



Comparison of two DNA extraction methods widely used in aquatic microbial ecology

Erick Mateus-Barros^{a,b}, Aylan K. Meneghini^a, Inessa Lacativa Bagatini^c, Camila C. Fernandes^{d,e}, Luciano T. Kishi^{d,e}, Armando A.H. Vieira^c, Hugo Sarmento^{a,*}

^a Universidade Federal de São Carlos (UFSCar), Department of Hydrobiology, Laboratory of Microbial Processes and Biodiversity, São Carlos, SP 13565-905, Brazil

^b Post Graduate Program in Ecology and Natural Resources (PPGERN), UFSCar, São Carlos, SP 13565-905, Brazil

^c UFSCar, Department of Botany, Laboratory of Phycology, São Carlos, SP 13565-905, Brazil

^d Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Departamento de Tecnologia, Laboratório de Bioquímica de Microrganismos e Plantas – LBMP, Jaboticabal, SP 14884-900, Brazil

^e UNESP, Faculdade de Ciências Agrárias e Veterinárias, Departamento de Tecnologia, Laboratório Multiusuário Centralizado para Sequenciamento de DNA em Larga Escala e Análise de Expressão Gênica – LMSeq, Jaboticabal, SP 14884-900, Brazil

ARTICLE INFO

Keywords:

Aquatic microbial ecology
High-throughput sequencing
Next-generation sequencing
Phenol:chloroform extraction
PowerSoil extraction kit

ABSTRACT

In recent years, the rapid advances of culture-independent methods and new molecular tools have revolutionized our understanding of microbial biodiversity and ecological functions. DNA extraction from microbial communities is a critical step in this process and several methods have been proposed and used, but the influence of the extraction method on the outcome and ultimately on ecological inferences from the results is not yet precisely determined. Here, we compared two of the most commonly used extraction methods in aquatic microbial ecology, and investigated whether the two methods yielded comparable results for community ecology analyses. We extracted DNA from 15 different shallow lakes with phenol:chloroform, a classical and widely used extraction method, and with the PowerSoil DNA isolation Kit, often suggested as the standard DNA extraction method, with some adaptations for aquatic environments. We found that although only 5% of all OTUs showed significant differences in pairwise comparisons (using the 15 lakes as replicates), these OTUs accounted for > 35% (on average) of the relative abundance. Diversity and richness did not differ significantly between the two extraction methods, but the beta-dispersion of the communities indicated that the organic extraction yielded more homogeneous communities, while the kit extraction generated variability. Consequently, we conclude that despite the small number of OTUs with significant differences, their impact on the community composition obtained was not negligible, and therefore the results from these two extraction methods were not comparable.

1. Introduction

High-throughput sequencing has become widely and routinely used by microbial ecologists. In order to achieve good sequencing results, the microbial community DNA obtained should have high quality and a sufficient quantity for downstream analyses of microbial communities, such as amplicon sequencing, metagenomics and metatranscriptomics (Logares et al., 2012).

In aquatic samples, the microbial biomass is usually obtained by filtering a certain volume of water through membranes of different porosities (e.g. Mestre et al., 2017). Sometimes, these samples have a high content of humic substances that may limit the accuracy of the method, selecting certain organisms (Mestre et al., 2017) or inhibiting DNA amplification and compromising the subsequent analyses (e.g.

Sidstedt et al., 2015; Wang et al., 2017). Therefore, it is important to use an appropriate and rigorous methodological design in order to minimize the negative effects of these factors on downstream analyses (Olson and Morrow, 2012; Taberlet et al., 2012).

Some studies have previously explored methodological approaches through different purification procedures, in both metagenomic and amplicon sequencing, aiming to maximize the quality of the material obtained (Arbeli and Fuentes, 2007; Dai et al., 2016; Wang et al., 2017). Others have compared extraction methods in order to maximize DNA integrity (Li et al., 2015; Seumahu et al., 2012), as the quality of the extracted material may affect the ability to detect some organisms or genes (Deiner et al., 2015). In this study, we investigated if ecological inferences from microbial communities may be influenced by the DNA extraction method.

* Corresponding author.

E-mail address: hsarmento@ufscar.br (H. Sarmento).

<https://doi.org/10.1016/j.mimet.2019.02.005>

Received 23 November 2018; Received in revised form 5 February 2019; Accepted 5 February 2019

Available online 06 February 2019

0167-7012/ © 2019 Elsevier B.V. All rights reserved.

Table 1

Number and percentage of OTUs that differed significantly in the paired *t*-test for the two DNA extraction methods, considering all OTUs, abundant OTUs (those that contributed at least 1% of total abundance), and rare OTUs (those that contributed < 0.1% of total abundance).

	Number of significantly different OTUs		Percentage (%) of significantly different OTUs		Impact on total relative abundance (%)		Higher relative abundance	
	Kit	Organic	Kit	Organic	Kit	Organic	Kit	Organic
Significant	156		4.82		35.24	42.81	35	121
Within abundant (> 1%)	8	10	0.25	0.31	27.24	34.67	0	10
Within rare (< 0.1%)	127	125	3.93	3.86	1.70	1.73	33	93

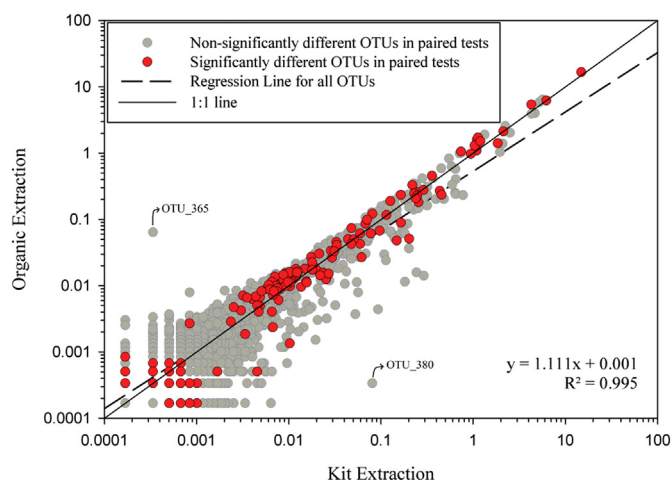


Fig. 1. Comparisons of mean relative abundances of OTUs (mean of 15 samples) obtained by two DNA extraction methods. OTUs that differed significantly in paired *t*-tests are highlighted in red. Some OTUs are far from the 1:1 line, but they were, surprisingly, not significantly different in the paired tests. This occurred because they had outliers (high abundance in a single sample), but the other samples were so homogeneous that the outliers did not affect the paired test results. OTU_365 and OTU_380 are two extreme examples of this (see Table S2 for further details). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

For this purpose, we evaluated two standard DNA extraction methods that are widely used in aquatic microbial ecology: 1) the PowerSoil DNA isolation kit (MoBio, Inc., Carlsbad, CA, USA), here termed kit extraction, proposed by the Earth Microbiome Project (Gilbert et al., 2010) and the Brazilian Microbiome Project (Pylro et al., 2014) as a standardized DNA extraction protocol; and 2) the Phenol:Chloroform extraction, here termed organic extraction.

2. Material and methods

2.1. Study sites and sampling

Water samples were taken from 15 shallow lakes located near river headwaters, dispersed throughout the state of São Paulo, Brazil. Most of these lakes are dammed creeks and are small and shallow. Several receive a considerable input of nutrients and organic carbon, making them rich in humic substances. The lakes are located between the coastal range and the Paraná River basin, in a transition between the Cerrado (Brazilian savannah biome) and the Atlantic Forest biome (semi-deciduous seasonal forest), but most of the area is now occupied by livestock and agricultural monocultures (mainly sugar cane).

In each lake, environmental data were recorded in situ and water samples were filtered to determine the chlorophyll *a* concentrations and perform DNA extraction and sequencing (see Table S1). For chlorophyll *a*, 100 to 500 ml of water was filtered through a glass-fiber filter (Macherey-Nagel® GF-6) and cooled until it was taken to the laboratory and stored in a freezer at -20°C . For molecular analyses, 500 ml was

pre-filtered through a glass-fiber filter with $1.2\mu\text{m}$ mesh (BOECO® MGC) to retain eukaryotes, large particles and attached prokaryotes. Then, 200 to 500 ml of filtrate was passed through $0.2\mu\text{m}$ polycarbonate membranes (Millipore® Isopore™ $0.2\mu\text{m}$ GTBP) to retain free-living prokaryotes, then the filters were cut in half with heat-sterilized scissors, immediately frozen, and stored at -80°C until DNA extraction.

In this study, we used only the free-living prokaryotes retained on the $0.2\mu\text{m}$ membranes. Each half-filter sampled in each shallow lake was used for one of the two DNA extraction methods described below.

2.2. Chlorophyll *a* extraction and trophic-state index measurement

Pigments were extracted with ethanol (90% v/v at 80°C) in the dark (Marker, 1980; Mush, 1980), quantified by spectrophotometry (Lorenzen, 1967), and used as a proxy of the trophic state (Carlson, 1977), in an adaptation for tropical environments (Cunha et al., 2013). We obtained a mix of samples from oligotrophic, mesotrophic and eutrophic water bodies, ensuring wide representation of different types of freshwater environments.

2.3. Extraction methods

The two extraction methods compared here are the PowerSoil DNA Isolation Kit and Phenol-Chloroform. The first was chosen from dozens of other kits because this is a method recommended by the Earth Microbiome Project (Gilbert et al., 2010, 2011) and the Brazilian Microbiome Project (Pylro et al., 2014) as the most suitable extraction method for the formulation of a standardized microbial data collection protocol applicable to any type of environmental sampling. The second was chosen because it is, historically, the extraction method most often used.

We understand that these methods have the potential to be the most often used in future studies comparing data collected and processed in different projects or studies that use data available in international databases, at least for microbial studies.

2.3.1. Extraction by PowerSoil DNA Isolation Kit

The Earth Microbiome Project (EMP) and the Brazilian Microbiome Project (BMP) propose a unified protocol for DNA extraction using a soil extraction kit, including for water samples (Gilbert et al., 2010, 2011; Pylro et al., 2014). The BMP uses the PowerSoil DNA Isolation Kit (MoBio®) for single samples, and the EMP emphasizes the protocol for plates (PowerSoil-htp 96 Well Soil DNA Isolation Kit, MoBio). These two protocols are essentially the same except with respect to the specifics for the plate or tubes and the centrifugation time. Since we used the PowerSoil DNA Isolation Kit for single samples, we followed the BMP protocol.

The PowerSoil kit is recommended as a substitute for the PowerWater kit because the former is capable of extracting DNA from filtered samples, with some simple adaptations of the original protocol, such as the centrifugation time or the incubation temperature; while the latter is not able to extract good-quality DNA from soil samples. This makes the PowerSoil kit the most suitable for the purposes of the EMP,

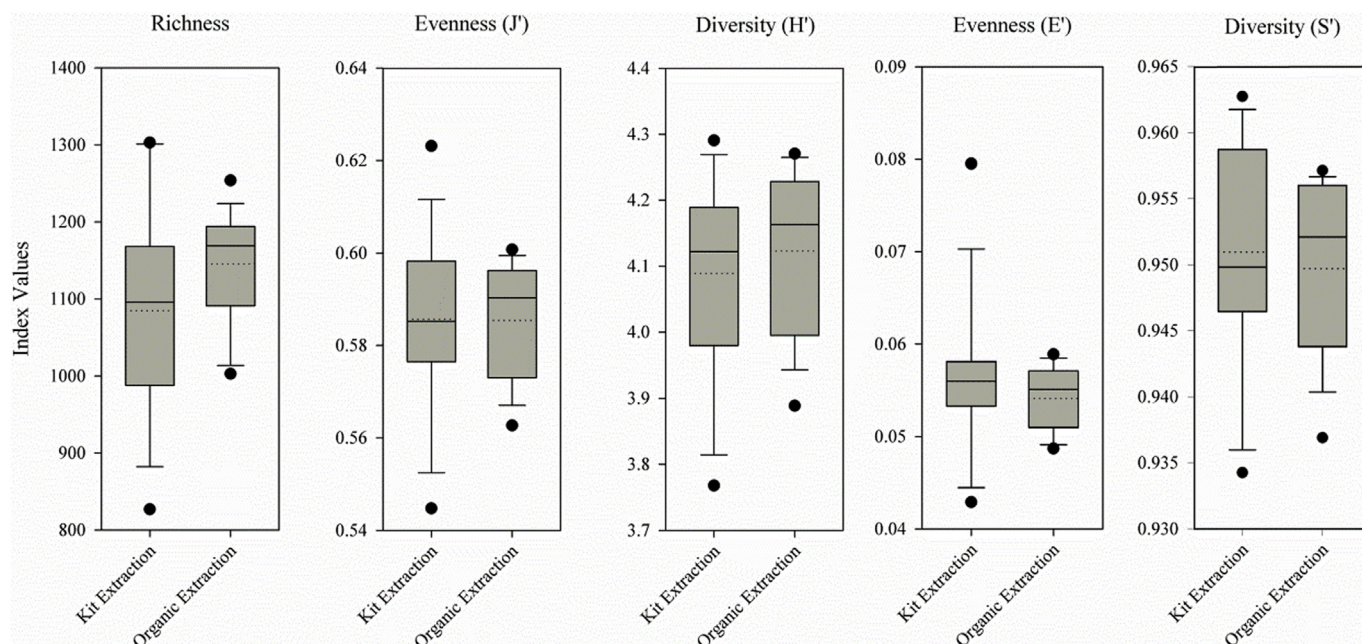


Fig. 2. Comparison of commonly used indexes in community ecology: richness, Shannon-Weaver (H') and Simpson (S') diversity indexes and Pielou (J') and Hill (E') evenness indexes were obtained for each DNA extraction method. No significant differences were observed by paired t -test comparisons. The central solid line indicates the median value, the dotted line indicates the arithmetic mean value, the boxes indicate the lower and upper quartiles, the vertical lines indicate the 10th and 90th percentiles, and the dots represent the 5th and 95th percentiles.

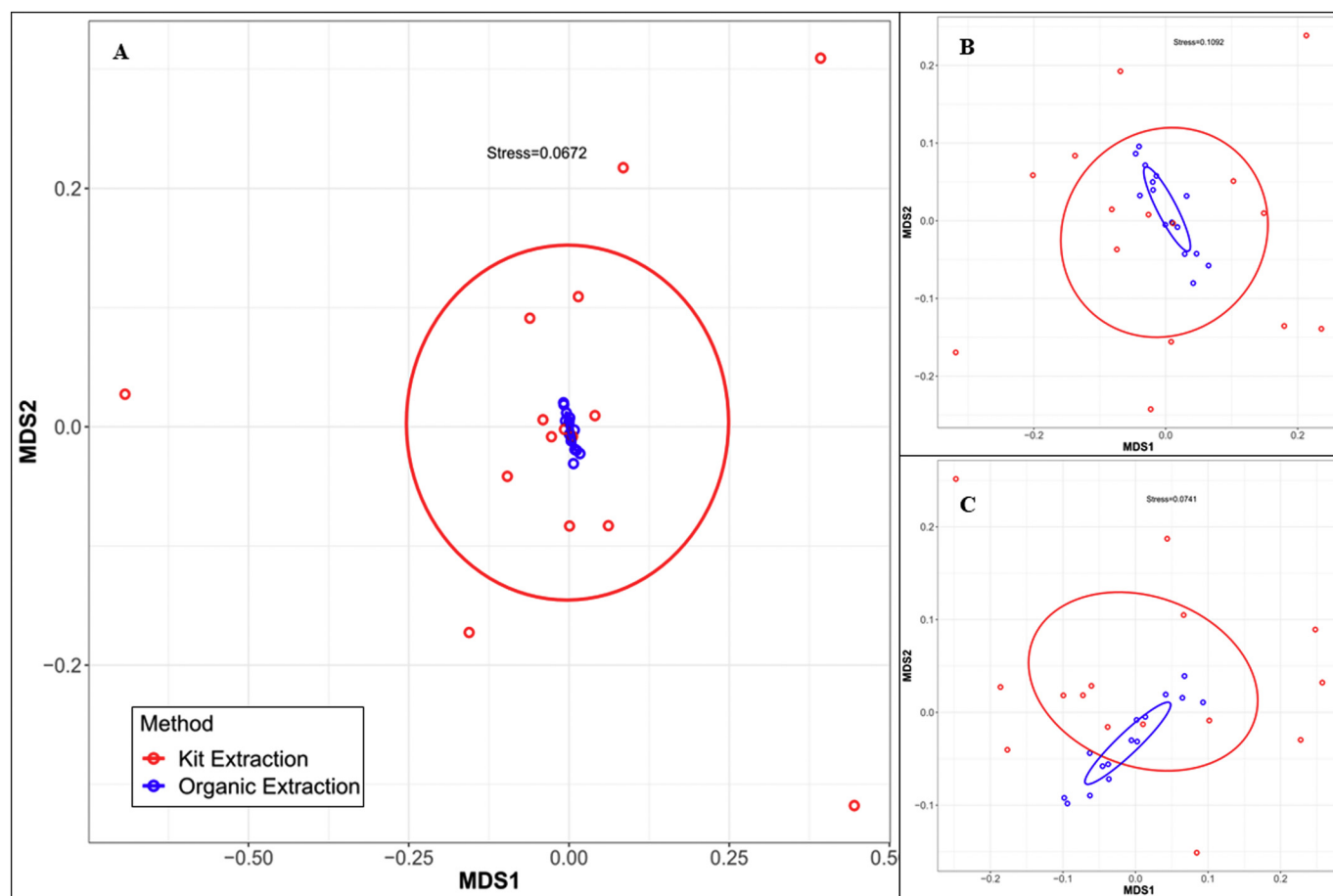


Fig. 3. Community dissimilarity between extraction methods. Bray-Curtis dissimilarity distance nMDS of bacterial community composition obtained by Kit Extraction (red) and Organic Extraction (blue) for OTUs (A), Classes (B) and Phyla (C). Circles are standard deviation ellipses with 95% confidence, calculated with the function `ordellipse` of the `vegan` package in R. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to create a standard protocol for any type of environment.

Below we describe the steps of the modified protocol used for our samples, aiming to maximize the yield of material from samples with high humic-acid contents. The fully detailed BMP protocol can be found at (<http://www.brmicrobiome.org/dna-mobio—water>).

Each half-filter membrane containing free-living bacteria was inserted into the “Power Bead Tubes” and then manually mixed. After the C1 solution was added, the samples were incubated at 65 °C in a water bath for 10 min (according to the EMP protocol), vortexed for 5 min and centrifuged for 1 min at 10,000 g. Following the protocol, we added the C2 solution and incubated the samples at –20 °C for 5 min. This incubation step was added after laboratory tests that showed a slight increase in the quantity and quality of genetic material obtained from filtered samples.

2.3.2. Extraction by phenol:chloroform method

For the extraction by the phenol:chloroform protocol, each half-filter was cut into small pieces and incubated with 3 ml of lysis buffer (40 nM of EDTA, 50 nM of Tris pH 8.3, 0.75 M of sucrose) and 75 µl of lysozyme (2 mg of lysozyme, 50 µl of lysis buffer) for 45 min at 37 °C, and then with 75 µl of proteinase K (0.4 mg proteinase K and 50 µl lysis buffer) and 300 µl of sodium dodecyl sulfate (SDS) 10% (300 ml of sterile water and 30 g SDS) for 60 min at 55 °C.

The incubated material was centrifuged three times, for 10 min each time: the first two after adding 2 ml of phenol-chloroform-isoamyl mixture (25:24:1) and the last one after the addition of 2 ml of chloroform-isoamyl mixture. After each centrifugation step, the supernatant was separated from the waste material. The extracted material was then purified in 100 KDa Amicon® columns (Millipore® 100,000 MWCO) with sterile water.

In order to clear the samples that were rich in humic substances, we performed a purification method after the extraction, using 10% cetyl trimethyl ammonium bromide (CTAB) (Schneegurt et al., 2003). This method consists of one incubation step with 5 M NaCl and CTAB, followed by centrifugation with 128 µl of chloroform-isoamyl. After the addition of 256 µl of 100% ethanol, the samples were incubated at –20 °C for at least 15 min and then centrifuged for 5 min. Then, the extraction continued with two centrifugation rounds of 2 min after the addition of 125 µl of 70% ethanol. After each centrifugation, the supernatant was removed and the pellet retained. Finally, the material was dried to eliminate the remaining ethanol and resuspended in ultrapure water.

2.4. Amplification and sequencing

Despite the purification procedures described above, we still had to dilute the samples in ultrapure water in a 1:5 proportion in order to prevent PCR inhibition by humic substances.

The V3-V4 regions of the 16S rRNA gene were amplified with the 341F (5'-CCTACGGGNGGCWGCAG-3') and 805R (5'-GACT ACHVGG-GTATCTAATCC-3') primers (Herlemann et al., 2011) with a 10 µM concentration, using the KAPA HiFi HotStart ReadyMix PCR Kit (Kapa Biosystems®). The thermocycling followed the protocol proposed by the PCR kit manufacturer, with 55 °C annealing temperature.

We then conducted a bead purification step and multiplex indexation with the Nextera XT Index Kit v2 (Illumina®), followed by another bead purification. In the bead purification, we added the same volume of magnetic beads to each sample (Agencourt® AMPure® XP, Beckman Coulter®). For bead purifications, the samples were incubated for 2 min at room temperature, placed on a magnetic rack, and incubated for 5 more min. The supernatant was then discarded. With the samples still in the magnetic rack, we added 200 µl of 80% ethanol, incubated for 30 s, and discarded the supernatant. This step was performed two times. After drying the pellet for 15 min, we removed the material from the magnetic rack, added 20 µl of ultrapure water, incubated for 5 min, replaced the samples in the magnetic rack, incubated for another 5 min,

and finally transferred the samples to other tubes.

Samples were sequenced in an Illumina MiSeq platform (2 × 250 bp) at the Universidade Estadual Paulista – UNESP, Departamento de Tecnologia, Laboratório Multiusuário Centralizado para Sequenciamento de DNA em Larga Escala e Análise de Expressão Gênica – LMSeq, located in Jabotical (São Paulo), Brazil. The 16S rRNA amplicon results have been deposited in the NCBI repository under accession number PRJNA411849.

2.5. Bioinformatics and statistics

Sequence data were processed using a UPARSE-based pipeline internally implemented (Logares et al., 2014; Logares, 2017). This pipeline uses UPARSE routines (Edgar, 2013) to recognize the paired sequences, clear of singletons and chimeras, and compare with the SILVA database, version 128 (Quast et al., 2013; Yilmaz et al., 2014), to determine OTUs at the 97% similarity threshold.

The statistical analyses were carried out in R Software, version 3.3.1 (R Core Team, 2016). Total reads were obtained after rarefaction by the sample with the lowest number of sequences, using the “rarefy” function from the “vegan” package (Oksanen et al., 2016). Richness, alpha-diversity and evenness indexes were calculated with the “BiodiversityR” package (Kindt and Coe, 2005), using the Simpson and Shannon-Weaver indexes for diversity and Hill's and Pielou's indexes for evenness. The R scripts used in these analyses can be found in: <https://github.com/LMPB/comparison-of-DNA-extraction-methods>.

3. Results

We obtained a mean of 184,114 reads per sample, and after rarefaction, all samples were normalized with 39,599 reads, the number in the sample with the lowest coverage. A matrix of 3224 OTUs (2761 for organic extraction; 2878 for kit extraction) with a mean of 1115 OTUs per sample (1169 for organic extraction; 1096 for kit extraction) was constructed. The minimum number of OTUs/sample was 827 (1003 for organic extraction; 827 for kit extraction) and the maximum was 1303 (1254 for organic extraction; 1303 for kit extraction).

Paired *t*-tests (or Wilcoxon signed-rank tests for those that were not normally distributed) for each OTU, taking the 15 samples as replicates, indicated that < 5% of the OTUs (156 of the total 3234 OTUs) had significantly different relative abundances. Still, these OTUs comprised 36% of the kit and 43% of the organic extraction mean relative abundances (Tables 1 and S2). Of these, 10 OTUs were in the abundant group (those that contributed at least 1% of total abundance), and all had higher relative abundances in the organic extraction. Most differences were found in less-abundant taxa, and most of these also had higher relative abundances in the organic extraction (Table 1).

More OTUs had higher relative abundances in the organic extraction, and these were more aggregated and closer to the 1:1 line. OTUs with higher relative abundances in the kit extraction were distributed more heterogeneously, with the points more scattered below the 1:1 line of the plots (Figs. 1 and S2).

Classes that had higher relative abundances in the organic extraction were representatives of the phyla Actinobacteria, Bacteroidetes, Cyanobacteria, Chloroflexi, Parcubacteria, Planctomycetes, Saccharibacteria and Verrucomicrobia. The exception was the class Chloroflexia, a rare group of the phylum Chloroflexi, which was over-represented by the kit extraction (see Fig. S3 and Table S3 for further details). Phyla that had higher relative abundances in the organic extraction were Planctomycetes and Chloroflexi; the only such phylum in the kit extraction was Verrucomicrobia (see Fig. S4 and Table S4).

We also compared some indexes that are commonly used in community ecology. Although all indexes obtained with the extraction kit varied more widely (Fig. 2), the results of the two extraction methods did not differ significantly in richness, alpha diversity or evenness (Table S5).

However, a more thorough analysis of community composition revealed a different situation. The beta-dispersion permutation test gave significantly different results for OTUs, classes and phyla ($P < .05$; Figs. 3 and S5). The PERMANOVA test was significant only at the OTU level ($P = 9.999 \times 10^{-5}$) and not significant at the class ($P = .089$) and phylum ($P = .059$) levels. These results must be interpreted cautiously because of the significant differences in beta-dispersion. If we consider beta-dispersion as a proxy for beta diversity (Anderson et al., 2006), we can infer that the kit extraction resulted in higher dissimilarity (Figs. 3 and S5).

4. Discussion

Although most of the ecological indexes of diversity were similar (except in beta-diversity inferences), these two methods indicated significantly different community compositions. These results resemble those reported for a marine environment by Djurhuus et al. (2017) and to a lesser extent by Janabi et al. (2016), who reported differences in results, including for diversity indexes.

These results allow us to reach two opposite interpretations, one that would attest the superiority of the organic extraction and another that advocates in favor of the kit extraction.

The first possible interpretation of these results is that the organic method is more efficient in extracting some of the most abundant OTUs, which are therefore overrepresented in relation to the kit extraction method. The data appear to indicate this difference, since most OTUs are overrepresented by the organic extraction (Table 1 and Fig. 1), and this tendency can also be observed at the level of classes (Table S3) and phyla (Table S4).

The alternative explanation for these results is that the observed tendency of OTUs to fall above the 1:1 line may have been caused by an artifact derived from the compositional nature of sequencing data. High-throughput sequencing technologies have limited sequence-reading capacity, and therefore the final result is a random sampling that represents the relative abundance of the sequences obtained in a fixed number of reads that the equipment can achieve (Gloor et al., 2017). As the values obtained are relative instead of absolute, samples with different organism abundances may show similar relative abundances (and vice-versa). In our case, the kit extraction showed more variation in the less-abundant OTUs (Figs. 1 and S2), which would eventually push the abundant OTUs downward. As the organic extraction resulted in less variation in the number of OTUs, the abundant OTUs would appear to be more abundant in relation to the rarer OTUs, and the comparison between the values obtained with the different extraction methods would create the illusion that these more-abundant organisms would be better extracted by organic extraction.

These two interpretations explaining the over- or under-representation from the organic and kit methods are equally plausible, and our data do not provide sufficient information to eliminate either of them. Although it is impossible to determine if one method is more accurate than other, it is clear that they may lead to different interpretations of ecological results. The finding of similar results in environments as diverse as shallow lakes (this study), seawater (Djurhuus et al., 2017) and fecal samples (Janabi et al., 2016) suggests that this method-dependent variation does not change with the habitat. Therefore, mixing samples drawn with these different types of extraction methods is not recommended.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.mimet.2019.02.005>.

Acknowledgments

We thank three anonymous reviewers who helped to improve our manuscript. This research was funded by the São Paulo Research Foundation - FAPESP (projects 2014/14139-3 and 2011/50054-4). We thank the LMSeq Laboratory of UNESP for sequencing (grants 2009/

53984-2 and 2016/07679-7). EMB was funded by a CAPES scholarship. HS was supported by CNPq (grant 309514/2017-7). ILB was supported by FAPESP (grant 2013/18083-0). Finally, we are grateful to Cilene C. Mori, Fabrício S. Macccheri, Guilherme P. de Moraes, Helena H. Vieira, Ingrid C. Moreira, Letícia P. Tessarolli, Lucas S. Tiberti, Luiz A. Joaquim, Naiara C. P. Archanjo, Rodrigo V. de Mello and Thaís G. da Silva, who participated in the field sampling.

Conflict of interests

None declared.

References

- Anderson, M.J., Ellingsen, K.E., McArdle, B.H., 2006. Multivariate dispersion as a measure of beta diversity. *Ecol. Lett.* 9, 683–693. <https://doi.org/10.1111/j.1461-0248.2006.00926.x>.
- Arbeli, Z., Fuentes, C.L., 2007. Improved purification and PCR amplification of DNA from environmental samples. *FEMS Microbiol. Lett.* 272, 269–275. <https://doi.org/10.1111/j.1574-6968.2007.00764.x>.
- Carlson, R.E., 1977. A trophic state index for lakes. *Limnol. Oceanogr.* 22, 361–369. <https://doi.org/10.4319/lo.1977.22.2.0361>.
- Cunha, D.G.F., Calijuri, M.d.C., Lamparelli, M.C., 2013. A trophic state index for tropical/subtropical reservoirs (TSITSr). *Ecol. Eng.* 60, 126–134. <https://doi.org/10.1016/j.ecoleng.2013.07.058>.
- Dai, X., Chen, S., Li, N., Yan, H., 2016. Quantitative and qualitative validations of a sonication-based DNA extraction approach for PCR-based molecular biological analyses. *Anal. Biochem.* 501, 44–46. <https://doi.org/10.1016/j.ab.2016.01.003>.
- Deiner, K., Walser, J.-C., Mächler, E., Altermatt, F., 2015. Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biol. Conserv.* 183, 53–63. <https://doi.org/10.1016/j.biocon.2014.11.018>.
- Djurhuus, A., Port, J., Closek, C.J., Yamahara, K.M., Romero-Maraccini, O., Walz, K.R., Goldsmith, D.B., Michisaki, R., Breitbart, M., Boehm, A.B., Chavez, F.P., 2017. Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. *Front. Mar. Sci.* 4 (314), 1–11. <https://doi.org/10.3389/fmars.2017.00314>.
- Edgar, R.C., 2013. UNPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. <https://doi.org/10.1038/nmeth.2604>.
- Gilbert, J.A., Meyer, F., Knight, R., Field, D., Kyrpides, N., Yilmaz, P., Wooley, J., 2010. Meeting report: GSC M5 roundtable at the 13th International Society for Microbial Ecology meeting in Seattle, WA, USA August 22–27, 2010. *Stand. Genomic Sci.* 3, 235–239. <https://doi.org/10.4056/sigs.1443528>.
- Gilbert, J.A., Bailey, M., Field, D., Fierer, N., Fuhrman, J.A., Hu, B., Jansson, J., Knight, R., Kowalchuk, G.A., Kyrpides, N.C., 2011. The earth microbiome project: the meeting report for the 1st international earth microbiome project conference, Shenzhen, China, June 13th–15th 2011. *Stand. Genomic Sci.* 5, 243–247. <https://doi.org/10.4056/sigs.2134923>.
- Gloor, G.B., MacLaim, J.M., Pawlowsky-Glahn, V., Egozcue, J.J., 2017. Microbiome datasets are compositional: and this is not optional. *Front. Microbiol.* 8, 2224. <https://doi.org/10.3389/fmicb.2017.02224>.
- Herlemann, D.P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., Andersson, A.F., 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* 5, 1571–1579. <https://doi.org/10.1038/ismej.2011.41>.
- Janabi, A.H.D., Kerkhof, L.J., McGuinness, L.R., Biddle, A.S., McKeever, K.H., 2016. Comparison of a modified phenol/chloroform and commercial kit methods for extracting DNA from horse fecal material. *J. Microbiol. Methods* 129, 14–19. <https://doi.org/10.1016/j.mimet.2016.07.019>.
- Kindt, R., Coe, R., 2005. *Tree Diversity Analysis: A Manual and Software for Common Statistical Methods for Ecological and Biodiversity Studies*. World Agroforestry Centre (ISBN: 929059179X).
- Li, P., Yang, S.F., Lv, B.B., Zhao, K., Lin, M.F., Zhou, S., Song, X., Tang, X.M., 2015. Comparison of extraction methods of total microbial DNA from freshwater. *Genet. Mol. Res.* 14, 730–738. <https://doi.org/10.4238/2015.January.30.16>.
- Logares, R., 2017. Ramalok/amplicon_processing: Workflow for Analysing MiSeq Amplicons Based on Uparse. <https://doi.org/10.5281/zenodo.259579>.
- Logares, R., Haverkamp, T.H., Kumar, S., Lanzén, A., Nederbragt, A.J., Quince, C., 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. <https://doi.org/10.1016/j.mimet.2012.07.017>.
- Logares, R., Sunagawa, S., Salazar, G., Cornejo-Castillo, F.M., Ferrera, I., Sarmiento, H., Hingamp, P., Ogata, H., de Vargas, C., Lima-Mendez, G., Raes, J., Poulain, J., Jaillon, O., Wincker, P., Kandels-Lewis, S., Karsenti, E., Bork, P., Acinas, S.G., 2014. Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environ. Microbiol.* 16, 2659–2671. <https://doi.org/10.1111/1462-2920.12250>.
- Lorenzen, C.J., 1967. Determination of chlorophyll and phaeo-pigments: spectrophotometric equations. *Limnol. Oceanogr.* 12, 343–346. <https://doi.org/10.4319/lo.1967.12.2.0343>.
- Marker, A., 1980. The measurement of photosynthetic pigments in freshwaters and standardization of methods: conclusions and recommendations. *Arch. Hydrobiol.*

- Ergebn. Limnol. 14, 91–106.
- Mestre, M., Borrell, E., Sala, M., Gasol, J.M., 2017. Patterns of bacterial diversity in the marine planktonic particulate matter continuum. *ISME J.* 11, 999–1010. <https://doi.org/10.1038/ismej.2016.166>.
- Mush, E., 1980. Comparison of different methods for chlorophyll and phaeopigment determination. *Arch. Hydrobiol. Beih.* 14, 14–36.
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2016. *Vegan: Community Ecology Package*. <https://CRAN.R-project.org/package=vegan>.
- Olson, N.D., Morrow, J.B., 2012. DNA extract characterization process for microbial detection methods development and validation. *BMC Res. Notes* 5, 668. <https://doi.org/10.1186/1756-0500-5-668>.
- Pylro, V.S., Roesch, L.F.W., Ortega, J.M., do Amaral, A.M., Tótola, M.R., Hirsch, P.R., Rosado, A.S., Góes-Neto, A., da Costa Silva, A.L., Rosa, C.A., Morais, D.K., Andreote, F.D., Duarte, G.F., de Melo, I.S., Seldin, L., Lambais, M.R., Hungria, M., Peixoto, R.S., Kruger, R.H., Tsai, S.M., Azevedo, V., 2014. Brazilian microbiome project: revealing the unexplored microbial diversity—challenges and prospects. *Microb. Ecol.* 67, 237–241. <https://doi.org/10.1007/s00248-013-0302-4>.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glockner, F., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. <https://doi.org/10.1093/nar/gks1219>.
- R. Core Team, 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Schneegurt, M.A., Dore, S.Y., Kulpa Jr., C.F., 2003. Direct extraction of DNA from soils for studies in microbial ecology. *Curr. Issues Mol. Biol.* 5, 1–8.
- Seumahu, C.A., Suwanto, A., Rusmana, I., Solihin, D.D., 2012. Comparison of DNA extraction methods for microbial community analysis in Indonesian tempe employing amplified ribosomal intergenic spacer analysis. *Hayati J. Biosci.* 19, 93–98. <https://doi.org/10.4308/hjb.19.2.93>.
- Sidstedt, M., Jansson, L., Nilsson, E., Noppa, L., Forsman, M., Rådström, P., Hedman, J., 2015. Humic substances cause fluorescence inhibition in real-time polymerase chain reaction. *Anal. Biochem.* 487, 30–37. <https://doi.org/10.1016/j.ab.2015.07.002>.
- Taberlet, P., Coissac, E., Hajibabaei, M., Rieseberg, L.H., 2012. Environmental DNA. *Mol. Ecol.* 21, 1789–1793. <https://doi.org/10.1111/j.1365-294X.2012.05542.x>.
- Wang, H., Qi, J., Xiao, D., Wang, Z., Tian, K., 2017. A re-evaluation of dilution for eliminating PCR inhibition in soil DNA samples. *Soil Biol. Biochem.* 106, 109–118. <https://doi.org/10.1016/j.soilbio.2016.12.011>.
- Yilmaz, P., Parfrey, L., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., Glockner, F., 2014. The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Res.* 42, D643–D648. <https://doi.org/10.1093/nar/gkt1209>.