DNA extraction bias is more pronounced for microbial eukaryotes than for prokaryotes

Mia Maria Bengtsson[1] and Anne Brauer[2]

[1]Affiliation not available
[2]Institute of Microbiology, Institute of Microbiology, University of Greifswald, University of Greifswald

February 22, 2024

DNA extraction bias is more pronounced for microbial eukaryotes than for prokaryotes

Anne Brauer
Institute of Microbiology, University of Greifswald

Corresponding author:
Mia M. Bengtsson*
Institute of Microbiology, University of Greifswald
&
Institute of Marine Biotechnology

Address:
Institute of Microbiology, University of Greifswald
Felix-Hausdorff-Straße 8
17489 Greifswald
Germany

Tel: +49 (0)3834 420 5918
Fax: +49 (0)3834 420 5909
*mia.bengtsson@uni-greifswald.de
https://orcid.org/0000-0002-2115-9139

Running head: Extraction bias in microbial eukaryotes

SUMMARY

DNA extraction and preservation bias is a recurring topic in DNA sequencing-based microbial ecology. Different methodologies can lead to distinct outcomes, which has been demonstrated especially in studies investigating prokaryotic community composition. Eukaryotic microbes are ubiquitous, diverse, and increasingly a subject of investigation in addition to bacteria and archaea. However, little is known about how the choice of DNA preservation and extraction methodology impacts perceived eukaryotic community composition. In this study, we compared the effect of two DNA preservation methods and 6 DNA extraction methods on the community profiles of both eukaryotes and prokaryotes in phototrophic biofilms on seagrass (Zostera marina) leaves from the Baltic Sea. We found that, whereas both DNA preservation and extraction method caused significant bias in perceived community composition for both eukaryotes and prokaryotes, extraction bias was more pronounced for eukaryotes than for prokaryotes. Especially soft-bodied or hard-shelled eukaryotes like nematodes and diatoms, respectively, were differentially abundant depending on the extraction method. We conclude that careful consideration of DNA preservation and extraction methodology is crucial to achieving representative community profiles of eukaryotes in marine biofilms, and likely all other habitats containing diverse eukaryotic microbial communities.

INTRODUCTION

Advances in sequencing technology and paradigm shifts in microbial ecology have led to a prolific rise in studies that use metagenomics and marker gene PCR amplicon sequencing to assess microbial communities in various environments. Essential to all of these efforts is the preservation and extraction of DNA from environmental or host-associated microbial communities. It is well-known that the choice of DNA preservation and extraction method can impact the perceived relative abundance of microbial taxa in microbial communities (e.g., Martin-Laurent et al. 2001). Differences in community composition depending on the DNA extraction method are referred to as extraction bias, which can have various causes, many of which are linked to the ability to lyse microbial cells (Koid et al. 2012). A wide variety of commercial kits and custom protocols have been developed to provide representative and reproducible DNA extraction from different sample types. For some environments, extraction bias has been evaluated by comparing the outcome of different extraction protocols, in some

cases leading to general recommendations on method choice (e.g., Albertsen et al. 2015, Weber et al. 2017). A majority of existing studies have focused on prokaryotic communities, reflecting an emphasis on bacteria and archaea in molecular microbial ecology.

However, in most natural environments, microbial eukaryotes are abundant, diverse, and play essential roles in ecosystem processes. Whereas they have traditionally been studied using microscopic methods, studies using molecular methods have revealed novel taxa that escape microscopic detection or identification (Liu et al. 2009, Jones et al. 2011). In the wake of numerous influential studies on prokaryote diversity in various ecosystems, microbial eukaryotes are receiving renewed attention by taking advantage of available high-throughput sequencing technologies (Lima-Mendez et al. 2015, Delmont et al. 2022).

Due to a high diversity of cell envelopes found in microbial eukaryotes, ranging from single membranes in amoeboid protists to silica frustules of diatoms or thick cellulose cell walls of green algae, effective cell lysis and subsequent DNA recovery pose unique challenges. Despite this, extraction bias has so far received little attention in surveys of microbial eukaryotes (but see Vesty et al. 2017, Koid et al. 2012, Santos et al. 2015, Donn et al. 2007, Mäki et al. 2017). In addition, microbial eukaryotes and prokaryotes are intermingled in most microbial communities, and extraction methods that recover DNA well from a variety of eukaryotes and prokaryotes are needed to achieve an accurate representation of microbial community composition.

Here, we compared the effect of different popular commercial and custom DNA extraction methods on the perceived community composition of prokaryotes and eukaryotes in marine phototrophic biofilms growing on seagrass leaves. We aimed to assess whether extraction bias affects microbial eukaryotes and prokaryotes at a similar magnitude in the same environment and whether this bias depends on the sample preservation method.

Phototrophic biofilms are known to be complex microbial ecosystems including members of all three domains of life, encompassing several trophic levels (Bengtsson et al. 2018). This is a property that they share with many other microbial habitats, including soils, sediments, and plankton. Biofilm material from leaves of the seagrass *Zostera marina* was rubbed off with a cotton swab. We used two different methods to preserve the DNA in the biofilms prior to extraction: Biofilms were either suspended in sterile seawater, pelleted by centrifugation, frozen in liquid N_2, and stored at -20°C, or they were suspended in RNAlater, pelleted and stored at +4°C. To ensure comparable results, the different extraction methods

started with pellets (in triplicate) of similar mass from the same suspension (one sterile seawater suspension and one RNAlater suspension). The 6 different extraction methods that were tested (summarized and detailed in Table S1) varied in lysis method (5 mechanical vs. 1 enzymatic), lysing matrix, and intended sample material (soil, biofilm, general). We used Illumina MiSeq sequencing of amplicons of SSU rRNA gene fragments of prokaryotes (16S rRNA) and eukaryotes (18S rRNA) to assess the microbial community composition of the biofilms (see the supplementary material for detailed descriptions of extraction methods and sequencing).

RESULTS AND DISCUSSION

Extraction bias is more pronounced for eukaryotes than for prokaryotes

The extraction method explained a significant amount of variation (PERMANOVA $p < 0.05$) in both eukaryotes and prokaryotes, confirming the presence of extraction bias for both groups (Fig. 1). However, extraction bias was more pronounced for eukaryotes (22.7 % of variation explained, $p < 0.01$) than for prokaryotes (15.3 % of variation explained, $p < 0.05$). Two of the tested extraction methods, the InnuSpeed Soil DNA kit (Analytic Jena; referred to as InnuSpeed) and the QuickDNA Universal kit (Zymo Research; referred to as QuickDNA) gave rise to more distinct eukaryote community compositions compared to the other four methods, especially for seawater-suspended biofilms (Fig. 1a). These two methods were characterized by more gentle lysis conditions, weak bead beating (smaller beads than in the other tested methods, see table S1) and enzymatic lysis, respectively, compared to the other methods that use harsh bead beating, indicating that incomplete lysis of some eukaryotic cells may underlie the observed pattern. However, when investigating which eukaryotic taxa were differentially abundant in these methods, we found that metazoans, especially nematodes and annelids, and rhizarian (*Cercozoa*) amplicon sequence variants (ASVs) were overrepresented in samples from the QuickDNA method compared to the PowerSoil DNA isolation kit (MoBio; referred to as PowerSoil) (Fig. 2e), a representative example of the methods based on mechanical lysis. Nematodes and annelids are generally soft-bodied, and should therefore not require harsh mechanical lysis for DNA recovery. Hence, their overrepresentation in the QuickDNA method may in part reflect a higher recovery of PCR-amplifiable nematode DNA, perhaps due to selective fragmentation of nematode DNA in the other, mechanical lysis-based, methods. In contrast, several diatom sequence variants were

underrepresented in samples extracted with the QuickDNA method (Fig. 2e). Indicating that enzymatic lysis might inefficiently lyse their silica frustules. This result was also supported by an underrepresentation of diatom plastid sequence variants (16S rRNA, fig. 2f) with the QuickDNA method, while *Rubritaleaceae* ASVs (*Verrucomicrobia*) were overrepresented. With the InnuSpeed kit *Polychaeta* (*Metazoa*) and *Cercozoa* (*Rhizaria*) ASVs were overrepresented, while diatom ASVs and some nematode (*Metazoa*) ASVs were underrepresented (Fig. 2c). For example, also an ASV classified as *Halomonhystera disjuncta* (nematode), which was overrepresented in the QuickDNA method. Several diatom plastid sequences were underrepresented with the Innu Speed kit, indicating that the weak bead-beating was not sufficient to completely lyse the silica frustules (Fig. 2d).

The preservation method has a stronger influence on community composition than the extraction method

Preservation protocol was the strongest explanatory variable for both prokaryotic (33.1 % of variation explained, $p<0.05$) and eukaryotic communities (33.9 % of variation explained, $p<0.01$) illustrated by a clear separate clustering of RNAlater- and seawater-suspended samples in the nMDS ordinations (Fig. 1). Preservation bias affected mainly Diatoms, *Alveolata*, *Cnidaria* and *Bacillariophyta* which were overrepresented in the RNAlater preserved samples, while Nematodes, *Cercozoa* and *Rubritaleaceae* (*Verrucomicrobia*) were underrepresented (Fig. 2a&b). A possible cause could be different GC content of DNA in the different organisms, as Gray et al. (2013) showed that bacteria with a high GC content are poorly recovered from samples conserved with RNAlater. However, the overall community composition pattern remained comparable (Fig S4).

DNA yield does not impact community composition

The DNA yield differed significantly among extraction methods (Kruskal-Wallis rank-sum test, $p < 0.05$), with the highest DNA yields observed for the PowerSoil and DNASpin kits in the seawater-suspended samples (Figure S1). The QuickDNA kit was the only one that resulted in a higher yield on RNAlater preserved than flash-frozen samples. DNA yield did not significantly explain variation in perceived community composition across prokaryotic and eukaryotic samples (Permanova, $p> 0.2$ and $p > 0.05$, respectively), indicating that factors that affect overall yield are different from those giving rise to DNA extraction bias. This is reassuring since

extraction yield can vary substantially even between replicate samples under the same extraction method (see e.g., PowerBiofilm method, fig. S1), but this does not compromise the reproducibility of community composition patterns (Vishnivetskaya et al. 2014).

CONCLUSIONS

Most microbial DNA extraction methods have been developed and optimized for prokaryotes and may therefore be inadequate for microbial eukaryotes which have a high diversity of cell envelopes posing unique challenges for effective cell lysis and subsequent DNA recovery. It is unlikely that we will ever arrive at one optimal methodology that captures all organism groups without bias. It is also not the aim of this study to offer specific recommendations for DNA preservation or extraction methods. However, in the light of our results, we recommend that the extraction and preservation method should be chosen carefully depending on the specific groups of interest in the focal ecosystem. For example, soft-bodied eukaryotes such as nematodes may benefit from more gentle enzymatic lysis methods while the tough silica frustules of diatoms may require mechanical lysis.

Acknowledgments

Thanks to Janina Brakel, Thorsten Reusch, and Florian Weinberger for access to sample material and Tim Urich for generosity. This study was funded by a stipend from the German Federal Environmental Foundation to A. Brauer and the SeaStore project (BMBF 03F0859C).

Data availability statement

The datasets generated and analyzed during the current study are available at https://doi.org/10.13140/RG.2.2.28409.54888. DNA sequences generated are available in the NCBI short read archive under the project number PRJNA389390 and accession numbers SRX29110 92 & 93, SRX29111 20-29, 50-59, 76-79 & 90-99 for eukaryotes, and SRX29110 98 & 99, SRX29111 00-19, 40-49 & 60-69 for prokaryotes.

References

Albertsen, M., Karst, S.M., Ziegler, A.S., Kirkegaard, R.H., and Nielsen, P.H. (2015) Back to Basics – The Influence of DNA Extraction and Primer Choice on Phylogenetic Analysis of Activated Sludge Communities. PLOS ONE 10: e0132783.

Bengtsson, M.M., Wagner, K., Schwab, C., Urich, T., and Battin, T.J. (2018) Light availability impacts structure and function of phototrophic stream biofilms across domains and trophic levels. Molecular Ecology 27: 2913–2925.

Delmont, T.O., Gaia, M., Hinsinger, D.D., Fremont, P., Vanni, C., Guerra, A.F., et al. (2022) Functional repertoire convergence of distantly related eukaryotic plankton lineages revealed by genome-resolved metagenomics.Cell Genomics (in press).

Donn, S., Griffiths, B.S., Neilson, R., and Daniell, T.J. (2008) DNA extraction from soil nematodes for multi-sample community studies. Applied Soil Ecology 38: 20–26.

Gray, Michael A., Pratte, Zoe A., and Kellogg , Christina A. (2013) Comparison of DNA preservation methods for environmental bacterial community samples. FEMS Microbiology Ecology 83: 468–477.

Jones, M.D.M., Forn, I., Gadelha, C., Egan, M.J., Bass, D., Massana, R., and Richards, T.A. (2011) Discovery of novel intermediate forms redefines the fungal tree of life. Nature 474: 200–203.

Koid, A., Nelson, W.C., Mraz, A., and Heidelberg, K.B. (2012) Comparative Analysis of Eukaryotic Marine Microbial Assemblages from 18S rRNA Gene and Gene Transcript Clone Libraries by Using Different Methods of Extraction. Appl Environ Microbiol 78: 3958–3965.

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., et al. (2015) Determinants of community structure in the global plankton interactome. Science 348: 1262073.

Liu, H., Probert, I., Uitz, J., Claustre, H., Aris-Brosou, S., Frada, M., et al. (2009) Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. Proc Natl Acad Sci USA 106: 12803–12808.

Mäki, A., Salmi, P., Mikkonen, A., Kremp, A., and Tiirola, M. (2017) Sample Preservation, DNA or RNA Extraction and Data Analysis for High-Throughput Phytoplankton Community Sequencing. Front Microbiol 8:.

Martin-Laurent, F., Philippot, L., Hallet, S., Chaussod, R., Germon, J.C., Soulas, G., and Catroux, G. (2001) DNA Extraction from Soils: Old Bias for New Microbial Diversity Analysis Methods. Applied and Environmental Microbiology 67: 2354–2359.

Santos, S.S., Nielsen, T.K., Hansen, L.H., and Winding, A. (2015) Comparison of three DNA extraction methods for recovery of soil protist DNA. Journal of Microbiological Methods 115: 13–19.

Vesty, A., Biswas, K., Taylor, M.W., Gear, K., and Douglas, R.G. (2017) Evaluating the Impact of DNA Extraction Method on the Representation of Human Oral Bacterial and Fungal Communities. PLOS ONE 12: e0169877.

Vishnivetskaya, T.A., Layton, A.C., Lau, M.C.Y., Chauhan, A., Cheng, K.R., Meyers, A.J., et al. (2014) Commercial DNA extraction kits impact observed microbial community composition in permafrost samples. FEMS Microbiology Ecology 87: 217–230.

Weber, L., DeForce, E., and Apprill, A. (2017) Optimization of DNA extraction for advancing coral microbiota investigations. Microbiome 5: 1–14.

Figures

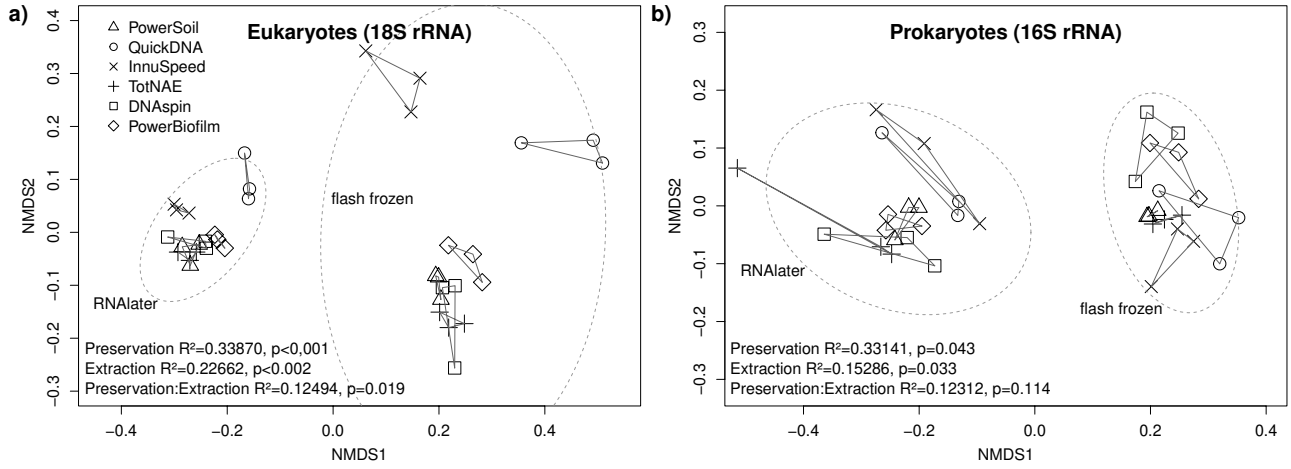

Figure 1: Comparison of communities of epibiotic microbial eukaryotes (a) and prokaryotes (b) on *Zostera marina* treated with different DNA preservation protocols and DNA extraction methods. Biofilm material from leaves of the seagrass *Z. marina* was rubbed off with a cotton swab and was suspended in sterile seawater, pelleted by centrifugation, frozen in liquid nitrogen, and stored at -20°C (flash-frozen) or suspended in RNAlater, pelleted and stored at +4°C (RNAlater). The 6 different extraction methods (different shapes) that were tested are summarized and detailed in Table S1. Samples were extracted in triplicates and 16S- and 18S rRNA genes were amplified and sequenced with Illumina MiSeq technology. Non-metric multidimensional scaling (nMDS) ordinations based on Bray-Curtis distances were calculated from Hellinger transformed sequence variant counts, dashed lines indicate the 95 % confidence interval of the factor preservation method.

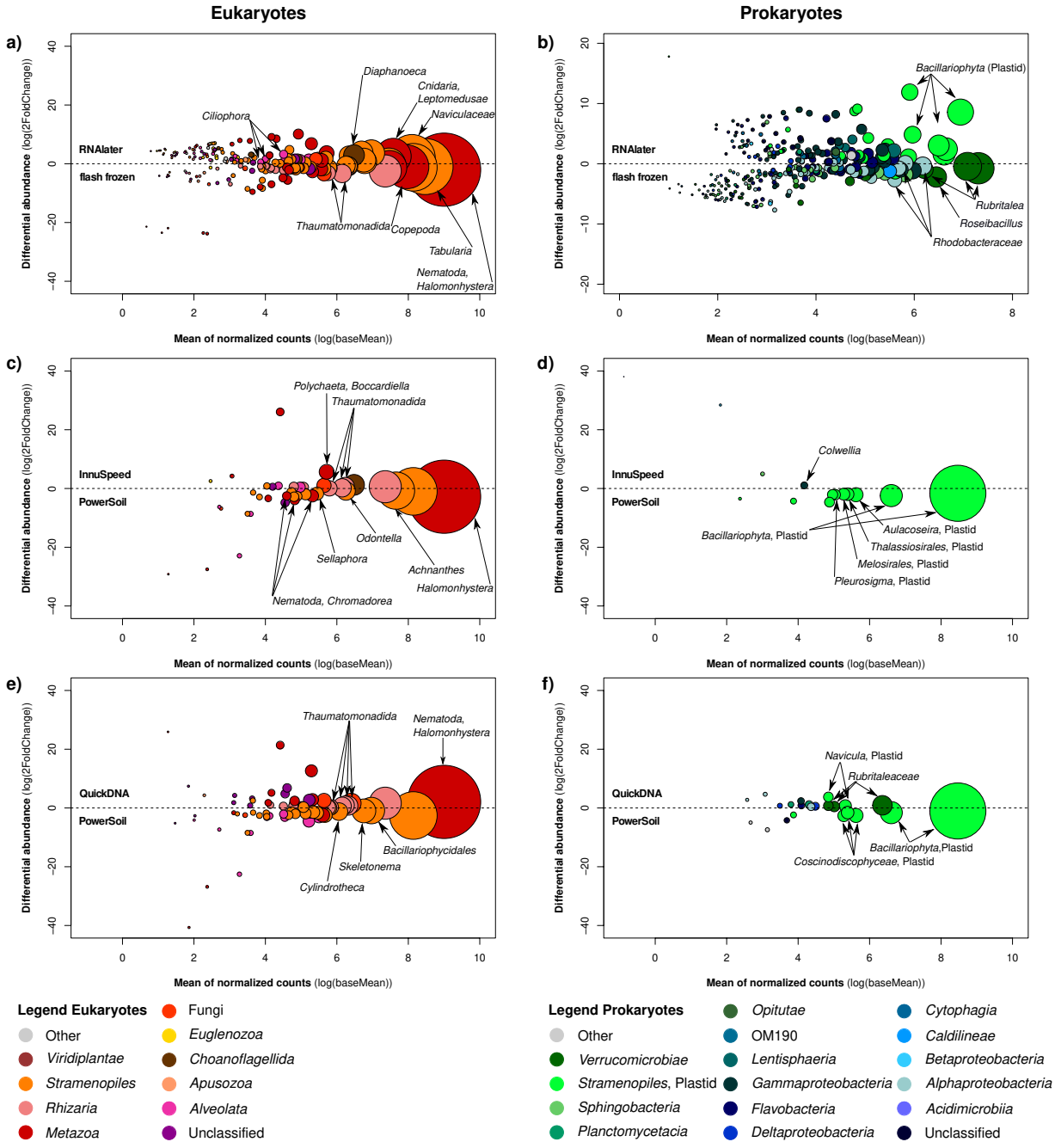

Figure 2: Significantly differentially abundant taxa (ASVs, p<0.01 are shown) in the epibiotic microbial eukaryotic (a, c, e) and prokaryotic (b, d, f) communities on *Zostera marina* treated with the two different preservation (a and b) or selected DNA extraction methods (c-f) as detected by DeSeq2 parametric Wald test. Point diameter is scaled by the abundance of the ASVs. c&d) communities extracted by the InnuSpeed method compared to the PowerSoil method. e&f) communities extracted by the QuickDNA method compared to the PowerSoil method. Taxa names on arrows indicate the finest taxonomic resolution for selected ASVs. Selected pairwise comparisons are shown here, see Fig. S5 for the remaining comparisons.