNA and Protein Isolation from Marine Macrophyte Surfaces

[1†]

† To whom correspondence should be addressed: marino.korlevic@irb.hr

1. Ruđer Bošković Institute, Center for Marine Research, G. Paliaga 5, Rovinj, Croatia

2. University of Vienna, Department of Limnology and Bio-Oceanography, Althanstraße 14, Vienna, Austria

Abstract

Introduction

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

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

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

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

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

Materials and Methods

Sampling

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

DNA Isolation, Illumina 16S rRNA Sequencing and Sequences Analysis

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

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

An aliquot of isolated DNA was treated with RNase A (final concentration 200 $\mu g\,ml^{-1}$) for 2 h at 37 °C. The DNA concentration was determined using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, USA) according to the manufacturer's instructions and diluted to 1 $ng\,\mu l^{-1}$. The V4 region of the 16S rRNA gene was amplified using a two-step PCR procedure. In the first PCR the 515F (5'-GTGYCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACNVGGGTWTCTAAT-3') primers from the Earth Microbiome Project (http://press.igsb.anl.gov/earthmicrobiome/protocols-and-standards/16s/) were used to amplify the target region (Caporaso *et al.*, 2012; Apprill *et al.*, 2015; Parada *et al.*, 2016). These primers 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 postive control a mock community composed of evenly mixed DNA material originating from 20 bacterial strains (ATCC MSA-1002, ATCC, USA) was used. The sequences obtained in this study have been submitted to the European Nucleotide Archive (ENA) under accession numbers ?????.

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   query = col_character(),
##   reference = col_character()
## )

## See spec(...) for full column specifications.

## Parsed with column specification:
## cols(
##   taxlevel = col_double(),
##   rankID = col_character(),
##   taxon = col_character(),
##   daughterlevels = col_double(),
```

```
##   total = col_double(),
##   `40` = col_double(),
##   `41` = col_double(),
##   `42` = col_double(),
##   `43` = col_double(),
##   `61` = col_double(),
##   `62` = col_double(),
##   `63` = col_double(),
##   ATCC_1 = col_double()
## )
```

Obtained sequences were analyzed on the computer cluster Isabella (University Computing Center, University of Zagreb) using mothur (version 1.43.0)(Schloss *et al.*, 2009) according to the MiSeq Standard Operating Procedure (MiSeq SOP; https://mothur.org/wiki/MiSeq_SOP) (Kozich *et al.*, 2013) and recommandations 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 132; http://www.arb-silva.de) was used (Quast *et al.*, 2013; Yilmaz *et al.*, 2014). Sequences classified as chloroplsts by mothur were exported, aligned using SILVA incremental aligner (SINA, version 1.6.0) (Pruesse *et al.*, 2012) against the same SILVA SSU Ref NR 99 database (release 132) and imported into ARB (version 6.0.6) (Ludwig *et al.*, 2004) for further phylogentic analysis using the same database. Reference sequences close to imported were selected and used to calculate a phylogentic tree using the Maximum Likelihood algorithm RAxML (version 7.0.3) with 1000 tree replicates (Stamatakis, 2006). Imported partial chloroplast sequences were added to the tree using the maximum parsimony criteria and not allowing changes to tree topology. Pipeline data processing and visualizaion was done using R (version 3.6.0) (R Core Team, 2019), package tidyverse (version 1.2.1) (Wickham *et al.*, 2019) and multiple other packages (Xie, 2014, 2015, 2019a, 2019b; Xie *et al.*, 2018; Zhu,

2019; Allaire *et al.*, 2019). The detailed analysis procedure including the R Markdown file for this paper are available as a GitHub repository (?????). Based on the ATCC MSA-1002 mock community included in the analysis a sequencing error of 0.009 % was detemined that is in line with previously reported values for next-generation sequencing data (Kozich *et al.*, 2013; Schloss *et al.*, 2016). In addition, the negative control processed together with the samples yielded only 44,056 sequences.

Protein Isolation and Metaproteomics

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

Confocal Microscopy

Results and Discussion

Conclusions

Materials and Methods

References

Allaire, J.J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., et al. (2019) rmarkdown: Dynamic Documents for R.

Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews* **59**: 143–169.

Andersson, A.F., Riemann, L., and Bertilsson, S. (2010) Pyrosequencing reveals contrasting seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *The ISME journal* **4**: 171–181.

Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015) Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology* **75**: 129–137.

Burke, C., Kjelleberg, S., and Thomas, T. (2009) Selective extraction of bacterial DNA from the surfaces of macroalgae. *Applied and environmental microbiology* **75**: 252–256.

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011a) Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 14288–14293.

Burke, C., Thomas, T., Lewis, M., Steinberg, P., and Kjelleberg, S. (2011b) Composition, uniqueness and variability of the epiphytic bacterial community of the green alga Ulva australis. *The ISME journal* **5**: 590–600.

Cai, X., Gao, G., Yang, J., Tang, X., Dai, J., Chen, D., and Song, Y. (2014) An ultrasonic method for separation of epiphytic microbes from freshwater submerged macrophytes. *Journal of*

196 *basic microbiology* **54**: 758–761.

197 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., et al.
198 (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq
199 platforms. *The ISME Journal* **6**: 1621–1624.

200 Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., et al.
201 (2010) Direct Cellular Lysis/Protein Extraction Protocol for Soil Metaproteomics. *Journal of*
202 *Proteome Research* **9**: 6615–6622.

203 Egan, S., Harder, T., Burke, C., Steinberg, P., Kjelleberg, S., and Thomas, T. (2013) The
204 seaweed holobiont: understanding seaweed-bacteria interactions. *FEMS microbiology reviews* **37**:
205 462–476.

206 Gilbert, J.A., Field, D., Swift, P., Newbold, L., Oliver, A., Smyth, T., et al. (2009) The
207 seasonal structure of microbial communities in the Western English Channel. *Environmental*
208 *Microbiology* **11**: 3132–3139.

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

211 Hultman, J., Waldrop, M.P., Mackelprang, R., David, M.M., McFarland, J., Blazewicz, S.J., et
212 al. (2015) Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature*
213 **521**: 208–212.

214 Jiang, Y.-F., Ling, J., Dong, J.-D., Chen, B., Zhang, Y.-Y., Zhang, Y.-Z., and Wang, Y.-S.
215 (2015) Illumina-based analysis the microbial diversity associated with Thalassia hemprichii in
216 Xincun Bay, South China Sea. *Ecotoxicology (London, England)* **24**: 1548–56.

217 Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013)
218 Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon

sequence data on the MiSeq Illumina sequencing platform. *Applied and environmental microbiology* **79**: 5112–5120.

Longford, S., Tujula, N., Crocetti, G., Holmes, A., Holmström, C., Kjelleberg, S., et al. (2007) Comparisons of diversity of bacterial communities associated with three sessile marine eukaryotes. *Aquatic Microbial Ecology* **48**: 217–229.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004) ARB: a software environment for sequence data. *Nucleic Acids Research* **32**: 1363–1371.

Massana, R., Murray, A.E., Preston, C.M., and DeLong, E.F. (1997) Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. *Applied and environmental microbiology* **63**: 50–56.

Nõges, T., Luup, H., and Feldmann, T. (2010) Primary production of aquatic macrophytes and their epiphytes in two shallow lakes (Peipsi and Võrtsjärv) in Estonia. *Aquatic Ecology* **44**: 83–92.

Parada, A.E., Needham, D.M., and Fuhrman, J.A. (2016) Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology* **18**: 1403–1414.

Pruesse, E., Peplies, J., and Glöckner, F.O. (2012) SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics (Oxford, England)* **28**: 1823–1829.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* **41**: D590–6.

R Core Team (2019) R: A Language and Environment for Statistical Computing, Vienna,

241 Austria: R Foundation for Statistical Computing.

242 Richter-Heitmann, T., Eickhorst, T., Knauth, S., Friedrich, M.W., and Schmidt, H. (2016)
243 Evaluation of Strategies to Separate Root-Associated Microbial Communities: A Crucial Choice
244 in Rhizobiome Research. *Frontiers in Microbiology* **7**: 773.

245 Schloss, P.D., Girard, R.A., Martin, T., Edwards, J., and Thrash, J.C. (2016) Status of the
246 Archaeal and Bacterial Census: an Update. *mBio* **7**: e00201–16.

247 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al.
248 (2009) Introducing mothur: open-source, platform-independent, community-supported software
249 for describing and comparing microbial communities. *Applied and environmental microbiology*
250 **75**: 7537–7541.

251 Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses
252 with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.

253 Staufenberger, T., Thiel, V., Wiese, J., and Imhoff, J.F. (2008) Phylogenetic analysis of
254 bacteria associated with Laminaria saccharina. *FEMS Microbiology Ecology* **64**: 65–77.

255 Su, C., Lei, L., Duan, Y., Zhang, K.-Q., and Yang, J. (2012) Culture-independent methods
256 for studying environmental microorganisms: methods, application, and perspective. *Applied*
257 *microbiology and biotechnology* **93**: 993–1003.

258 Uku, J., Björk, M., Bergman, B., and Díez, B. (2007) Characterization and comparison of
259 prokaryotic epiphytes associated with three East African seagrasses. *Journal of Phycology* **43**:
260 768–779.

261 Weidner, S., Arnold, W., and Puhler, A. (1996) Diversity of uncultured microorganisms
262 associated with the seagrass Halophila stipulacea estimated by restriction fragment length
263 polymorphism analysis of PCR-amplified 16S rRNA genes. *Applied and environmental*

microbiology **62**: 766–771.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., et al. (2019) Welcome to the tidyverse. *Journal of Open Source Software* **4**: 1686.

Xie, Y. (2015) Dynamic Documents with {R} and knitr, 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC.

Xie, Y. (2014) knitr: A Comprehensive Tool for Reproducible Research in {R}. In, Stodden,V., Leisch,F., and Peng,R.D. (eds), *Implementing reproducible computational research*. Chapman; Hall/CRC.

Xie, Y. (2019a) knitr: A General-Purpose Package for Dynamic Report Generation in R.

Xie, Y. (2019b) tinytex: Helper Functions to Install and Maintain 'TeX Live', and Compile 'LaTeX' Documents.

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

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014) The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic acids research* **42**: D643–8.

Zhu, H. (2019) kableExtra: Construct Complex Table with 'kable' and Pipe Syntax.