Dynamics of Epiphytic Microbial Communities on 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

2 Introduction

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*Cymodocea nodosa* meadow in the Bay of Saline decayed to an extent that no leaves could be retrieved (Najdek *et al.*, unpublished data).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

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 mg ml^{-1}) and the mixture was incubated at 37 °C for 30 min. Subsequently, proteinase K (final concentration 0.5 mg ml^{-1}) and SDS (final concentration 1 %) were added and the samples were incubated at 55 °C for 2 h. Following the

5

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.

Illumina 16S rRNA Sequencing

An aliquot of isolated DNA was treated with RNase A (final concentration 200 µg 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 µl^{-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

6

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. The sequences obtained in this study

have been submitted to the European Nucleotide Archive (ENA) under accession numbers **TO BE**

ADDED LATER!.

Sequence Analysis

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 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 132; http://www.arb-silva.de) was used (Quast *et al.*,

2013; Yilmaz *et al.*, 2014). Sequences classified as chloroplasts 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 phylogenetic analysis using the same database. Reference sequences close to imported ones were selected and used to calculate a phylogenetic 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 visualization 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 (**TO BE ADDED LATER!**). Based on the ATCC MSA-1002 mock community included in the analysis a sequencing error rate of 0.01 % was determined, which 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 2 sequences after sequence quality curation.

Protein Isolation

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 $10,000 \times g$ and 4 °C for 40 min. The

8

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 samples were centrifuged at 20,000 × g and 4 °C for 5 min. Dried pellets were stored at −80 °C until further analysis.

Metaproteomics

Isolated proteins were whole trypsin digested using the FASP (filter-aided sample preparation) Protein Digestion Kit (Expedeon, UK) according to the manufacturer's instructions (Wiśniewski *et al.*, 2009) with small modifications. Before loading the solution to the column, protein pellets were solubilized in a urea sample buffer included in the kit amended with DTT (final concentration 100 mM) for 45 min at room temperature and centrifuged at 20,000 × g for 2 – 5 min at room temperature to remove larger particles. The first washing step after protein solution loading was repeated twice. In addition, centrifugation steps were prolonged if the column was clogged. Trypsin digestion was performed on column filters at 37 °C overnight for 18 h. The final filtrate containing digested proteins was acidified with 1 % (final concentration) trifluoroacetic acid, freezed at −80 °C for 15 min, lyophilized and sent to the Vienna Metabolomics Center (University of Vienna) for metaproteomic analysis. Peptides were resuspended in 1 % (final concentration) trifluoroacetic acid, desalted using the Pierce C18 Tips (Thermo Fisher Scientific, USA) according to the manufacturer's instructions and sequenced on a Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific, USA). Obtained MS/MS spectra were searched against a protein database from metagenomic assembly published in Burke *et al.* (2011) using SEQUEST-HIT engines and validated with Percolator in Proteome Discoverer 2.1 (Thermo Fisher Scientific, USA). The target-decoy approach was used to reduce the probability of false peptide identification. Results whose false discovery rate at the peptide level was <1 % were kept. For protein identification a minimum of of two peptides and one unique peptide were required. For protein quantification, a chromatographic peak area-based free quantitative method was applied.

Confocal Microscopy

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

Results

Discussion

Acknowledgements

References

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

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., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011) 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.

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

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

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

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013) 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.

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.

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, Austria: R Foundation for Statistical Computing.

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

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

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

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., et al. (2019)

12

Welcome to the tidyverse. *Journal of Open Source Software* **4**: 1686.

Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nature Methods* **6**: 359–362.

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.

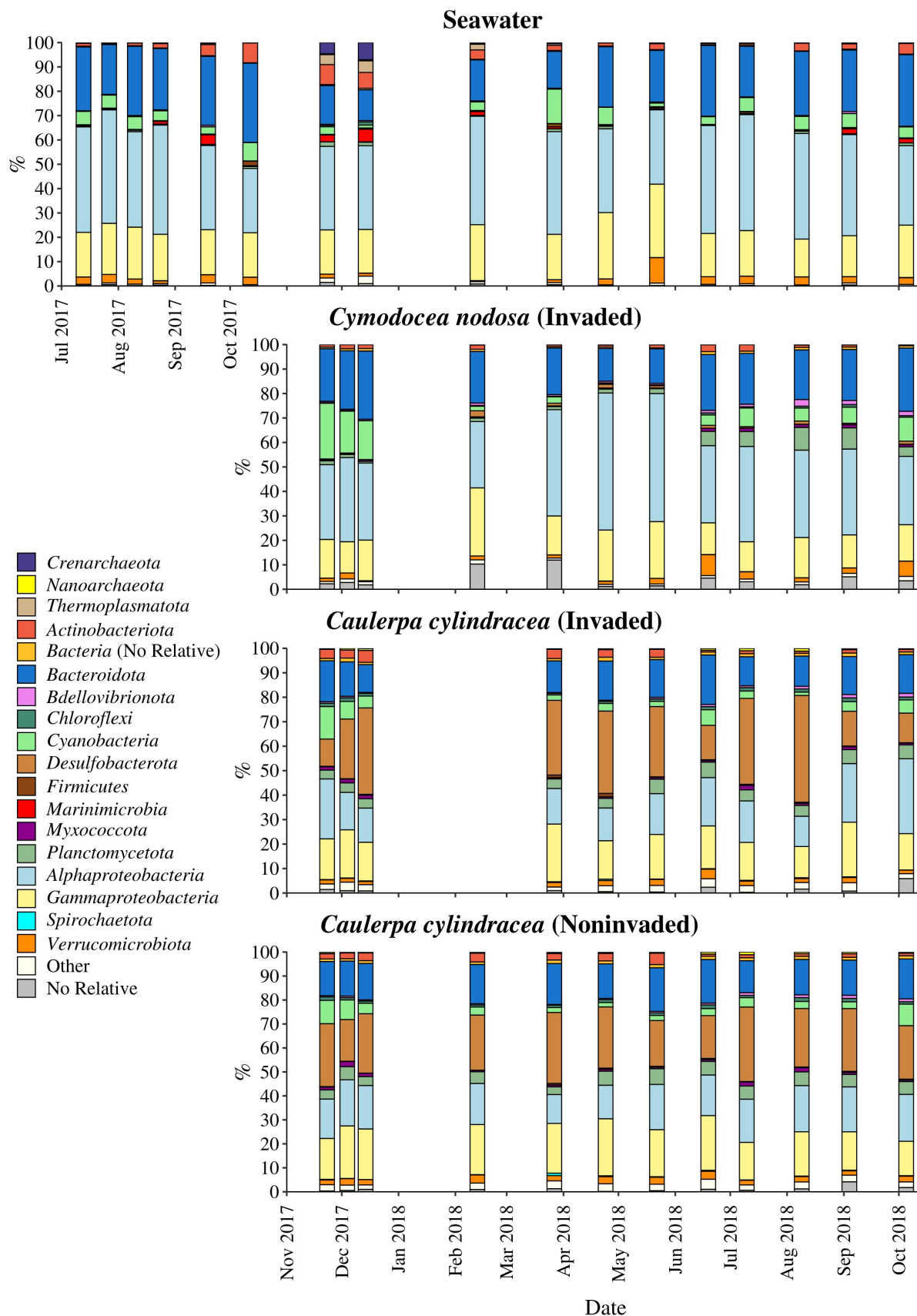Xie, Y., Allaire, J.J., and Grolemund, G. (2018) R Markdown: The Definitive Guide, Boca Raton, Florida: Chapman; Hall/CRC.

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.

Figure Captions

Figure 1. Taxonomic classification and relative contribution of the most abundant bacterial sequences.

Figures

Figure 1. Taxonomic classification and relative contribution of the most abundant bacterial sequences.

16