



**HAL**  
open science

# Performance of the melting seawater-ice elution method on the metabarcoding characterization of benthic protist communities

Albert Reñé, Adrià Auladell, Guillaume Reboul, David Moreira, Purificación López-García

## ► To cite this version:

Albert Reñé, Adrià Auladell, Guillaume Reboul, David Moreira, Purificación López-García. Performance of the melting seawater-ice elution method on the metabarcoding characterization of benthic protist communities. *Environmental Microbiology Reports*, 2020, 12 (3), pp.314-323. 10.1111/1758-2229.12834 . hal-02988536

**HAL Id: hal-02988536**

**<https://hal.science/hal-02988536>**

Submitted on 4 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Performance of the melting seawater-ice elution method on the metabarcoding characterization of benthic protist communities

Albert Reñé 1,2\*, Adrià Auladell 2, Guillaume Reboul 1, David Moreira 1, Purificación López-García 1

1 Unité d'Ecologie, Systématique et Evolution, CNRS, Université Paris-Sud, AgroParisTech, Université Paris-Saclay, 91400 Orsay, France

2 Departament de Biologia Marina i Oceanografia, Institut de Ciències del Mar (CSIC), Barcelona, Catalonia, Spain

### \* Corresponding author:

Albert Reñé

Institut de Ciències del Mar

Pg. Marítim de la Barceloneta, 37-49

08003 Barcelona (Spain)

Tel: +34932309500

Fax: +34932309555

[albertrene@icm.csic.es](mailto:albertrene@icm.csic.es)

**Running title:** Benthic protist metabarcoding after seawater-ice elution methods

## ORIGINALITY-SIGNIFICANCE STATEMENT

While metabarcoding studies to characterize the diversity of planktonic protists are numerous, those focused on benthic protists are scarce and require specific methods. Assessing the performance of such methods to characterize benthic communities while limiting biases is therefore required. The present study provides information about the efficiency and influence of the melting seawater-ice elution method for 18S rRNA gene metabarcoding analysis of benthic protist communities, as compared with the most commonly used direct DNA purification from sediments. The observed differences between the two approaches will contribute to decision-making in future studies, providing elements to choose the most appropriate method depending on the objective and target organisms.

### Summary

Massive amplicon sequencing approaches to characterize the diversity of microbial eukaryotes in sediments are scarce and controls about the effects introduced by different methods to recover DNA are lacking. In this study, we compare the performance of the melting seawater-ice elution method on the characterization of benthic protist communities by 18S rRNA gene metabarcoding with results obtained by direct cell lysis and DNA purification from sediments. Even though the most abundant operational taxonomic units were recovered by both methods, eluted samples yielded higher richness than samples undergoing direct lysis. Both treatments allowed recovering the same taxonomic groups, although we observed significant differences in terms of relative abundance for some of them. Dinoflagellata and Ciliophora strongly dominated the community in eluted samples (> 80% reads). In directly lysed samples, they only represented 37%, while groups like Fungi and Ochrophytes were highly represented (> 20% reads respectively). Our results show that the elution process yields a higher protist richness estimation, most likely as a result of the higher sample volume used to recover organisms as compared to commonly used volumes for direct benthic DNA purification. Motile groups, like dinoflagellates and ciliates, are logically more enriched during the elution process.

## Introduction

Studies on marine benthic protists have traditionally focused on the characterization of the diversity, distribution and function in ecosystems of morphological species based on traditional microscopy observations and cell counting (Mare, 1942; Dragesco, 1965; Fenchel, 1969). However, these studies are much scarcer than those from planktonic organisms due to difficulties in collecting, analysing and most notably, separating the cells from sediments, making it difficult to quantify them (Bak and Nieuwland, 1989). Methodologies to separate cells from soil and sediment have been developed over decades. Density gradient centrifugation has been tested to separate bacteria (Courtois et al., 2001) and protists (Starink et al., 1994) from substrate using different media. Depending on the sediment type, the cell recovery is usually high, but several limitations and uncertainties exist, like the recovery rate, biases in the recovered groups or possible adverse effects on the integrity of living cells (Robe et al., 2003; Parent et al., 2018). Some other methods to separate cells from substrate involve suspension of sediment in filtered seawater, followed by successive filtration through mesh nets of specific size pores to remove the sediments and recover and concentrate the organisms, or yet placing coverslips on top of the sediment and recovering the organisms that attach to them (Webb, 1956). However, those methods do not fully remove remaining sediment in the final sample and might result in a low and biased recovery of cells. Also frequently used, the traditional seawater ice 'Uhlig' method consists of melting seawater ice on top of a tube filled with sediment; upon melting, organisms that flow down accumulate in a Petri dish (Uhlig, 1964). Even though the recovery of cells using this method is reputed to be relatively low, it is commonly used in taxonomical studies focused on some specific groups of protists like ciliates, dinoflagellates, diatoms and other groups of flagellates (Saburova et al., 1995; Azovsky et al., 2013; Hoppenrath et al., 2014).

Since the early 2000s, traditional methods used to characterize protist communities, like microscopy, have been complemented and largely displaced by molecular methods based on the use of conserved gene markers, which sidestep many difficulties associated with morphological identification (Díez et al., 2001; López-García et al., 2001; Moonvan der Staay et al., 2001). Currently, 18S rRNA gene metabarcoding using high-throughput sequencing (HTS) techniques provides a fast, cost effective and highly sensitive method for characterizing protist diversity in natural samples (Logares et al., 2012). These metabarcoding approaches are being widely applied to marine planktonic protist communities, providing insights in their diversity, composition, spatial distribution (at global or local scale) and temporal dynamics (de Vargas et al., 2015; Massana et al., 2015; Malviya et al., 2016; Piredda et al., 2016). However, studies characterizing benthic protist communities using metabarcoding are still scarce (Chariton et al., 2010; Quaiser et al., 2011; Bik et al., 2012; Gong et al., 2015; Forster et al., 2016; Pan et al., 2019; Salonen et al., 2019), and biases likely higher. Indeed, while genomic DNA from plankton is usually obtained from biomass retained after filtering large seawater volumes (usually litres), DNA from benthic samples is usually obtained with a direct-lysis of cells from a relatively low sediment volume or mass. Furthermore, in addition to cell lysis and DNA purification as variability sources in assessing microbial community composition, soils and marine sediments can contain detrimental amounts of potential inhibitors for downstream molecular analyses. Indeed, direct extraction methods provide higher DNA yields but lower purity, while indirect methods, which require a previous specific sample treatment to separate cells from sediment, provide lower DNA yields but of higher purity, although it is time consuming and might induce biases in microbial community characterization (Steffan et al., 1988; Robe et al., 2003). The yield of different DNA extraction methods from sediments, as well as the impact on inferences of protist diversity and community composition, has been previously assessed for cloning libraries or denaturing gel gradient electrophoresis (Lekang et al., 2015 and references therein). Some studies have focused on the effect of other factors that can greatly influence richness, sample dispersion and the structure of microbial communities. These include using different soil sample sizes (Penton et al., 2016), increasing DNA extraction replicates of marine sediments (Lanzén et al., 2017) or increasing the sequencing efforts

and the number of polymerase chain reaction (PCR) replicates (Smith and Peay, 2014). Bacterial diversity studies comparing direct (direct lysis from soil) or indirect (previous cell separation using Nycodenz gradient) treatments yielded similar results for the two methods when using similar amounts of soil (Courtois et al., 2001; Delmont et al., 2011a,b). In any case, the standard methodology to obtain genomic DNA from sediment and soil consists of the direct cell lysis and DNA extraction from small amounts of sample (e.g., < 1 g) (Salonen et al., 2019), even though it has been proved that larger sample sizes provide a better capture of total diversity (Delmont et al., 2011a; Penton et al., 2016; Nascimento et al., 2018). Despite of the existing literature, the performance of the seawater-ice elution method for metabarcoding purposes has never been evaluated before. This study aims to determine how this sample treatment affects the inferred richness and composition of marine benthic protist communities. To explore this, we characterized by 18S rRNA gene metabarcoding protist communities of coastal sediment samples from the Mediterranean Sea by applying two different sample treatments: DNA purified after direct cell lysis in sediments, hereafter referred to as 'direct-lysis' samples, and DNA purified from cells separated from sediment using the 'Uhlrig method', hereafter referred as 'eluted' samples.

## Results and discussion

In order to test the performance of the melting seawater-ice (Uhlrig) elution method to study the diversity of benthic protists, we carried out 18S rRNA gene metabarcoding analysis on a total of 72 samples issued from 18 sampling outings at different dates and three localities in the NW Mediterranean Sea (Supporting Information Table S1). Sediment from each sampling trip was subjected to the two treatments (elution and direct lysis) with two replicates per treatment, resulting in 72 samples (see Experimental procedures section in Supporting Information). We then generated 18S rRNA gene amplicons of approximately 550 bp encompassing the hypervariable V4 region using broadrange primers for microbial eukaryotes and sequenced them (MiSeq Illumina). Thirteen of the 72 samples issued from the direct-lysis method did not yield amplicons and were subjected to re-amplification by semi-nested PCR (Supporting Information Table S1). After quality trimming, clustering of sequences in 'swarm' operational taxonomic units (OTUs), removal of singletons and exclusion of amplicons not corresponding to protists (Supporting Information), the first inspection of the read abundances and OTU composition showed that 10 of those samples were composed of very few OTUs, mostly belonging to Fungi. We interpret this as the result of very little protist biomass per volume unit in these samples and subsequent nested-PCR-associated biases. Consequently, those samples were removed from the data set for further comparative analyses, and their counterpart replicates treated under elution method were also removed in order to have the same number of samples for each treatment. We thus retained 26 samples that corresponded to DNA purified from small sediment volumes by direct lysis, and 26 samples that corresponded to DNA obtained from protists eluted from larger sediment volumes using the melting seawater ice method, that is, a total of 52 samples. In summary, those samples yielded 3,609,403 reads clustered into 12,518 OTUs. Samples corresponding to eluted samples yielded 1,986,751 reads and 10,447 OTUs and those corresponding to direct-lysis yielded 1,622,652 reads and 3,598 OTUs. To avoid biases in some analyses introduced by the comparison of sequence data sets of different size, we rarefied our sequence data sets when needed to the minimum number of reads observed in a sample (22,056 reads), resulting in a global data set of 1,146,912 reads and 10,142 OTUs.

## Effects on the determination of the community richness

Regarding community richness, rarefaction curves showed that the diversity estimated from direct-lysis sediment samples completely saturated, while eluted samples appeared near saturation (Fig. 1A). By contrast, species accumulation curves with the addition of samples were not saturated, showing that an increase in the sampling effort would increase the observed richness (Supporting Information Fig. S1). The evaluation of OTU abundance distribution showed some OTUs comprising most reads, and many

'rare' OTUs comprising a low number of reads in eluted samples. Direct-lysis sediment samples also showed some dominant OTUs, but in comparison, the number of 'rare' OTUs was much lower (Fig. 1B). We observed higher OTU richness in eluted samples than in direct-lysis ones (ANOVA  $F_{1,50} = 88.34$ ;  $p < 0.0001$ ) (Fig. 1C), and in Chao1 index ( $F_{1,50} = 113.4$ ;  $p < 0.0001$ ). However, they showed no significant differences when comparing Shannon ( $p > 0.1$ ), and Simpson ( $p > 0.5$ ) indexes (Fig. 1E and F), pointing out that even though eluted samples had higher richness, many OTUs were represented by a low number of reads. Even though DNA recovered using elution method does not correspond to all organisms present in the sediment volumes, but only to those successfully eluted, all differences observed in alpha-diversity among treatments could be attributed to the diverging initial sample volume used, with sample size of  $\sim 80 \text{ cm}^3$  for eluted and  $\sim 1 \text{ cm}^3$  for direct-lysis sediment samples, thus being close to 1:100 between both treatments.

Analysis of beta-diversity (Supporting Information Fig. S2) showed the dispersion of samples differed between treatments (ADONIS  $R = 0.11$ ,  $p < 0.001$ ; BETADISPER  $F_{1,50} = 7.18$ ,  $p < 0.01$ ), and those obtained by elution showed lower dispersion among them than the obtained by direct lysis in the multivariate dispersion analysis (MVDISP,  $\text{IMD} = 0.562$ , dispersion eluted = 0.72; direct lysis = 1.28). All sample replicates clustered together. However, eluted replicates showed lower dissimilarity among them than replicates of direct-lysis samples (Fig. 2).

The used seawater-ice method does not recover all organisms present in the sample, and thus, not all DNA is eluted from the sediment volume used. However, the higher richness obtained in eluted samples confirms a higher capture of diversity when larger sample volumes are used. Given that the sequencing depth was the same for both treatments, this suggests that differences in OTU richness and saturation were due to the different volume of sample used and highlights the impact of this parameter in determining the diversity of benthic communities. For this purpose, the elution process represents around 1–2 h of time (the needed for the seawater-ice to melt and the posterior filtration of eluted sample), and does not imply remarkable extra costs than using standard direct-lysis methods.

Direct-lysis sediment samples showed a higher dispersion than eluted ones, and replicates from eluted samples, which were obtained from two different sediment cores, showed higher similarity among them than replicates obtained from subsamples from the same sediment core in samples from direct lysis. This reflects that low amounts of sediment, like those used in standard methods of direct lysis and DNA extraction ( $< 1 \text{ g}$ ), can lead to an incomplete characterization of the protist community, although this might ultimately depend on the density of protist cells per volume unit. Furthermore, and in contrast with previous studies (Courtois et al., 2001; Robe et al., 2003), the DNA yield obtained for eluted samples was high, probably as an effect of the higher sample volume used. The lower yield obtained for directly lysed samples might reflect a relatively low protist density per sediment volume unit (less eukaryotic DNA amount) and/or a lower DNA purity (e.g., metal cations or organic acids inhibiting the Taq polymerase, depending on the used DNA purification method). This likely explains the failure to amplify 18S rRNA genes by direct PCR in some samples. In agreement with our results, Penton and colleagues (2016) demonstrated that the use of larger sample sizes (e.g., 10 g in front of 0.25/1/5 g) allowed the capture of irregularly distributed abundant and rare organisms (bacterial and fungal). Nascimento and colleagues (2018) also showed that sample volume affected all protist diversity metrics investigated, being higher when increasing volumes, suggesting that sample volumes  $> 10 \text{ g}$  are needed to achieve a representative assessment of alpha- and beta-diversity of microorganisms that are non-homogeneously distributed in sediments. Even if Lanzén and colleagues (2017) obtained a better representation of diversity values when increasing DNA extraction replicates than using higher amounts of sample volume, our results are congruent with those claiming that most-commonly used methods in HTS for the characterization of benthic protists lead to incomplete community determination due to the low amount of sample used and the heterogeneity of organism's distribution in the sediment. Likewise, Delmont and colleagues (2011a) concluded that the use of sample of  $\sim 100 \text{ g}$  was sufficient to capture the majority of bacterial diversity,

such that this could be used rather than increasing sampling effort, and that the major player in the estimation of community descriptors was the DNA extraction method (including direct and indirect ones). Nascimento and colleagues (2018) also demonstrated that the larger the sample volume, the more similar samples were among them. These results suggest that in protists, differences attributed to patchy distributions might often obscure non-representative pictures of community composition due to insufficient sample volumes.

#### Effect on the composition of protist community

Eluted samples were clearly dominated by Alveolata: ciliates and dinoflagellates (mean of 46.6% and 36.8% reads respectively). These two groups were also abundant in direct-lysis sediment samples (28.1% and 9.2% respectively). Other groups, such as fungi (22.1%) and ochrophytes (23.0%) were also relatively abundant in direct-lysis samples, but represented lower abundances in eluted samples (Fig. 3A, Table 1). Some eukaryotic groups were present at lower percentages in both treatments without remarkable differences (Fig. 3B). The groups presenting significant differences between treatments were Dinophyta, Fungi, Ochrophyta and Apusomonadidae, (Fig. 3), confirming the selection effect (positive for Dinophyta and negative for the others) of the elution method (Supporting Information Fig. S3). The comparison of the community composition inferred from both methodologies showed that 1,459 OTUs were shared among treatments, representing 43.4% of those present in direct-lysis sediment samples and 15% of eluted ones. However, those shared OTUs represented 84% and 75% of reads, respectively, showing that most dominant OTUs were obtained by both methods, and most non-shared OTUs comprised low number of reads. This was also observed for the different taxonomic groups: shared OTUs comprised a high percentage of reads in most groups, and usually represented a fraction > 80% in direct-lysis samples (Table 1). The richness inferred for each taxonomic group (number of OTUs, regardless of their abundance) was higher (more OTUs) in eluted samples, except for Ochrophytes, Radiolaria and Apusomonadidae, which yielded a higher number of OTUs in direct-lysis samples (Fig. 4). All taxonomic groups showed similar proportions of OTUs shared between both treatments, and those unique for the direct-lysis treatment. Likewise, the exceptions were Ochrophytes and Fungi, which showed a higher proportion of OTUs unique for the direct-lysis treatment, in agreement with the higher representation in this treatment. At any rate, this difference cannot be explained by the difference in reads obtained for the two treatments. Actually, Fungi showed a similar richness of OTUs in direct-lysis samples compared to eluted samples, but the proportion of reads was much higher in direct-lysis samples (22.1%) than in eluted ones (0.73%) (Table 1). This might be explained by the lack (fungi) or limited (ochrophytes) mobility of these groups or their larger size, which might hamper their elution from the sediment. Also, it might be that members of these groups are not (or less) active and correspond to resting stages more difficult to retrieve by the elution process. Conversely, Dinophyta and Ciliophora showed higher richness of OTUs in eluted samples, in agreement with the relative abundances obtained (Fig. 3). Finally, some eukaryotic groups showed unexpected richness despite they were represented by a low number of reads in both treatments. This is the case of Cercozoans or Katablepharidophyta, which represent the third and fourth most diverse group in eluted samples, comprising 1.7% and 2.5% of reads respectively (Table 1).

The melting seawater-ice (Ühlig) elution method is supposed to select organisms with active motility, even though many other groups of organisms can be partially recovered just by the water flow created in the sediment column. Additionally, the mesh pore used to separate cells from the sediment only allowed to recover organisms < 60 µm. The posterior filtration on 3.0 µm filters removed those below this size. Thus, it was expected to predominantly recover motile organisms with body sizes from 3 to 60 µm, and in fact, eluted samples were dominated by taxonomic groups agreeing with those characteristics, such as ciliates and dinoflagellates (> 80% of reads). By contrast, direct DNA extraction from sediments should affect less the original composition of organisms. In direct-lysis samples, ciliates and dinoflagellates represented 37% of reads, confirming their important contribution to the community composition.

However, other groups like Ochrophytes and Fungi were also highly represented in direct-lysis samples but not in eluted samples (Fig. 3), suggesting that part of this component was not recovered when using the seawater ice separation method. In any case, shared OTUs between both treatments comprised ~80% of reads in both data sets, confirming that, although differing in their relative abundance, most abundant OTUs were recovered using both methods, all lineages detected were present in both data sets and differences observed in terms of richness and taxonomic composition corresponded to low abundant OTUs.

#### Dinoflagellate and ciliate community composition

Given that ciliates and dinoflagellates dominated the eluted samples and were a significant component of 'direct-lysis' sediment samples, their diversity and relative abundance was specifically compared to test possible differences among treatments (Table 2). Of all OTUs (6,385) belonging to ciliates (Ciliophora) or dinoflagellates (Dinophyceae), 483 were shared among treatments, representing 8.1% of those from eluted samples and 53.9% of those from direct-lysis in sediments (Fig. 4). All major taxonomic groups were present in both data sets, while those represented at low relative abundances (< 0.02% and all belonging to Ciliophora) were only present in eluted samples (not shown), or showed low levels of shared OTUs (e.g., Prorocentrales and Ciliophora group 7). In most cases, those shared OTUs comprised more than 85% of all reads obtained in direct-lysis samples for that taxonomic group and the percentages were generally higher than those of eluted samples. Even though some groups like Peridiniales or 'Uncertain naked dinoflagellates' showed significant differences in their relative abundances among treatments, all taxonomic subgroups were present at abundances within the same range, confirming their dominance or rareness in the community. Thus, both treatments appear to yield a reliable characterization of dinoflagellate and ciliate communities. But, as observed for the entire community, the elution method allowed capturing higher richness of ciliates and dinoflagellates, confirming that the seawater ice 'Uhlrig' treatment should be chosen when the objective of the study focuses on characterizing the community of dinoflagellates or ciliates.

**Concluding remarks** We have carried out a study of the microbial eukaryotic diversity inferred by 18S rRNA gene metabarcoding in sediments after elution of protist cells by melting water in comparison with results from direct-lysis and DNA purification from sediments. We have shown that (i) alphadiversity obtained for the elution method is much higher than the obtained for direct-lysis, likely as a result of the larger sediment volume used to obtain DNA samples. Additionally, (ii) eluted samples showed a higher similarity among them and, accordingly, reduced variability owing to stochastic subsampling effects, or patchiness of benthic communities, implying that standard methods used for metabarcoding based on small sample volumes (especially in cases of low protist density) can lead to an inadequate characterization of sample richness. We also show that although the seawater-ice elution method enriches some motile groups, it allows to recover most abundant OTUs of all taxonomic groups, although relative abundances are biased towards some of them in eluted samples. Anyway, most abundant OTUs were present in both data sets. Consequently, (iii) the seawater-ice elution seems a time and cost-efficient method that provides a more complete determination of total protist richness, especially for dinoflagellates and ciliates.

#### Acknowledgements

We thank R. Gallisai (ICM-CSIC) and T. Slámová (Univ. Prague) for assistance during samplings and P. Bertolino (CNRS – Univ. Paris-Sud) for technical assistance during molecular analyses. G.R., D.M. and P. L.-G. were funded by European Research Council Grant ProtistWorld (no. 322669) under the European Union's Seventh Framework Program. A.R. was granted by MINECO Grant COPAS 'Understanding top-



down control in coastal bloom-forming protists' (CTM2017-86121-R), and a MECD grant 'Estancia de Movilidad en el extranjero José Castillejo' (CAS17/00237).

#### Conflict of interest

The authors have no conflict of interest to declare.

#### References

- Azovsky, A., Saburova, M., Tikhonenkov, D.V., Khazanova, K., Esaulov, A., and Mazei, Y. (2013) Composition, diversity and distribution of microbenthos across the intertidal zones of Ryazhkov Island (the White Sea). *Eur J Protistol* 49: 500-515.
- Bak, R.P.M., and Nieuwland, G. (1989) Seasonal fluctuations in benthic protozoan populations at different depths in marine sediments. *Netherlands Journal of Sea Research* 24: 37-44.
- Bik, H.M., Sung, W., De Ley, P., Baldwin, J.G., Sharma, J., Rocha-Olivares, A., and Thomas, W.K. (2012) Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. *Mol Ecol* 21: 1048-1059.
- Chariton, A., Court, L., Hartley, D., Collof, M., and Hardy, C. (2010) Ecological assessment of estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. *Frontiers in Ecology and the Environment* 8: 233-238.
- Courtois, S., Frostegard, A., Göransson, P., Depret, G., Jeannin, P., and Simonet, P. (2001) Quantification of bacterial subgroups in soil: comparison of DNA extracted directly from soil or from cells previously released by density gradient centrifugation. *Environmental Microbiology* 3: 431-439.
- de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R. et al. (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science* 348: 1-11.
- Delmont, T.O., Robe, P., Clark, I.M., Simonet, P., and Vogel, T.M. (2011a) Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J Microbiol Methods* 86: 397-400.
- Delmont, T.O., Robe, P., Cecillon, S., Clark, I.M., Constancias, F., Simonet, P. et al. (2011b) Accessing the soil metagenome for studies of microbial diversity. *Appl Environ Microbiol* 77: 1315-1324.
- Díez, B., Pedrós-Alió, C., and Massana, R. (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Appl Environ Microbiol* 67: 2932-2941.
- Dragesco, J. (1965) Étude cytologique de quelques flagellés mésopsammiques. *Cah Biol Mar* 6: 83-115.
- Fenchel, T. (1969) The ecology of marine microbenthos. IV. Structure and function of the benthic ecosystem, its chemical and physical factors and the microfauna communities with special reference to the ciliated protozoa. *Ophelia* 6: 1-182.
- Forster, D., Dunthorn, M., Mahé, F., Dolan, J.R., Audic, S., Bass, D. et al. (2016) Benthic protists: the under-charted majority. *FEMS Microbiol Ecol* 92: 1-11.
- Gong, J., Shi, F., Ma, B., Dong, J., Pachiadaki, M., Zhang, X., and Edgcomb, V.P. (2015) Depth shapes  $\alpha$ - and  $\beta$ -diversities of microbial eukaryotes in surficial sediments of coastal ecosystems. *Environmental Microbiology* 17: 3722-3737.
- Hoppenrath, M., Murray, S., Chomérat, N., and Horiguchi, T. (2014) Marine benthic dinoflagellates - unveiling their worldwide biodiversity: Kleine Senckenberg-Reihe 54. Senckenberg Gesellschaft für Naturforschung, Senckenberganlage 25 Frankfurt am Main, Germany.
- Lanzén, A., Lekang, K., Jonassen, I., Thompson, E.M., and Troedsson, C. (2017) DNA extraction replicates improve diversity and compositional dissimilarity in metabarcoding of eukaryotes in marine sediments. *PLoS ONE* 12: e0179443.

- Logares, R., Haverkamp, T.H.A., Kumar, S., Lanzén, A., Nederbragt, A.J., Quince, C., and Kausrud, H. (2012) Environmental microbiology through the lens of high-throughput DNA sequencing: Synopsis of current platforms and bioinformatics approaches. *J Microbiol Methods* 91: 106-113.
- López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C., and Moreira, D. (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409: 603-607.
- Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Poulain, J. et al. (2016) Insights into global diatom distribution and diversity in the world's ocean. *Proceedings of the National Academy of Sciences*: E1516-E1525.
- Mare, M.F. (1942) A study of a marine benthic community with special reference to the micro-organisms. *J Mar Biol Ass UK* 25: 51-554.
- Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C. et al. (2015) Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environmental Microbiology* 17: 4035-4049.
- Moon-van der Staay, S.Y., De Watcher, R., and Vaultot, D. (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* 409: 607-610.
- Nascimento, S.M., Lallias, D., Bik, H.M., and Creer, S. (2018) Sample size effects on the assessment of eukaryotic diversity and community structure in aquatic sediments using high-throughput sequencing. *Scientific Reports* 8: 11737.
- Pan, Y., Yang, J., McManus, G.B., Lin, S., and Zhang, W. (2019) Insights into protist diversity and biogeography in intertidal sediments sampled across a range of spatial scales. *Limnol Oceanogr* doi: 10.1002/lno.11375.
- Parent, B., Barras, C., and Jorissen, F. (2018) An optimised method to concentrate living (Rose Bengal-stained) benthic foraminifera from sandy sediments by high density liquids. *Mar Micropaleontol* 144: 1-3.
- Penton, C.R., Gupta, V.V.S.R., Yu, J., and Tiedje, J.M. (2016) Size matters: Assessing optimum soil sample size for fungal and bacterial community structure analyses using high throughput sequencing of rRNA gene amplicons. *Frontiers in Microbiology* 7: 824.
- Piredda, R., Tomasino, M.P., D'Erchia, A.M., Manzari, C., Pesole, G., Montresor, M. et al. (2016) Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. *FEMS Microbiol Ecol* 93: fiw200.
- Quaiser, A., Zivanovic, Y., Moreira, D., and López-García, P. (2011) Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. *The Isme Journal* 5: 285-304.
- Robe, P., Nalin, R., Capellano, C., Vogel, T.M., and Simonet, P. (2003) Extraction of DNA from soil. *Eur J Soil Biol* 39: 183-190.
- Saburova, M., Polikarpov, I.G., and Burkovsky, I.V. (1995) Spatial structure of an intertidal sandflat microphytobenthic community as related to different spatial scales. *Mar Ecol Prog Ser* 129: 229-239.
- Salonen, I.S., Chronopoulou, P.M., Leskinen, E., and Koho, K.A. (2019) Metabarcoding successfully tracks temporal changes in eukaryotic communities in coastal sediments. *FEMS Microbiol Ecol* 95: fiy226.
- Smith, D.P., and Peay, K.G. (2014) Sequence Depth, Not PCR Replication, Improves Ecological Inference from Next Generation DNA Sequencing. *PLoS ONE* 9: e90234.
- Starink, M., Bär-Gilissen, M.-J., Bak, R.P.M., and Cappenberg, T. (1994) Quantitative centrifugation to extract benthic protozoa from freshwater sediments. *Appl Environ Microbiol* 60: 167-173.
- Steffan, R.J., Goksoyr, J., Bej, A.K., and Atlas, R.M. (1988) Recovery of DNA from soils and sediments. *Appl Environ Microbiol* 54: 2908-2915.
- Uhlig, G. (1964) Eine einfache methode zur extraktion der vagilen, mesopsammalen microfauna. *Helgol Wiss Meeresunters* 11: 178-185.
- Webb, M.G. (1956) An ecological study of brackish water ciliates. *J Anim Ecol* 25: 148-175.

**Table 1: Relative abundance of reads (%) of major taxonomic groups in samples pooled by sample treatment (eluted and direct-lysis).** The ‘Total’ columns show the percentage of reads of each taxonomic group in the whole community. The ‘Shared’ columns display the percentage of reads belonging to OTUs shared between both treatments. Only taxonomic groups >0.1% in both treatments are shown.

Superphylum	Phylum	Total		Shared	
		Eluted	Direct lysis	Eluted	Direct lysis
Alveolata	Ciliophora	44.05	28.08	76.54	85.78
	Dinophyta	37.55	9.21	77.96	87.00
	Perkinsea	0.17	0.43	30.93	62.81
Archaeplastida	Chlorophyta	2.54	1.33	87.21	96.19
	Streptophyta	0.29	1.80	4.60	85.60
Hacrobia	Katablepharidophyta	2.59	0.55	45.33	96.26
	Cryptophyta	0.88	0.45	75.42	95.06
	Telonemia	0.30	0.48	85.97	85.31
Opisthokonta	Fungi	0.92	22.07	37.87	78.86
	Mesomycetozoa	0.23	0.40	91.50	99.83
	Choanoflagellida	0.12	0.18	4.55	19.58
Rhizaria	Cercozoa	1.99	3.79	38.88	64.81
	Radiolaria	0.16	0.91	36.41	37.25
Stramenopiles	Ochrophyta	5.08	23.01	86.15	86.61
	MAST	1.13	1.27	68.55	81.16
	Pirsonia clade	0.79	1.13	49.38	77.29

**Table 2: Relative abundance of reads (%) of different ciliate and dinoflagellate taxa in samples pooled by sample treatment (eluted and direct-lysis).** The ‘Total’ columns display the percentage of reads of each taxonomic group in the whole community. The ‘Shared’ columns represent the percentage of reads belonging to OTUs shared between both treatments in relation to their totality. Taxonomic groups <0.02% in both treatments are omitted. Shaded area: no shared OTUs.

Taxonomic group		Total		Shared	
		Eluted	Direct lysis	Eluted	Direct lysis
Ciliophora	Spirotrichea	28.91	20.19	75.3	88.2
	Oligohymenophorea	7.34	1.26	67.3	75.9
	Ciliophora group 5	4.45	3.22	99.9	100
	Prostomatea	2.79	2.61	81.5	61.0
	Phyllopharyngea	0.18	0.41	70.6	69.3
	Colpodea	0.17	0.1	43.9	100
	Litostomatea	0.09	0.14	74.8	42.0
	Heterotrichea	0.05	0.05	8.7	98.9
	Karyorelictea	0.03	0.02	2.5	100
	Ciliophora group 7	0.02	0.08		
Dinophyceae	Peridinales	15.38	2.34	88.0	96.5
	Uncertain Naked	12.84	0.37	76.9	99.1
	Uncertain	3.64	1.54	75.0	88.2
	Gymnodinales	3.33	1.45	50.9	83.5
	Gonyaulacales	1.23	1.14	92.2	94.2
	Dinophysiales	0.34	0.16	45.2	94.0
	Suessiales	0.09	0.12	50.7	87.8
	Prorocentrales	0.08	0.02	2.2	7.6
	Uncertain Thecate	0.06	0.08	85.3	95.2

## Figure legends

Figure 1. Comparison of OTU richness in 'eluted' and 'direct-lysis' sediment samples. A) Distribution of the read abundances per OTU. B) Richness observed for eluted and direct-lysis samples. The horizontal lines in the density distribution area represent the median, 75th and 25th percentile. C) Rarefaction curves of both groups of samples, relating the increase in the number of reads with the number of OTUs for the complete dataset. Direct-lysis samples are represented in orange and eluted ones in blue.

Figure 2. Beta-diversity analyses. Heatmap showing the dissimilarity level among all samples organized by hierarchical clustering.

Figure 3. Comparison of community structure for different eukaryotic phyla between eluted and sediment samples. A) Relative abundance of dominant phyla, and B) phyla showing lower relative abundances (only those >2% are shown). Note the logarithmic scale. Each boxplot presents the median and interquartile range of the distribution of data points shown in grey. Whiskers represent 1.5 times the interquartile range. The colored area represents the density distribution of data points.

Figure 4. Distribution of OTUs between treatments and taxonomic groups. Observed OTUs by Phylum only present in either the elution (blue), the direct-lysis treatment (orange) or in both treatments (green).

**FIG 1**

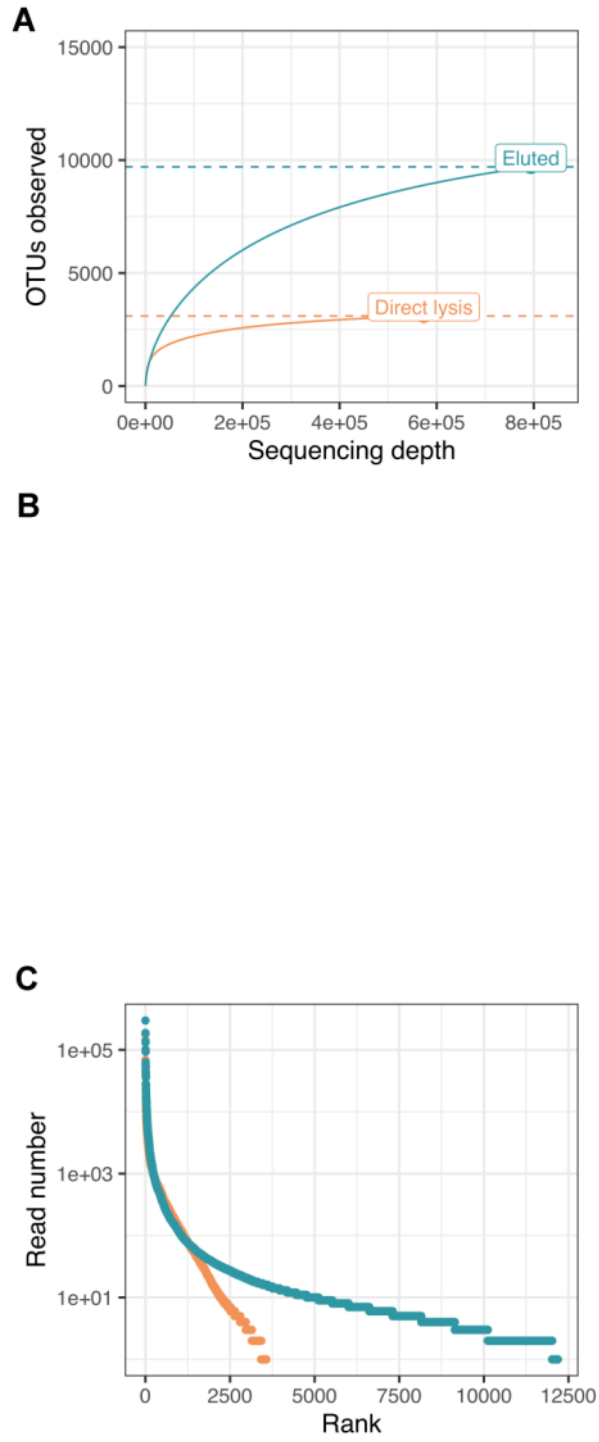


FIG 2

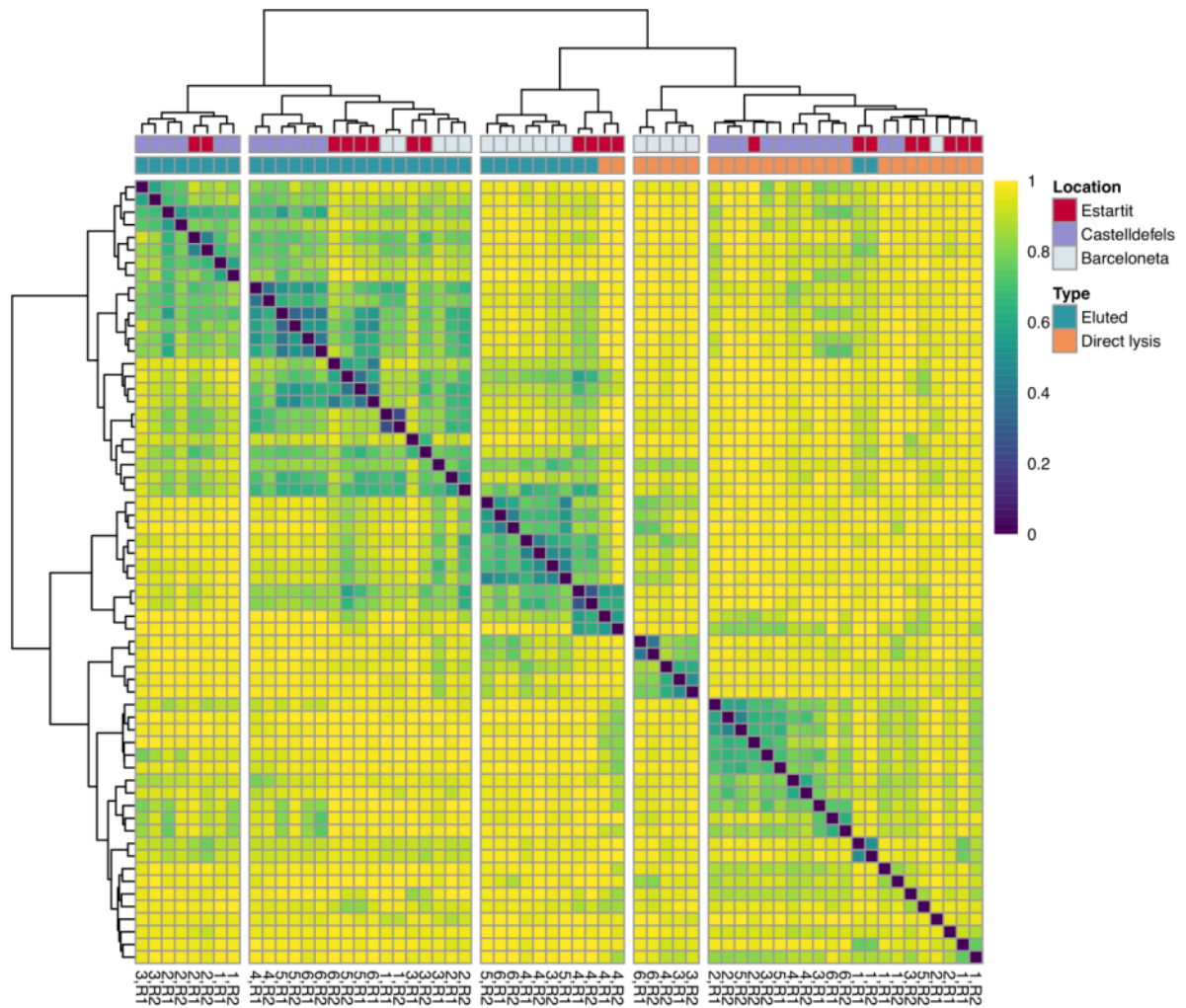


FIG 3

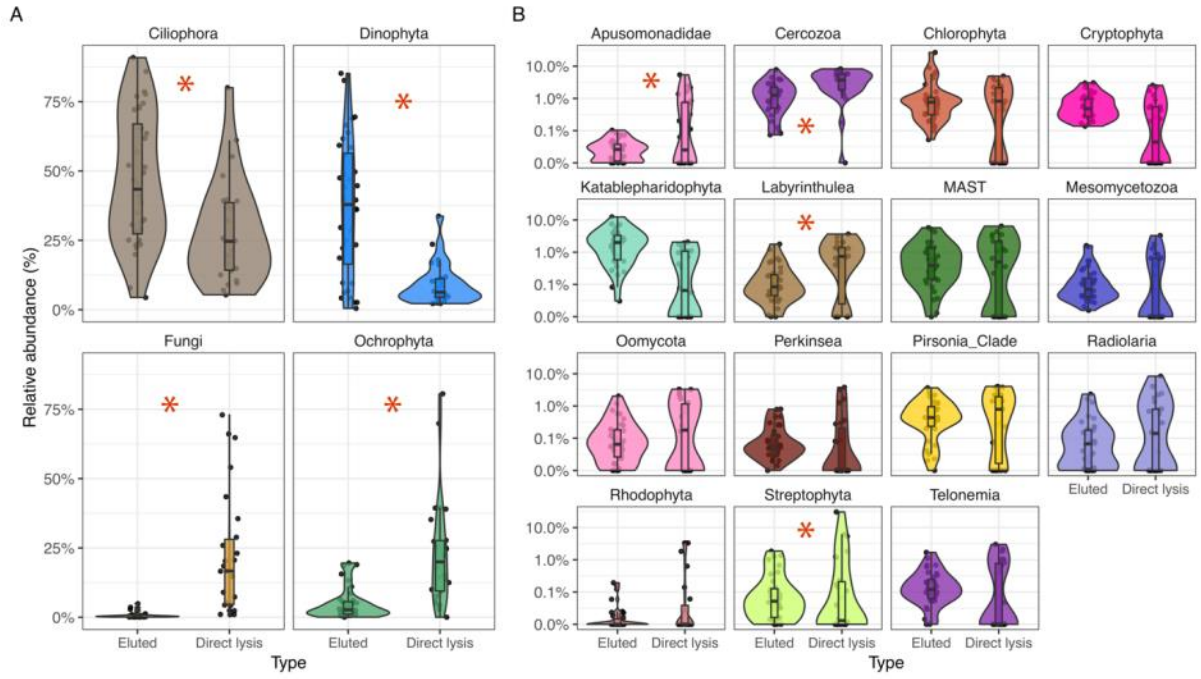
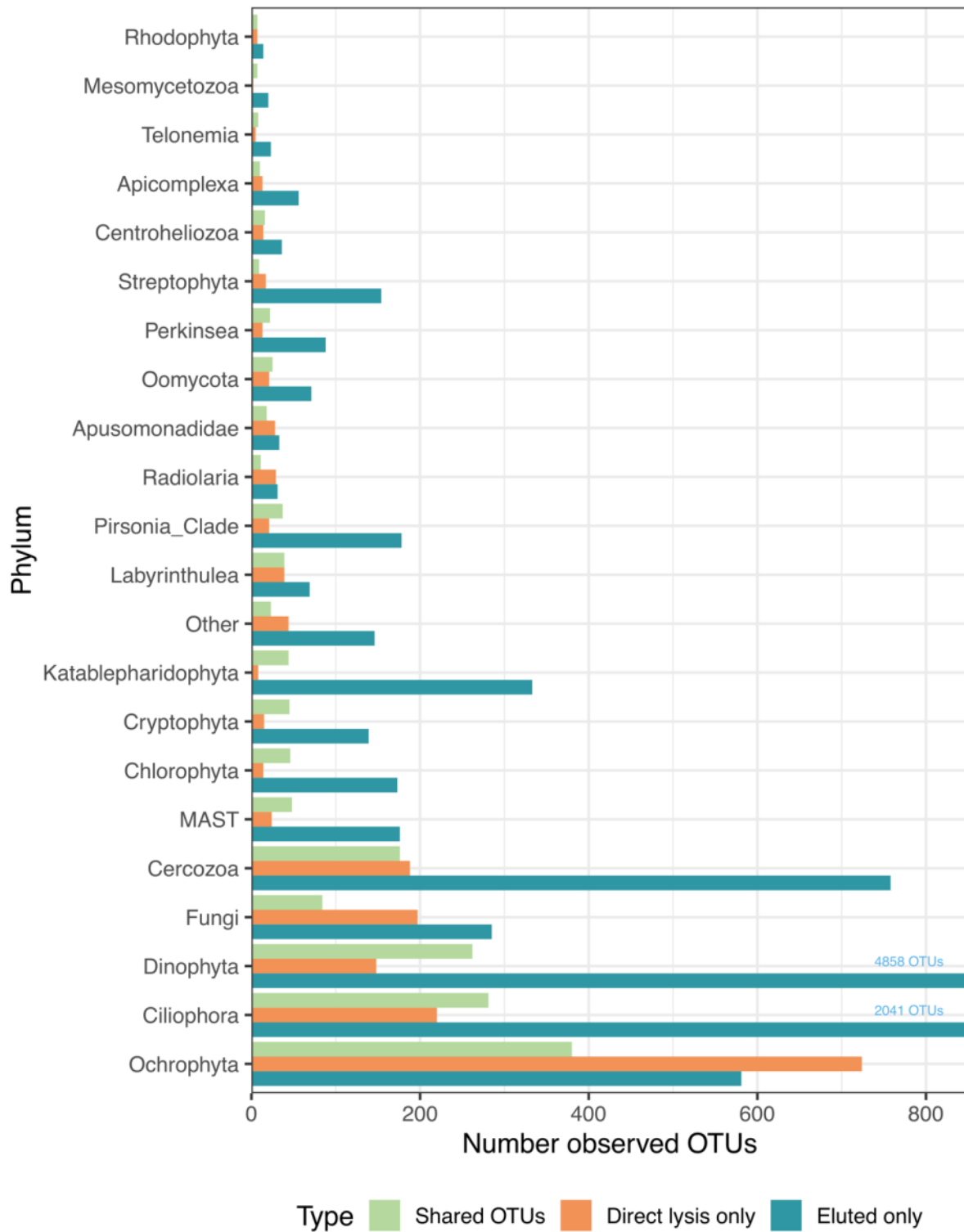




FIG 4



Supporting Information for Environmental Microbiology

**Performance of the melting seawater-ice elution method on the characterization of benthic protists community using metabarcoding analysis**

Albert Reñé <sup>1,2\*</sup>, Adrià Auladell <sup>2</sup>, Guillaume Reboul <sup>1</sup>, David Moreira <sup>1</sup>, Purificación López-García <sup>1</sup>

<sup>1</sup>Unité d'Ecologie, Systématique et Evolution, CNRS, Université Paris-Sud, AgroParisTech, Université Paris-Saclay, 91400 Orsay, France

<sup>2</sup>Departament de Biologia Marina i Oceanografia, Institut de Ciències del Mar (CSIC), Barcelona, Catalonia, Spain

## **EXPERIMENTAL PROCEDURES**

### **Sampling**

Sediment samples were collected monthly from April to September 2017 at three beaches from the NW Mediterranean Sea (Catalan coast): L'Estartit (42° 3' 5"N; 3° 12' 9"E), Barcelona (41°23'5" N; 2°11'55"E) and Castelldefels (41°15'35"N; 1°55'51"E). Sediment cores were individually obtained by snorkelling, close to the shoreline at approximate 2 meters depth. At each sampling visit, the microorganisms from the first 5 cm (80 cm<sup>3</sup> and approximately 100 gr of wet weight) of two sediment cores, considered as replicates, were eluted using the Uhlig method. Briefly, the sediment was placed in a PVC tube of 4.5 cm diameter, with a mesh of 60 µm attached to the bottom. The mesh was in contact with a Petri dish filled with previously 0.2 µm-filtered and autoclaved seawater. Two seawater-ice cubes, previously prepared with the same 0.2 µm-filtered and autoclaved seawater, were placed on top of the sediment for melting. The ice cubes melt in around 1 hour. Afterwards, the volume of seawater containing eluted cells recovered from the Petri dish was filtered by gravity through a 3.0 µm polycarbonate filter. The filters were stored into cryovials ('eluted' treatment). The sediment from a replicate core was gently mixed and two subsamples of approximately 1 cm<sup>3</sup> were placed in centrifuge tubes ('direct-lysis' treatment). Filters and sediment were fixed with absolute ethanol and stored at -80°C until analysed. A total of 72 samples were analysed in this study (Supporting Information Table S1).

### **Metabarcoding**

DNA was extracted using the PowerSoil DNA isolation kit (MoBio Laboratories, Carlsbad, CA, USA) following the manufacturer's instructions. However, the two types of samples required different pre-treatments. For direct lysis samples, sediment aliquots were gently mixed and approximately 1 g of sediment was transferred to a 2-mL Eppendorf tube.

Additionally, 900  $\mu$ L of a whitish precipitate present on top of the fixed sediments was also transferred to the tube. This material was centrifuged at 12,500 rpm for 3 min and ethanol removed by pipetting. Subsequently, 200  $\mu$ L of the kit resuspension buffer was added to the pellet and stored at 4°C for at least 2 hours to allow full rehydration. For eluted samples, filters retaining the biomass of microbes separated from sediment were cut into small pieces using a sterile scalpel and transferred to the first tube of the DNA purification kit containing microbeads and the first resuspension buffer. The remaining ethanol was transferred to 2-mL Eppendorf tubes, centrifuged at 12,500 rpm for 5 min and, after ethanol removal, the pellet was resuspended in the same buffer and added to the microbead-containing tubes. These were also stored at 4°C for at least 2 h for rehydration and subsequent DNA purification. PCR amplification of a 18S rRNA gene fragment containing the hypervariable V4 region was performed using MID (Molecular Identifiers) tagged primers EK-565F (Simon et al., 2015) and 18S-EUK-1134R-UnonMet biased against metazoans (Bower et al., 2004), yielding fragments of ca. 550 bp. Four independent PCRs were conducted for each sample following the protocol from Simon et al. (2015) and then pooled together. Thirteen of the sediment samples did not yield 18S rRNA gene amplicons. In these cases, we carried out a first PCR reaction using primers EK-82F and EUK-1134R-UnonMet followed by a (semi)nested-PCR with MID tagged primers as previously described, resulting in successful amplification. Amplicons were purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany), quantified, and equivalent amounts from the 72 samples were pooled together for paired-end (2 x 300 bp) MiSeq Illumina sequencing (Eurofins Genomics, Germany).

### **Metabarcoding pipeline to define OTUs from Illumina Mi-Seq Paired-end reads**

The paired-end reads obtained after sequencing were treated using an in-house pipeline following a standard protocol. First, the paired-end reads were merged using FLASH v1.2.11

(Magoč and Salzberg, 2011) with `--min-overlap` and `--max-overlap` parameters set to 5 and 50 respectively. Then, the full amplicons were checked for quality trimming in three steps: i) only merged amplicons were kept if their complete MID sequence of 10 nucleotides was found in 5' and 3', then ii) searching and removing MID sequences as well as PCR primer sequences using `cutadapt v1.14` (Martin, 2011) with an error rate of 10%. Finally, iii) the full amplicons for which reverse PCR primer was found in 5' and forward primer in 3' were reversed and complemented. In addition, full amplicons in which their corresponding MID was found in the DNA sequence after this quality control treatment were discarded to avoid possible chimeras. Only CMRs (clean and merged reads) above 450 nucleotides of length were selected for subsequent dereplication into clusters of CMRs sharing exact same sequence and exact same length using `vsearch v2.3.4` (Rognes et al., 2016), option `--derep_fulllength`, and option `--sizeout`, which allowed us to retrieve the information of cluster abundance. A further potential chimera detection was computed on those clusters with `vsearch v2.3.4` (Rognes et al., 2016) (options `--uchime_denovo` and the default parameters of `--minh` and `--xn`). CMRs not tagged as chimeric were kept and clustered using `swarm` (Mahé et al., 2015) in fastidious mode (`-d 1 -f`). All singleton (made of only one CMR) swarms (operational taxonomic units, OTUs) were discarded from the downstream analyses. All 18,470 non-singleton OTUs were blasted against home-made cultivated and environmental databases based on `SILVAv128` (Quast et al., 2013) and `PR2v4.5` (Guillou et al., 2013) using `vsearch v2.3.4` (Rognes et al., 2016) pairwise alignment; only the best hit was retrieved and used as a taxonomic proxy. OTUs corresponding to metazoans, nucleomorphs or bacteria were removed and those showing a similarity <75% to the closest match were manually reclassified as 'Unknown'. Files with raw sequences were deposited to NCBI's Short Read Archive under the study accession number `PRJNA531674`. Samples metadata is provided in Supporting Information Table S1.

## **Data analyses**

10 of the 72 samples issued from the direct-lysis method did not yield amplicons and were subjected to reamplification by semi-nested PCR (Supporting Information Table S1). After quality trimming, clustering of sequences in ‘swarm’ operational taxonomic units (hereafter simply referred as OTU), removal of singletons and exclusion of amplicons not corresponding to protists, the first inspection of the read abundances and OTU composition showed that those 10 samples were composed of very few OTUs, mostly belonging to Fungi. We interpret this as the result of very little protist biomass per sediment volume in these samples and subsequent nested-PCR-associated biases. Consequently, those samples were removed from the dataset for further comparative analyses, and their counterpart replicates treated under elution method were also removed in order to have the same number of samples for each treatment. We thus retained 26 samples that corresponded to DNA purified from small sediment volumes by direct lysis, and 26 samples that corresponded to DNA obtained from protists eluted from larger sediment volumes using the melting seawater ice method, i.e. a total of 52 samples. The dataset referred to as ‘reduced’ was generated by randomly reducing the number of reads of each eluted sample to 1:50, using the `rarefy_even_depth` function from `vegan` package (R software), in order to artificially simulate the same sample volume used for sediment samples analysed by the direct-lysis approach.

All statistical analyses were run under R software and data transformations using the package `phyloseq` [44] and `tidyverse` [45]. Data visualization was performed with the `ggplot2` package (from `tidyverse`). For beta diversity analyses, the rarefied data matrix (i.e. 22,056 reads per sample) was square-root transformed and Bray-Curtis dissimilarity calculated. Rarecurves, alpha diversity and beta diversity analysis, non-parametric statistical test of analysis of similarities (ANOSIM), multivariate homogeneity of group dispersions (BETADISPER) and

permutational multivariate analysis of variance (ADONIS) were performed through vegan package [46]. Multivariate dispersion test (MVDISP) and Index of Multivariate Dispersion (IMD) was run under Primer 6.1 software [47]. To test for differential abundance at the subphylum level, the corncob package was used [48]. The algorithm estimates the relative abundance of taxa through modelling the distribution with a beta binomial model, adequate for the compositional data from microbiome datasets.

## **RESULTS**

We compared the diversity values obtained by both methodologies successively reducing the number of reads of eluted samples by multiples of ten, from 1:10 to 1:100 in order to simulate the use of smaller sample sizes for eluted samples. Eluted and direct-lysis sediment samples showed similar richness values when reducing the number of reads of the first samples (Fig. S4A), with no significant differences between both treatments from 1:50 reads reduction in eluted samples and onwards (ANOVA Tukey HSD  $p > 0.1$ ). That effect was observed in the rarefaction curves, such that eluted samples showed a similar number of OTUs observed to the direct-lysis ones, and thus being far from saturation (Fig. S4B). The same trend could also be observed in terms of distribution of OTU abundances. Both datasets (the original one for directly-lysed samples, and the reduced to 1:50 reads for eluted samples) showed a similar pattern, with a strong reduction in the number of OTUs due to the loss of less-abundant ones in eluted samples (2,657 in eluted samples and 3,598 in direct-lysis sediment ones) (Fig. S4C).

After the reduction of reads to 1:50 in eluted samples, beta-diversity analyses (Fig. S5) still showed significant differences in sample dispersion between treatments (ADONIS  $R=0.106$ ,  $p < 0.001$ , BETADISPER  $F_{1,50} = 8.27$ ,  $p < 0.01$ ), and eluted samples having lower dispersion

between them than direct-lysis sediment samples (MVDisp, IMD = 0.493, Dispersion Eluted = 0.75; Direct lysis = 1.27).

When reducing the number of reads to artificially compensate the difference of sample volume between treatments, richness was not significantly different to direct-lysis sediment samples.

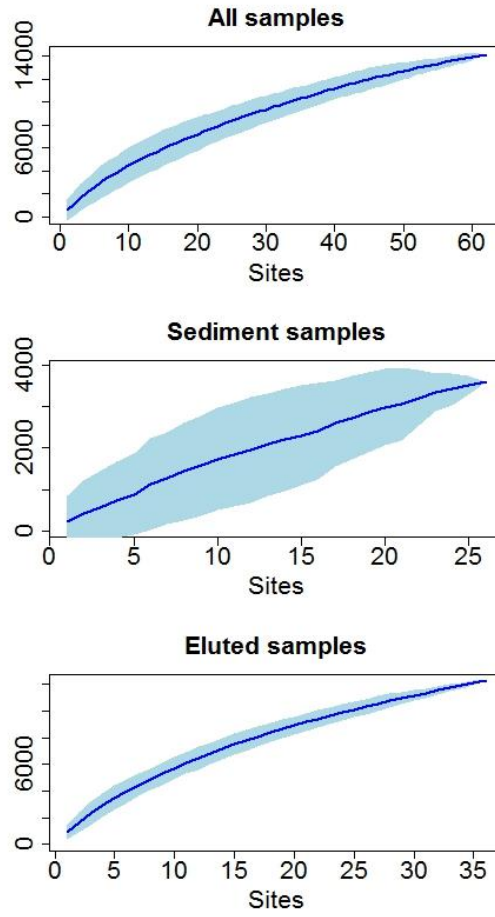
In any case, beta-diversity analyses performed after reducing the number of reads (1:50) to simulate the low amount of sediment used for direct lysis still showed a clear clustering by sample treatment, with stronger effect than other factors like locality or seasonality, and differences in dispersion between groups. Therefore, in addition to sample volume effects, the influence of the sample treatment in describing the community composition seems fundamental.



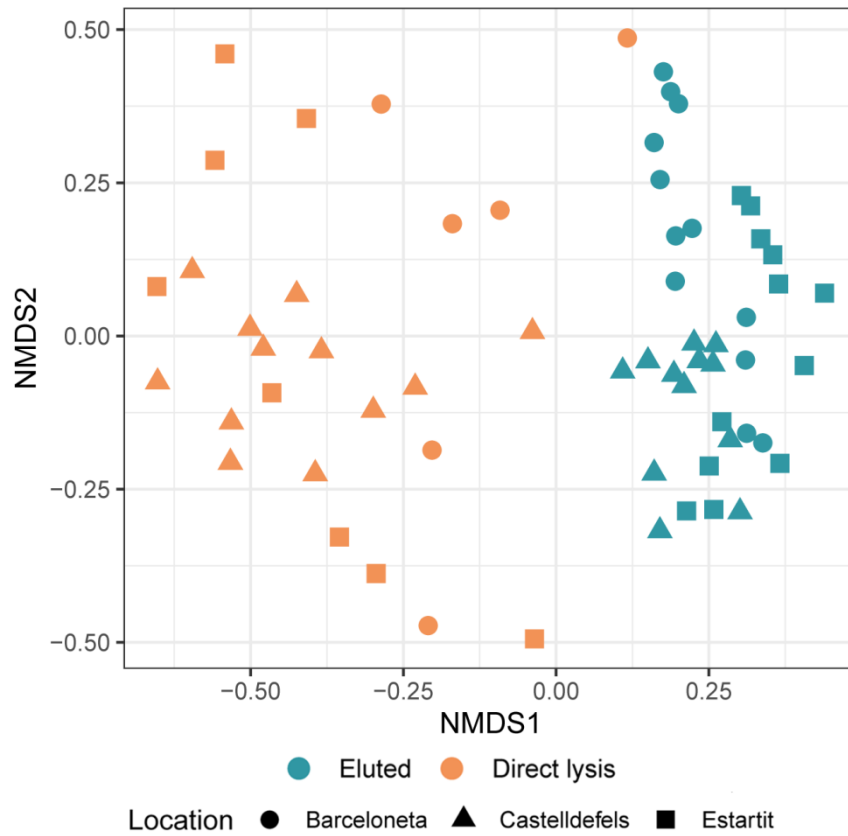
**Supplementary Table 1: List of samples analyzed.** The asterisks indicate the samples where a nested PCR was needed to amplify sufficient product for subsequent steps. Shaded rows indicate the samples that were removed for subsequent analyses.

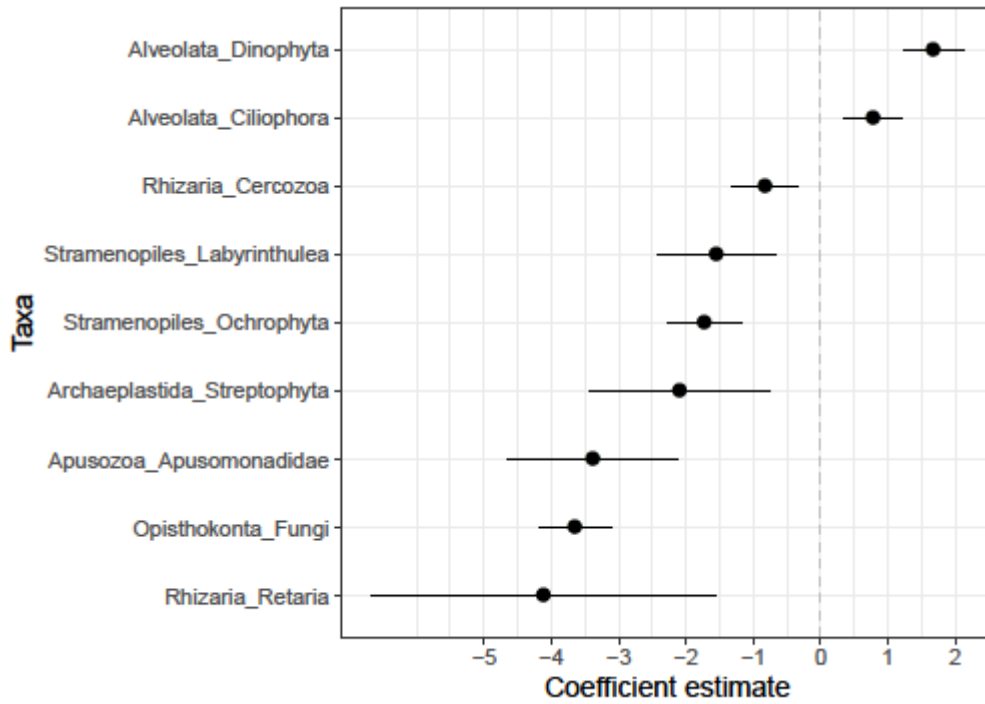
Sample	Location	Sampling Date	Treatment	Replicate	Cleaned Reads	Swarms	Nested PCR
EST1S1	Estartit	2017-04-20	Direct lysis	1	53266	324	
EST1S2	Estartit	2017-04-20	Direct lysis	2	82153	207	
EST2S1	Estartit	2017-05-10	Direct lysis	1	47164	128	
EST2S2	Estartit	2017-05-10	Direct lysis	2	32056	37	
EST3S1	Estartit	2017-06-01	Direct lysis	1	89726	14	*
EST3S2	Estartit	2017-06-01	Direct lysis	2	34879	86	
EST4S1	Estartit	2017-07-19	Direct lysis	1	56008	75	*
EST4S2	Estartit	2017-07-19	Direct lysis	2	36210	32	
EST5S1	Estartit	2017-08-21	Direct lysis	1	92525	8	*
EST5S2	Estartit	2017-08-21	Direct lysis	2	54573	73	
EST6S1	Estartit	2017-09-26	Direct lysis	1	81636	23	*
EST6S2	Estartit	2017-09-26	Direct lysis	2	103696	42	*
BCN1S1	Barcelona	2017-05-02	Direct lysis	1	74654	5	*
BCN1S2	Barcelona	2017-05-02	Direct lysis	2	79163	33	*
BCN2S1	Barcelona	2017-05-16	Direct lysis	1	42351	144	*
BCN2S2	Barcelona	2017-05-16	Direct lysis	2	17523	80	*
BCN3S1	Barcelona	2017-06-13	Direct lysis	1	81848	201	
BCN3S2	Barcelona	2017-06-13	Direct lysis	2	66521	222	
BCN4S1	Barcelona	2017-07-05	Direct lysis	1	93829	26	*
BCN4S2	Barcelona	2017-07-05	Direct lysis	2	54529	255	
BCN5S1	Barcelona	2017-08-07	Direct lysis	1	115753	5	*
BCN5S2	Barcelona	2017-08-07	Direct lysis	2	85244	79	*
BCN6S1	Barcelona	2017-09-13	Direct lysis	1	70792	1363	
BCN6S2	Barcelona	2017-09-13	Direct lysis	2	108719	112	
CSF1S1	Castelldefels	2017-05-02	Direct lysis	1	27352	94	
CSF1S2	Castelldefels	2017-05-02	Direct lysis	2	45311	51	
CSF2S1	Castelldefels	2017-05-25	Direct lysis	1	58279	62	
CSF2S2	Castelldefels	2017-05-25	Direct lysis	2	64350	96	*
CSF3S1	Castelldefels	2017-06-20	Direct lysis	1	22056	129	
CSF3S2	Castelldefels	2017-06-20	Direct lysis	2	71596	97	
CSF4S1	Castelldefels	2017-07-24	Direct lysis	1	97466	236	
CSF4S2	Castelldefels	2017-07-24	Direct lysis	2	49564	142	
CSF5S1	Castelldefels	2017-08-14	Direct lysis	1	53985	72	
CSF5S2	Castelldefels	2017-08-14	Direct lysis	2	107349	58	
CSF6S1	Castelldefels	2017-09-20	Direct lysis	1	143880	712	
CSF6S2	Castelldefels	2017-09-20	Direct lysis	2	60395	66	

EST1F1	Estartit	2017-04-20	Eluted	1	71192	994	
EST1F2	Estartit	2017-04-20	Eluted	2	33632	737	
EST2F1	Estartit	2017-05-10	Eluted	1	86463	1310	
EST2F2	Estartit	2017-05-10	Eluted	2	97150	1241	
EST3F1	Estartit	2017-06-01	Eluted	1	62285	805	
EST3F2	Estartit	2017-06-01	Eluted	2	69452	1380	
EST4F1	Estartit	2017-07-19	Eluted	1	67362	725	
EST4F2	Estartit	2017-07-19	Eluted	2	66298	633	
EST5F1	Estartit	2017-08-21	Eluted	1	63667	963	
EST5F2	Estartit	2017-08-21	Eluted	2	74847	901	
EST6F1	Estartit	2017-09-26	Eluted	1	118443	1343	
EST6F2	Estartit	2017-09-26	Eluted	2	64715	766	
BCN1F1	Barcelona	2017-05-02	Eluted	1	79115	851	
BCN1F2	Barcelona	2017-05-02	Eluted	2	67100	956	
BCN2F1	Barcelona	2017-05-16	Eluted	1	83341	1229	
BCN2F2	Barcelona	2017-05-16	Eluted	2	91565	1120	
BCN3F1	Barcelona	2017-06-13	Eluted	1	73673	1167	
BCN3F2	Barcelona	2017-06-13	Eluted	2	94889	1025	
BCN4F1	Barcelona	2017-07-05	Eluted	1	56544	1265	
BCN4F2	Barcelona	2017-07-05	Eluted	2	96224	515	
BCN5F1	Barcelona	2017-08-07	Eluted	1	144597	775	
BCN5F2	Barcelona	2017-08-07	Eluted	2	79284	906	
BCN6F1	Barcelona	2017-09-13	Eluted	1	48709	820	
BCN6F2	Barcelona	2017-09-13	Eluted	2	23159	619	
CSF1F1	Castelldefels	2017-05-02	Eluted	1	91269	992	
CSF1F2	Castelldefels	2017-05-02	Eluted	2	103066	583	
CSF2F1	Castelldefels	2017-05-25	Eluted	1	91531	1107	
CSF2F2	Castelldefels	2017-05-25	Eluted	2	111632	604	
CSF3F1	Castelldefels	2017-06-20	Eluted	1	70753	721	
CSF3F2	Castelldefels	2017-06-20	Eluted	2	60886	771	
CSF4F1	Castelldefels	2017-07-24	Eluted	1	86912	1361	
CSF4F2	Castelldefels	2017-07-24	Eluted	2	80499	1354	
CSF5F1	Castelldefels	2017-08-14	Eluted	1	60404	666	
CSF5F2	Castelldefels	2017-08-14	Eluted	2	81911	728	
CSF6F1	Castelldefels	2017-09-20	Eluted	1	96811	628	
CSF6F2	Castelldefels	2017-09-20	Eluted	2	96142	522	

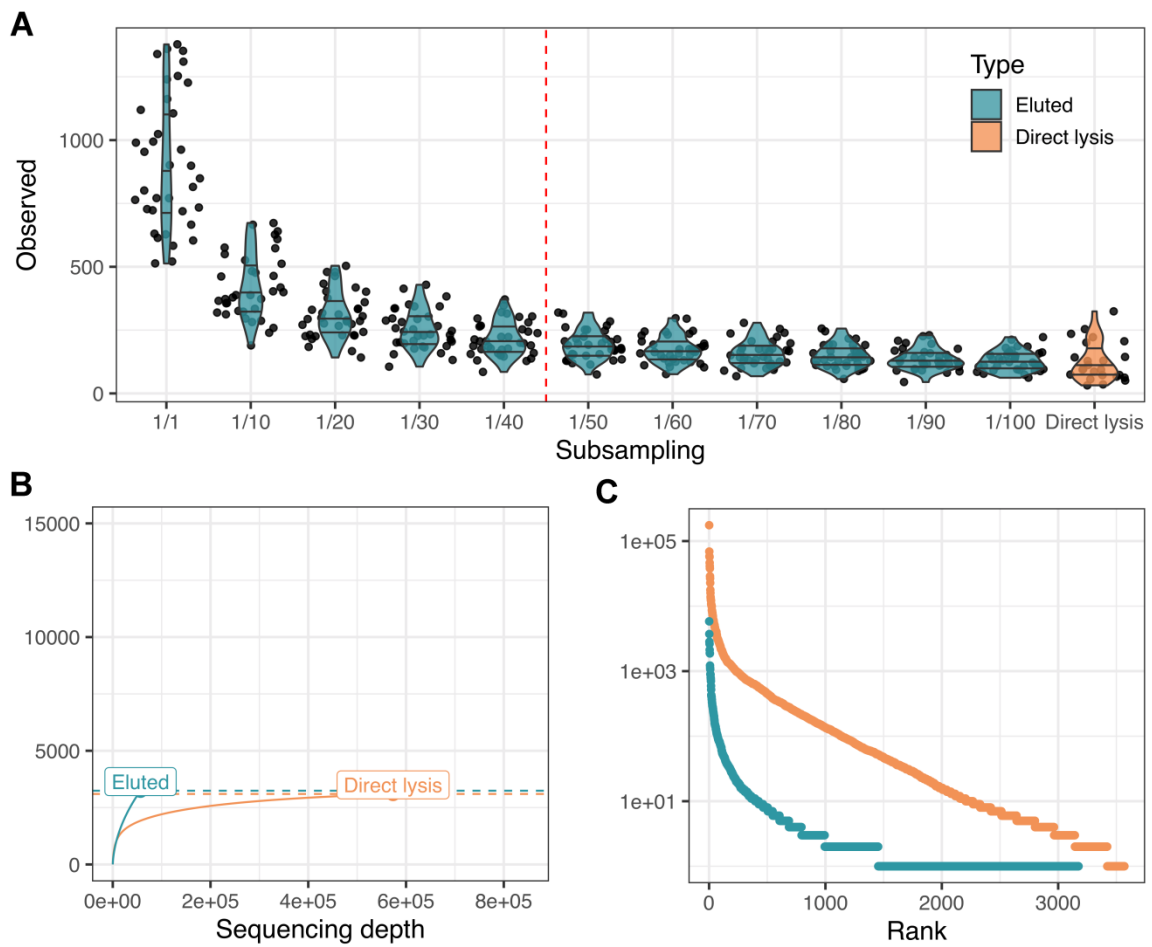


**Supplementary Figure 1: OTU accumulation curves based on the addition of samples calculated using the “random” method. A) Estimation for all samples. B) Estimation for only sediment samples. C) Estimation for only eluted samples.**

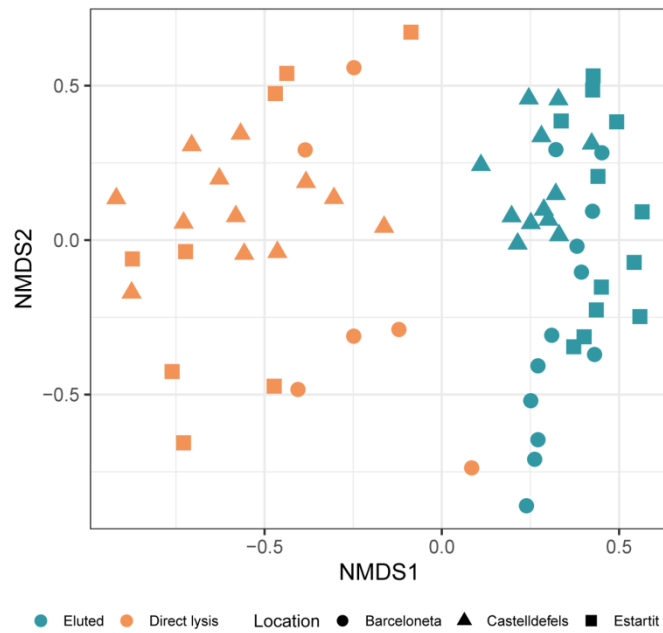




**Supplementary Figure 3: Model coefficients of the differential abundance test between eluted and direct-lysis treatments.** The line ranges are the standard error of the estimate coefficient (both positives and negatives). All results are corrected for false discovery rate <math><0.01</math>. Only significant results are shown.



**Supplementary Figure 4: Values of Chao1, Shannon and Simpson indices of eluted (blue) and direct-lysis samples (orange).** A reads reduction by multiples of ten was performed for eluted samples to simulate a reduction of sample size, from the original  $\sim 80 \text{ cm}^3$  to the  $\sim 1 \text{ cm}^3$  used for direct-lysis. The dotted red line represents the loss of significant differences (ANOVA Tukey HSD  $p < 0.05$ ) with direct-lysis dataset.



**Supplementary Figure 5: Beta-diversity analyses.** Non-metric multidimensional scaling of the dataset samples after performing a 1:50 reduction of reads in eluted samples. Blue dots represent eluted samples and orange dots direct-lysis samples.