



Rare Plankton Subcommunities Are Far More Affected by DNA Extraction Kits Than Abundant Plankton

Min Liu^{1,2}, Yuanyuan Xue^{1,2} and Jun Yang^{1*}

¹ Aquatic EcoHealth Group, Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China, ² University of Chinese Academy of Sciences, Beijing, China

OPEN ACCESS

Edited by:

Raju Sekar,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Jian Yang,
China University of Geosciences,
China
Yusuke Okazaki,
National Institute of Advanced
Industrial Science and Technology
(AIST), Japan

*Correspondence:

Jun Yang
jyang@iue.ac.cn

Specialty section:

This article was submitted to
Aquatic Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 25 October 2018

Accepted: 20 February 2019

Published: 11 March 2019

Citation:

Liu M, Xue Y and Yang J (2019)
Rare Plankton Subcommunities Are
Far More Affected by DNA Extraction
Kits Than Abundant Plankton.
Front. Microbiol. 10:454.
doi: 10.3389/fmicb.2019.00454

Advances in high-throughput sequencing technologies allow a more complete study of microbial plankton community composition and diversity, especially in the rare microbial biosphere. The DNA extraction of plankton is a key step for such studies; however, little is known about its influences on the abundant or rare microbial biosphere. Our aim was to quantify the influences of different DNA extraction kits on abundant and rare plankton in the surface waters of a reservoir and provide a reference for the comparisons between microbial community studies with different extraction methods. We evaluated the influence of five common commercial kits on DNA quality, microbial community diversity and composition, and the reproducibility of methods using both 16S and 18S rRNA genes amplicon sequencing. Our data showed that results of Fast DNA Spin Kit for Soil (MPF) had higher α diversity for bacteria and high DNA quality, indicating that it is the most suitable approach for bacterioplankton diversity study. However, DNeasy Blood & Tissue Kit (QD) and QIAamp DNA Mini Kit (QQ) methods could produce results that are easier to replicate for bacteria and eukaryotes, respectively, and were more comparable between studies. The use of different DNA extraction kits had larger influence on the rare taxa compared with abundant taxa. Therefore, the comparability between studies that employed different extraction methods can be improved by removing low-abundance or less-representative OTUs. Collectively, this study provides a comprehensive assessment of the biases associated with DNA extraction for plankton communities from a freshwater reservoir. Our results may guide researchers in experimental design choices for DNA-based ecological studies in aquatic ecosystem.

Keywords: DNA extraction, high-throughput sequencing, plankton, species diversity, community composition

INTRODUCTION

Recently, with the rapid development of sequencing technology DNA-based experiments have become routinely used to study the microbial community in various ecosystems (Shendure et al., 2017; Sinha et al., 2017; Knight et al., 2018). Microorganisms in aquatic ecosystems (i.e., oceans, rivers, lakes and reservoirs) which are some of the most studied ecosystems for their high microbial diversity, play key roles in biogeochemical processes (Liu et al., 2015; Sunagawa et al., 2015; Xue et al., 2018). Numerous studies have shown DNA-based approaches may provide unprecedented

insight into the dynamics of microbial communities, and reveal general principles about their ecology and mechanisms of community assembly. However, several technical limitations are still underexplored in acquiring the DNA from water samples and constitute a key constraint to the accuracy of the new findings (Deiner et al., 2015; Walden et al., 2017).

Generally, different DNA extraction methods are recommended based on the origin of the samples as there is not a single superior method which can be applied across all samples and microbes (Kuhn et al., 2017; Sinha et al., 2017; Walden et al., 2017). Studies have showed suitable extraction methods for specific samples such as human fecal (Costea et al., 2017), soil protist (Santos et al., 2015, 2017) and fish (Eichmiller et al., 2016). The variable structure of microbial cell walls can lead to a misrepresentation of specific microbial taxa, increasing the difficulty of data comparison between studies (Cabeen and Jacobs-Wagner, 2005; Santos et al., 2017). So far, few researches have evaluated the influences of different extraction methods on bacterial and eukaryotic plankton communities, respectively (Eland et al., 2012; Albertsen et al., 2015; Li et al., 2015; McCarthy et al., 2015; Walden et al., 2017). Further, we have very limited knowledge how DNA extraction methods influence the plankton community of bacteria and eukaryotes together, with an emphasis on the comparison between abundant and rare plankton.

An important aspect that should be considered is the influence of DNA extraction methods on the rare biosphere of the microbial community. The study of these rare taxa has become a very active research area, due to the development of fast and cheap high-throughput sequencing and their important and unknown ecological roles (Sogin et al., 2006; Pedrós-Alió, 2012; Liu et al., 2015, 2017, 2019; Lynch and Neufeld, 2015; Xue et al., 2018). However, until now, only a few studies have evaluated the influences of DNA extraction kits on rare taxa (**Supplementary Table S1**). Limited studies have showed that kits contamination of different methods may be different and would have a high influence on microbial samples with low biomass (Salter et al., 2014; Velásquez-Mejía et al., 2018). To better compare results collected from different research groups or studies, the evaluation of the influence of DNA extraction methods on taxa in different relative abundances is an urgent call.

The reproducibility of each DNA extraction method is another important aspect which should be considered and assessed. The reproducibility has a very important role in comparing data from different studies, labs, or even within the same study (Zhou et al., 2011; Sinha et al., 2017). Here we collected one surface water sample from a subtropical reservoir, and systematically tested how distinct DNA extraction methods influence on the taxa in different relative abundances (i.e., rare and abundant taxa), instead of the effect of downstream analysis (i.e., PCR, primer choice, DNA sequencing platform, and bioinformatics). Additionally, sequencing depth was also considered to better estimate the reproducibility of each DNA extraction method.

The aim of this study was to inform the experimental design by quantifying the relative influence of DNA extraction on plankton taxa with different relative abundances, and to provide a reference for the comparison of different studies on microbial

plankton communities, instead of obtain the “best” protocols for aquatic microbiome studies. We compared five commonly used DNA extraction kits, using the 16S rRNA and 18S rRNA gene amplicon sequences data as the readout, and evaluated taxonomic variability of both bacterial and microeukaryotic plankton. Specifically, we assessed: (i) the influence of DNA extractions on the bacterial and eukaryotic diversity; (ii) the influence of DNA extraction kits on plankton community composition based on different distances; (iii) the similarity of community composition detected from the replicate extractions, which indicates the reproducibility of each method.

MATERIALS AND METHODS

Experimental Design

Water samples (upper 50 cm, 15 litter) were collected by a 5-L polymethyl methacrylate sampler in Tingxi Reservoir (24°48'N, 118°08'E) on July 26, 2016. Then all waters were put into a 20-L PVC bottle and transferred to the laboratory as soon as possible. The samples were first filtered through a 200 μ m pore-size sieve to remove debris, large metazoans and grains. Then plankton communities (500 mL water) were collected on each 0.22 μ m pore-size polycarbonate filter (47 mm diameter, Millipore, Billerica, MA, United States). The water was well mixed in the 20-L PVC bottle before filtering to keep the uniformity of each filtered sample. The filters were then stored at -80°C until DNA extraction.

Five commercial DNA extraction kits from three companies which are commonly used to extract DNA from environmental samples (Lear et al., 2018), were evaluated in this study (**Table 1**). Commercial DNA extraction kits were used in this study for their standardized application in multiple-labs (Renshaw et al., 2015; Walden et al., 2017). The DNA extractions were carried out in triplicate for each kit, and subsequent analyses were performed on the 15 individual extracts. All triplicates were taken from the same water and same bottle (20-L PVC bottle). Negative control extractions, where new membranes without plankton samples were added, were also performed for each method in triplicate. Each membrane or filter represents one replication. We mostly followed the instructions of manufacturers but introduced changes to improve the comparability among different DNA extraction kits. For MBS kit, to make it comparable to other DNA extraction kits, instead of filtering by SterivexTM filter unit, the filtered membrane was cut into small pieces and put into a sterilization centrifuge tube. Then we added the 0.9 mL of solution ST1B, vortexed as instruction and followed by adding 0.9 mL of ST2. After incubation at 90°C for 5 min, the mixtures were cooled at room temperature for 2 min. Next, the mixtures were vortexed at maximum speed for 5 min. After this, the lysate was added into the 5 mL PowerWater[®] SterivexTM glass Bead Tube. The other steps followed the manufacturer instructions to extract the DNA. The extracted DNA was quantified by a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Pittsburgh, PA, United States) and stored at -20°C until further use. Extracts were considered to contain sufficiently pure genomic DNA when their A260/A280 nm ratio was between 1.8

TABLE 1 | The five common DNA extraction kits used in this study.

| Extraction kits | Abbreviation | Lysis method | Company | Headquarter address |
|---|--------------|---------------------------|-------------------------|------------------------------|
| PowerWater DNA isolation kit | MB | Mechanical, chemical | MoBio Laboratories Inc. | Carlsbad, CA, United States |
| PowerWater® Sterivex™ DNA isolation kit | MBS | Mechanical, chemical | MoBio Laboratories Inc. | Carlsbad, CA, United States |
| Fast DNA spin kit for soil | MPF | Mechanical, chemical | MP Biomedicals | Santa Ana, CA, United States |
| DNeasy blood and tissue kit | QD | Heat, chemical, enzymatic | Qiagen | Hilden, Germany |
| QIAamp DNA mini kit | QQ | Heat, chemical | Qiagen | Hilden, Germany |

and 2.0. Otherwise, the extracted DNA included proteins, phenols or other contaminants (Hermans et al., 2018).

PCR Amplification and Illumina Sequencing

We used primer pairs targeting the V3-V4 variable region of 16S rRNA gene in bacteria (341F, 5'-CCTAYGGGRBGCASCAG-3'; 806R, 5'-GGACTACNNGGTATCTAAT-3') (Yu et al., 2005; Sundberg et al., 2013; Guo et al., 2018) and the V9 variable region of the 18S rRNA gene in eukaryotes (1380F, 5'-CCCTGCCHTTTGTACACAC-3'; 1510R, 5'-CCTTCYGCAGGTTACCTAC-3') (Amaral-Zettler et al., 2009; Liu et al., 2017; Xue et al., 2018). Each DNA sample and negative control were run in triplicate. Each DNA sample was individually PCR-amplified in 30 μ L reactions included an initial denaturation at 98°C for 1 min, followed by 30 cycles of 10 s at 98°C, 30 s at 50°C, and 30 s at 72°C. At the end of the amplification, the amplicons were subjected to final 5 min extension at 72°C. The 30 μ L PCR mixture included 15 μ L of Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Beverly, MA, United States); 0.2 μ M of forward and reverse primers, and about 10 ng template DNA for both bacteria and eukaryotes. The triplicate PCR products were mixed in equimolar amounts and were confirmed after running in 1% agarose gel. Then the PCR products were isolated from the gel and purified with GeneJET Gel Extraction Kit (Thermo Fisher Scientific, Hudson, NH, United States). Sequencing libraries were constructed using NEB Next Ultra DNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, United States) following manufacturer's instructions, and barcodes were added. The library quality was estimated on the Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, United States) and Agilent Bioanalyzer 2100 system (Agilent Technologies, Palo Alto, CA, United States). At last, the library was sequenced on the Illumina MiSeq platform (Illumina, Inc., San Diego, CA, United States) using a 250 bp paired-end protocol (Liu et al., 2017).

Bioinformatics

Paired-end reads were merged by using FLASH (Magoč and Salzberg, 2011) and then assigned to each sample according to the unique barcodes. Sequence data were processed using quantitative insights into microbial ecology (QIIME v.1.8.0) software following standard protocols: maximum number of consecutive low-quality base = 3; minimum of continuous high-quality base = 75% of total read length; maximum number of ambiguous bases = 0 (Caporaso et al., 2010). Chimeric sequences

were identified by UCHIME and discarded before further analysis (Edgar et al., 2011). Quality-filtered sequences were then assigned to operational taxonomic units (OTUs) at 97% level of sequence similarity by using the pick_otus.py. The 97% threshold has been widely used for both of bacteria and eukaryotes (Henderson et al., 2013; Liu et al., 2015, 2017; Dai et al., 2016; Hermans et al., 2018; Xue et al., 2018). There is no general agreement on a standard definition to classify the eukaryotes into OTUs at species level. We selected 97% threshold for eukaryotic plankton to facilitate comparisons between studies because many previous studies used this threshold to define OTUs (Liu et al., 2017; Xue et al., 2018). Further, our previous study (Liu et al., 2017) found that the choice of different thresholds (97% vs. 99%) had no apparent effect on overall results and general conclusions in plankton community ecology. Sequences were taxonomically classified by the RDP classifier using the 80% confidence threshold against the Silva 123 for bacteria (Quast et al., 2013) and PR² for eukaryotes (Guillou et al., 2013), respectively. For bacteria, all eukaryota, chloroplasts, archaea, mitochondria, unknown sequences and singleton OTUs were excluded. For eukaryotes, unassigned and singleton OTUs were excluded. Finally, sequences data were normalized to 41,744 and 121,146 sequences per sample using the "sub.sample" command in MOTHUR v.1.33.3 (Schloss et al., 2009) for bacteria and eukaryotes, respectively.

All raw sequences from this study have been submitted to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database under the BioProject number PRJNA474064 and the accession number SRP149868.

Definition of Rare and Abundant Plankton

In order to evaluate the effects of DNA extraction kits on taxa in different relative abundances, we expanded the classification of microbial taxa based on the detected sequences and their relative abundance. This followed the definition of abundant (1%) and rare (0.01%) biosphere in previous studies (Pedrós-Alió, 2012; Liu et al., 2015; Xue et al., 2018). All OTUs were artificially defined and grouped into 6 exclusive categories following previous studies (Dai et al., 2016; Xue et al., 2018; Liu et al., 2019): (i) the OTUs with a relative abundance always $\geq 1\%$ in all replicates were regarded as always abundant taxa (AT); (ii) the OTUs with a relative abundance greater than 0.01% in all replicates and $\geq 1\%$ in some replicates but never rare ($< 0.01\%$) were regarded as conditionally abundant taxa (CAT); (iii) the OTUs with a relative abundance varying from rare ($< 0.01\%$) to abundant ($\geq 1\%$) were regarded as conditionally rare and abundant taxa (CRAT); iv) the OTUs with relative abundance between 0.01% and 1% in all

replicates were regarded as moderate taxa (MT); v) the OTUs with a relative abundance <0.01% in some replicates but never $\geq 1\%$ in all replicates were regarded as conditionally rare taxa (CRT); (vi) the OTUs with a relative abundance always <0.01% in all replicates were regarded as always rare taxa (RT).

Data Analyses

Rarefaction curves, and α -diversity indices were computed by MOTHUR v.1.33.3 (Schloss et al., 2009). The non-parametric Kruskal-Wallis test was used to evaluate the influences of DNA extraction kits on α -diversity indices and the quality and quantity of DNA by SPSS 22.0 (IBM Corp., Armonk, NY, United States).

Four dissimilarity matrices (Bray-Curtis, Jaccard, weighted unifracs and unweighted unifracs) within and between DNA extraction kits were calculated with the relative abundance-based OTUs of bacteria and eukaryotes. Analysis of similarities (ANOSIM) was used to estimate the significant differences among different DNA extraction kits. Complete separation is suggested by $R = 1$, with $R = 0$ representing no separation (Clarke and Gorley, 2015). We used the adonis function (vegan R packages) to run a PERMANOVA on the Bray-Curtis dissimilarity profiles using 10,000 permutations for assessing the effect of DNA extraction methods (Anderson, 2001). Similarity of percentages analysis (SIMPER) analysis was performed with PAST (Paleontological Statistics, version 3.01) software to identify the contribution of each OTU to the community dissimilarity (Hammer et al., 2001). The Bray-Curtis dissimilarity was used as the distance for ANOSIM, PERMANOVA and SIMPER analyses.

The reproducibility of each DNA extraction kit, evaluating the difference between plankton community compositions among triplicates within each DNA kit, was quantified by computing the average dissimilarity of each set of three replicates, using four dissimilarity matrices. The lower the dissimilarity measure, the more consistent that method was predicted to be. In addition, we randomly selected subsets of cleaned sequences (10,000, 20,000, 30,000, and 40,000 for bacteria; 10,000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000, and 100,000 for eukaryotes) from each replicate to estimate the influence of sequencing depth on the reproducibility. The one-way ANOVA was used to test the significant difference for reproducibility of DNA extraction kits.

A Venn diagram was constructed using the “Venn Diagram” package to compare the number of OTUs detected by using different DNA extraction kits. The “Niche breadth” approach of Levins (1968) was used to evaluate DNA extraction kits preference by the formula:

$$B_j = \frac{1}{\sum_{i=1}^N P_{ij}^2} \quad (1)$$

where B_j indicates the number of DNA extraction kits which the OTU occurred and P_{ij} represents the percentage of individuals belonging to species j present in a given DNA extraction kit i . Phylotypes characteristic of specific DNA extraction kit were identified using indicator species analysis (ISA). Those indicator OTUs with a P -value <0.05 and phylotypes with indicator values

> 50, were considered valid (Dufrene and Legendre, 1997). A heat map of the 20 most abundant OTUs and high-rank taxa was made using the “pheatmap” package in the R environment.

RESULTS

Effect of Extraction Kits on Plankton DNA Quality and Quantity

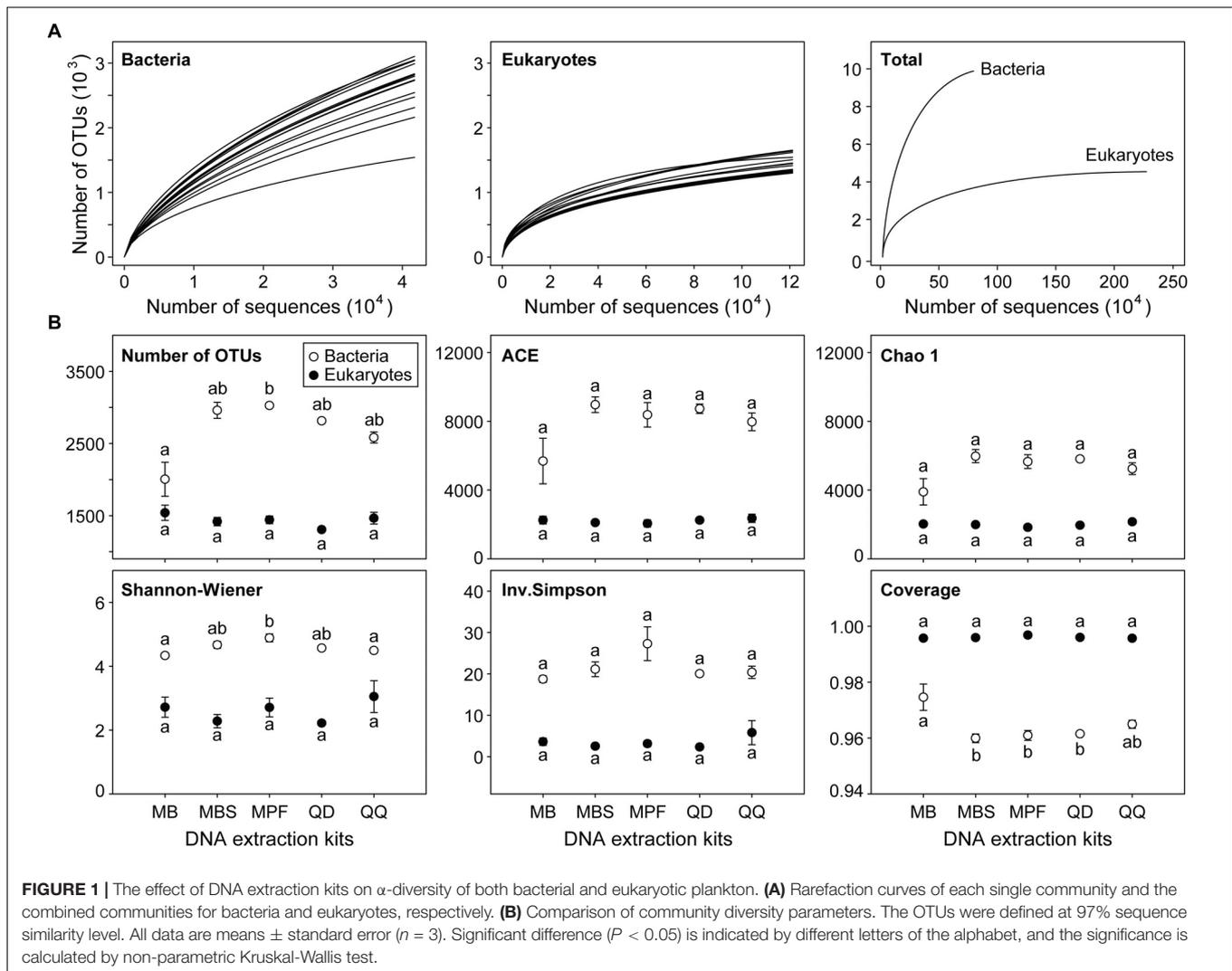
Although the DNA purity of the same filtered samples had no significant difference across different DNA extraction kits, the DNA concentrations varied greatly (Supplementary Figure S1). The DNA purity ranged from 1.90 to 2.15. The best results were obtained using PowerWater DNA Isolation Kit (MB) (mean, 1.97), PowerWater® Sterivex™ DNA Isolation Kit (MBS) (1.95) and Fast DNA Spin Kit for soil (MPF) (1.93). The highest DNA concentration was obtained using Fast DNA Spin Kit for soil (MPF) (mean \pm s.e., 64.80 \pm 1.12 ng/ μ L), and PowerWater DNA Isolation Kit (MB) (15.85 \pm 2.90 ng/ μ L) produced the lowest yield of DNA. Overall, MPF can maximize DNA concentration and purity. All negative extraction controls yielded DNA below the limit of detection.

Effect of Extraction Kits on Plankton Richness and α -Diversity

Most of the microbial plankton taxa had been recovered, as indicated by the species accumulation curves (Figure 1A), although no single replicate sample achieved a full saturation in the rarefaction curves. The sequencing depth for eukaryotic plankton (121,146) was higher than that of bacteria (41,744). Distinct influences of DNA extraction kits on α -diversity indices were observed for bacteria and eukaryotes (Figure 1B). For bacteria, only observed OTU number and Shannon-Wiener index had significant differences among DNA extraction kits ($P < 0.05$). MB showed the lowest OTU number (2005 \pm 235) and Shannon-Wiener (4.34 \pm 0.02), however the highest values of OTU number (3027 \pm 19) and Shannon-Wiener (4.89 \pm 0.12) were obtained from MPF. For eukaryotes, no significant difference was found for any α -diversity index among five different DNA extraction kits.

Reproducibility and Comparability of DNA Extraction Kits

To evaluate the effects of extraction kits on rare and abundant microbial taxa, we defined and grouped all OTUs into six exclusive categories based on their ranges of the relative abundance in 15 replicate samples (Supplementary Figure S2). Rare taxa had the highest OTUs number for both bacteria (86.62%) and eukaryotes (80.75%), whereas abundant taxa had the lowest OTUs number for both bacteria (0.04%) and eukaryotes (0%). For bacteria, conditionally abundant taxa showed the highest sequence number (42.65%), while the lowest sequence number was found in conditionally rare and abundant taxa (1.32%). For eukaryotes, conditionally abundant taxa also showed the highest sequence number (57.71%), while the lowest sequence number was found in abundant taxa (0%).



Community dissimilarity can be mainly due to the differences of DNA extraction methods (Figure 2A and Supplementary Figures S3A–S5A). The reproducibility of each DNA extraction kit showed significant difference ($P < 0.05$). QD and QQ exhibited the highest consistency for bacteria (0.62) and eukaryotes (0.76), respectively (Figure 2B). Different DNA extractions showed distinct patterns of consistency along relative abundance ranks (Figure 2 and Supplementary Figures S3–S5). Interestingly, RT, CRT and CRAT had lower consistency compared with AT, CAT and MT categories for both bacteria and eukaryotes (Figure 2B), indicating that DNA extractions have the larger effects on taxa with the lower relative abundance. In general, the higher the average between-replicate dissimilarity was, the larger the standard error of the dissimilarity was estimated (Figure 2 and Supplementary Figures S3–S5). Furthermore, we did not find that the sequencing depth showed any significant influence on the reproducibility of each DNA extraction kit for eukaryotes from 10,000 to 100,000 sequences (Supplementary Figure S6). For bacterial community, however, the reproducibility of both QD and QQ was significant higher

for 40,000 sequencing depth than 10,000 sequencing depth ($P < 0.05$), while sequencing depth exhibited no significant effect on the reproducibility of three other DNA extraction kits from 10,000 to 40,000 sequences (Supplementary Figure S6).

Effects of Extraction Kits on Community Composition

Four kinds of distance indices were considered to compare the community differences among the DNA extraction kits based on the variation of OTUs along relative abundance ranks, and in general they produced similar results (Figures 2C and Supplementary Figures S3C–S5C). For both bacteria and eukaryotes, most of the community differences were from rare taxa. By contrast, abundant, conditionally abundant, and moderate taxa showed higher similarity and higher consistency among different DNA extraction kits. In addition, two non-parametric multivariate statistical tests (ANOSIM and Adonis) showed that these observed differences were statistically significant (Table 2). Furthermore, the SIMPER analysis showed

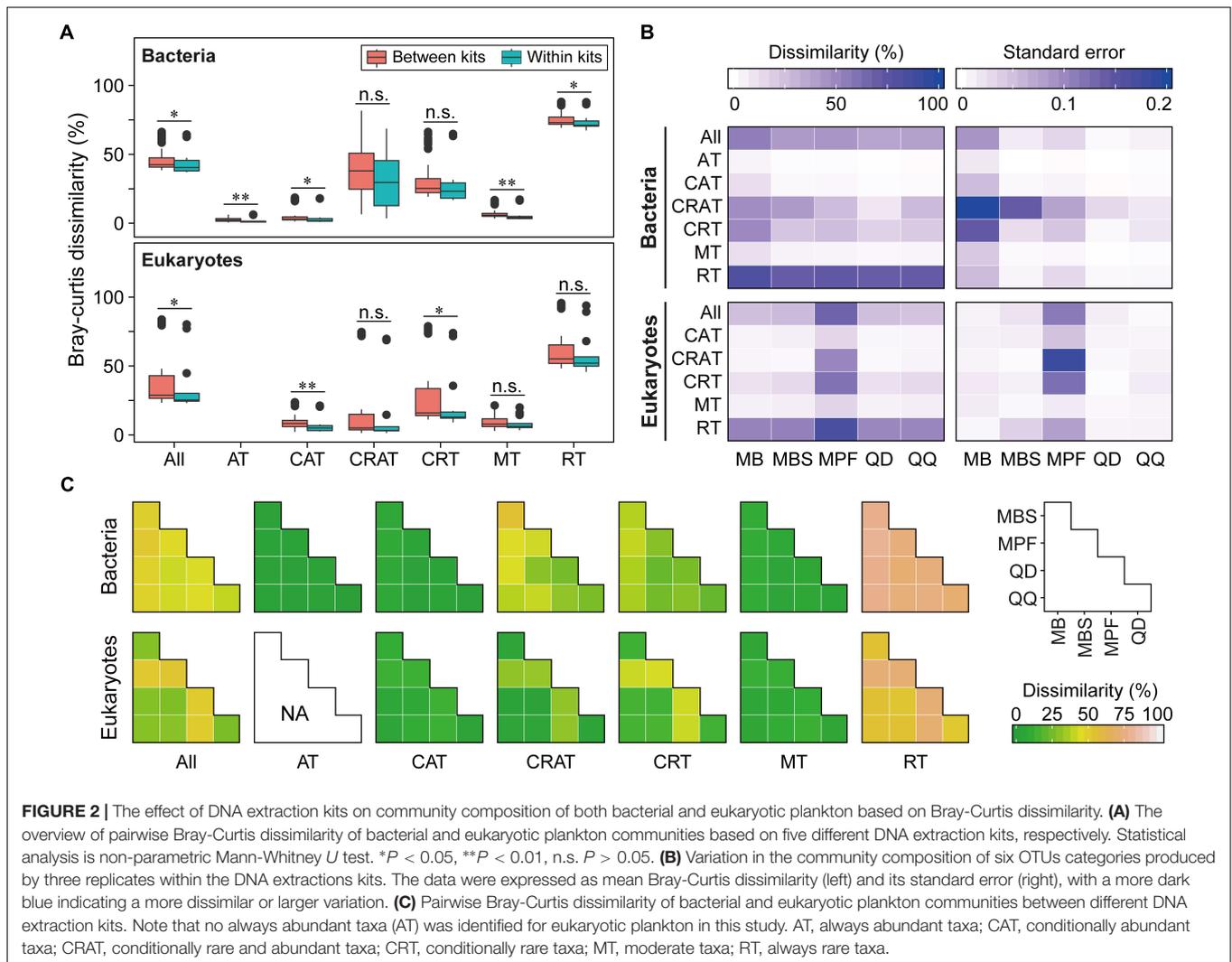
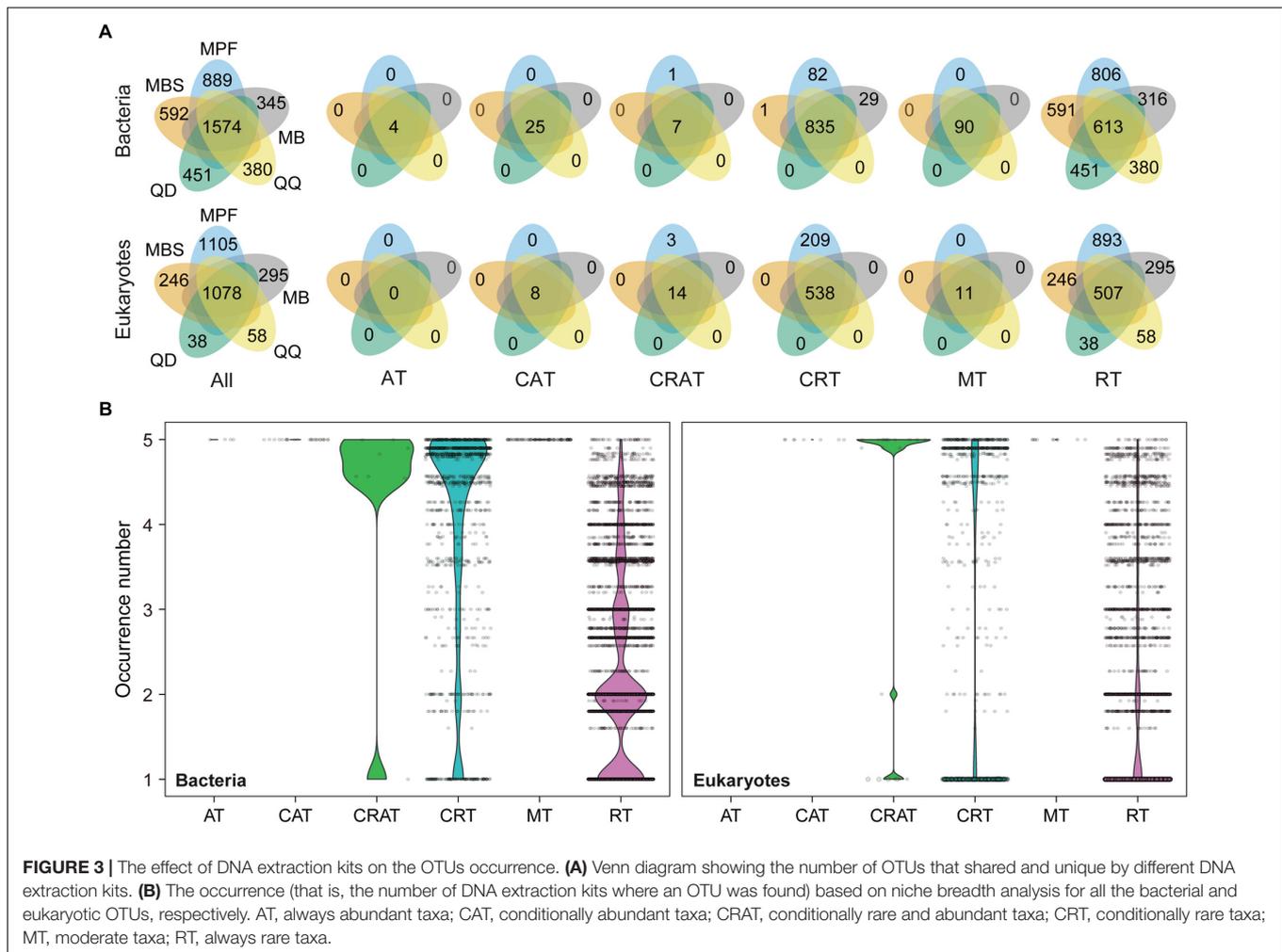


TABLE 2 | Analysis of similarity (ANOSIM) and permutational multivariate analysis of variance (PERMANOVA) for comparisons of microbial plankton communities among five different DNA extraction kits based on Bray-Curtis dissimilarity.

| Category | ANOSIM | | | | PERMANOVA | | | |
|----------|-------------|-----------------------|-------------|-----------------------|----------------------|-----------------------|----------------------|-----------------------|
| | Bacteria | | Eukaryotes | | Bacteria | | Eukaryotes | |
| | R | <i>P</i>-value | R | <i>P</i>-value | R² | <i>P</i>-value | R² | <i>P</i>-value |
| All | 0.34 | 0.004 | 0.32 | 0.001 | 0.33 | 0.006 | 0.37 | 0.002 |
| AT | 0.48 | 0.001 | – | – | 0.50 | 0.009 | – | – |
| CAT | 0.40 | 0.001 | 0.51 | 0.001 | 0.30 | 0.318 | 0.54 | 0.001 |
| CRAT | 0.26 | 0.008 | 0.32 | 0.002 | 0.40 | 0.073 | 0.37 | 0.002 |
| CRT | 0.28 | 0.005 | 0.33 | 0.003 | 0.33 | 0.150 | 0.38 | 0.005 |
| MT | 0.58 | 0.001 | 0.18 | 0.040 | 0.39 | 0.016 | 0.35 | 0.241 |
| RT | 0.35 | 0.001 | 0.26 | 0.002 | 0.31 | 0.003 | 0.34 | 0.002 |

Bold font indicates the significance at *P* < 0.05. Values show the *R* and *R*² values for ANOSIM and PERMANOVA, respectively. The operational taxonomic units (OTUs) were defined at 97% sequence similarity threshold. The ANOSIM statistic compares the mean of ranked dissimilarities among groups to the mean of ranked dissimilarities within groups. An *R* value close to “1” suggests dissimilarity between groups, while an *R* value near “0” suggests an even distribution of high and low ranks within and between groups. Negative *R* values indicate that dissimilarities are greater within groups than between groups. AT, always abundant taxa; CAT, conditionally abundant taxa; CRAT, conditionally rare and abundant taxa; CRT, conditionally rare taxa; MT, moderate taxa; RT, always rare taxa. Note that no always abundant taxa (AT) was identified for eukaryotic plankton in this study.



that RT and CRT accounted for the largest contribution for community differences for both bacteria (69.62% for RT, 27.86% for CRT) and eukaryotes (59.56% for RT, 38.24% for CRT) (**Supplementary Figure S7**).

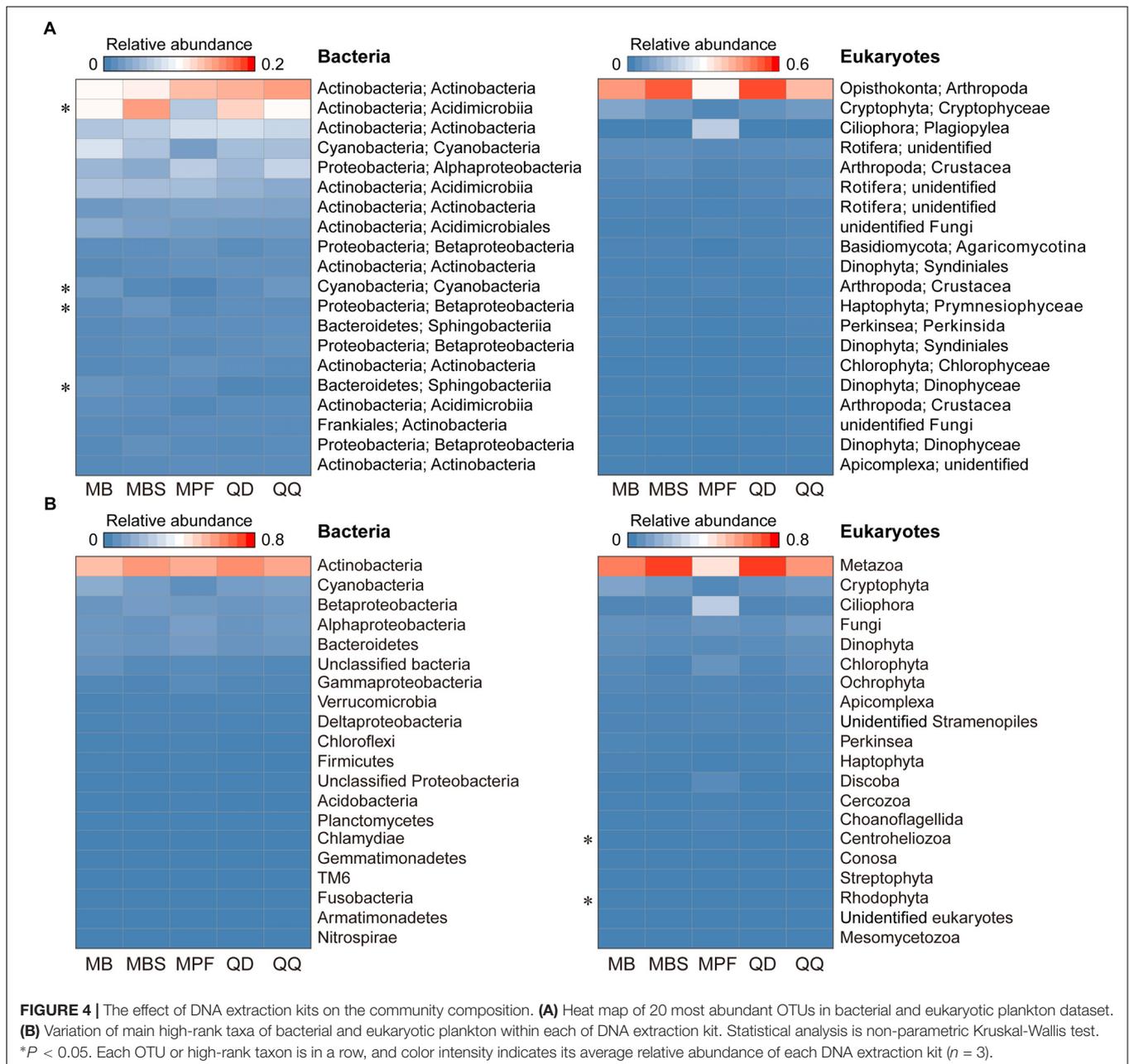
The influence of DNA extraction kits on community composition was identified at both OTU and higher taxonomic levels (**Figures 3, 4**). At OTU level, significant differences were found for specific taxa. First, regardless the microbial plankton types, most of the unique OTUs for each method belonged to the rare taxa (**Figure 3A**). In total, 15% bacterial OTUs and 23% eukaryotic OTUs were identified in at least four extraction kits, and less than 60% OTUs were found in no more than two methods for both bacteria and eukaryotes (**Figure 3B**). Second, indicator OTUs were identified for each DNA extraction kit (**Supplementary Table S2**). For bacteria, the numbers of indicator OTUs among kits were 1 for MB, 0 for MBS, 6 for MPF, 0 for QD, and 2 for QQ methods. For eukaryotes, the numbers of indicator OTUs were 4 for MB, 0 for MBS, 13 for MPE, 0 for QD, and 7 for QQ methods. At last, only 4 (*Actinobacteria*, *Cyanobacteria*, *Proteobacteria*, *Bacteroidetes*) of 20 most abundant OTUs were significantly different among DNA extraction kits for bacteria, whereas no difference was

identified for eukaryotic plankton (**Figure 4A**). At high-rank taxa level, no any significant difference was found for bacteria, while both *Rhodophyta* and *Centroheliozoa* were significantly different among five DNA extraction kits for eukaryotes (**Figure 4B**).

Finally, to improve data comparability, we removed singleton sequences and OTUs with less and equal to 5, 10, 50, 100, 500 sequences. We found that the larger number of sequences removed, the lower the community dissimilarity (**Figure 5A**) and the higher overlap of the OTUs (**Figure 5B**) for both bacteria and eukaryotes. This indicates that the low-abundance taxa had significant contribution to the community dissimilarity.

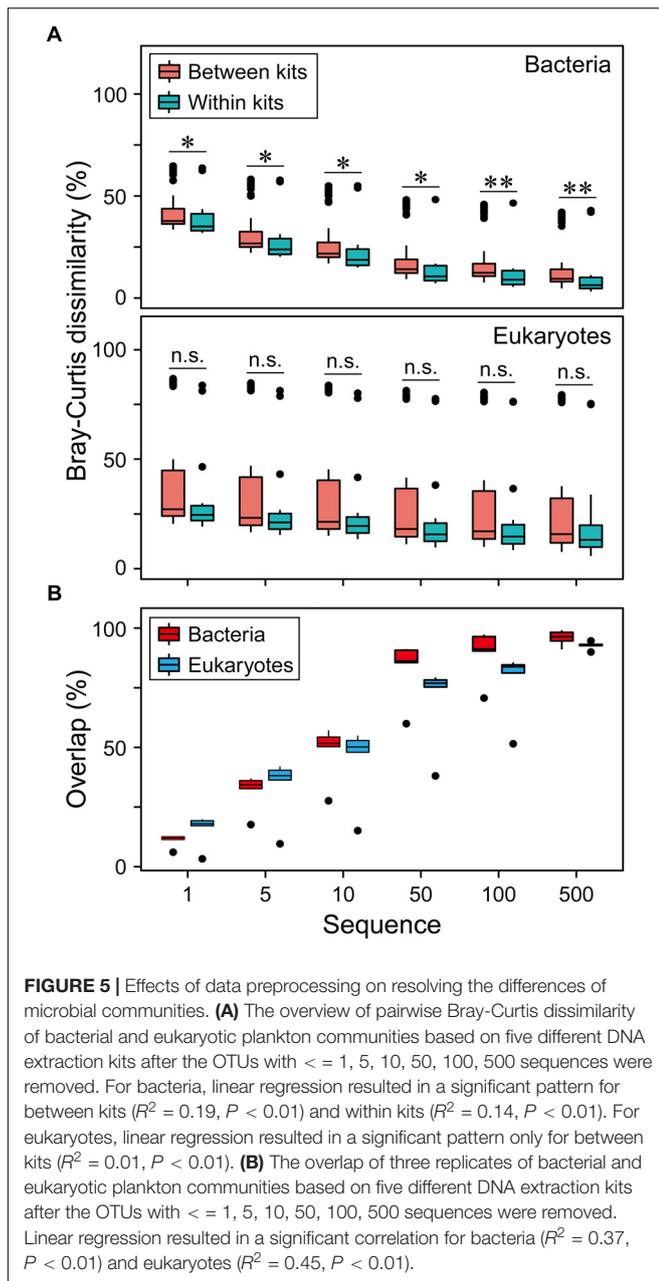
DISCUSSION

With the development of high-throughput technologies of target genes, researchers can now directly quantify the rare biosphere, which play very important ecological and biogeochemical roles in various ecosystems (Sogin et al., 2006; Pedrós-Alió, 2012; Liu et al., 2015, 2019; Lynch and Neufeld, 2015; Xue et al., 2018). However, great caution is needed when using DNA-based sequencing technologies for describing and



characterizing the diversity and composition of microbial community. The aim of this study is to comprehensively and systematically understand how low-abundance taxa were influenced by different DNA extraction kits. Here, we compared the variations of plankton communities from a single water sample using five of the most commonly used DNA extraction kits. We found that different DNA extraction kits had distinct influences on the alpha-diversity, community composition and specific taxa for both bacteria and eukaryotes. Specifically, low-abundance taxa were greatly affected by different DNA extraction kits. We also found that the community composition reproducibility varied among the DNA extraction kits.

The alpha-diversity of a microbial community could be influenced by DNA extraction kits to different degrees. Previous studies have found that the choice of DNA extraction methods could affect the alpha-diversity of bacterial communities from kick-net, leaf litter, soil and water (Purswani et al., 2011; Deiner et al., 2015; Hermans et al., 2018). However, no significant difference in alpha-diversity was obtained for eukaryotic (animal, fungal and plant) communities from soil and water samples (Hermans et al., 2018). Another study showed that no significant differences were found among DNA extraction kits when assessing diversity of eukaryotic microalgae from freshwater by using denaturant gradient gel electrophoresis (Eland et al., 2012). Compared with the previous studies mentioned above,



sequencing depths were higher in this study. Our results indicated that the alpha-diversity of bacteria was significantly affected by different DNA extraction kits, whereas there was no significant influence on the alpha-diversity of eukaryotic plankton community.

Additionally, significant differences were found for specific taxa at OTU or phylum level for both bacteria and eukaryotes. Previous studies have shown the preference in extraction specific taxa such as cyanobacteria, Excavata, Cercozoa and Amoebozoa (Singh et al., 2011; Santos et al., 2017). The differences in community composition detected may partly be due to the distinct susceptibility of bacterial and eukaryotic cell wall to different lysis methods or beads of different materials and sizes

(Bürgmann et al., 2001; Fredricks et al., 2005; Henderson et al., 2013; Velásquez-Mejía et al., 2018). Mechanical lysis is suitable for bacteria attributed to its harder lysing cells, whereas direct lysis of cells based on bead-beating protocols is considered to be more suited to the detection of microeukaryotes, as these cells often have fortified cell walls or very resilient cell membranes (Santos et al., 2017; Hermans et al., 2018). Otherwise, even for bacteria, methods without a bead beating or enzymatic treatment step generally extracted less DNA from Gram-positive bacteria (Knudsen et al., 2016). Other factors such as sample type or origin (Knudsen et al., 2016), inherent specimen properties (Knudsen et al., 2016; Sinha et al., 2017), primer choice (Albertsen et al., 2015) and bioinformatics protocol choices (Sinha et al., 2017) can also influence the outcome of microbial community analysis. It is difficult to find a single extraction method for all circumstances. Our results highlighted that it is important to design specific DNA extraction method to target particular taxonomic groups.

Previous studies have found the influence of DNA extraction kits on the whole or abundant taxa in microbial communities (Knudsen et al., 2016; Sinha et al., 2017; Hermans et al., 2018), however, the influence of extraction methods on low-abundance taxa were not comprehensively and systematically evaluated. Recently, Weiss et al. (2014) found that DNA extraction kits had significant influence on the resulting microbial community profile of low-biomass samples. Our results indicated that low-abundance taxa were largely influenced by the DNA extraction kits, and offered further evidence supporting that plankton taxa in different relative abundances were disproportionately influenced. One explanation for the results is that as the number of “true” taxa becomes less, the potential for contaminants occupying a larger fraction of the sequences will become greater (Weiss et al., 2014). Another explanation for this is that some of the community differences between DNA extraction kits may be actually “real” chance differences between replicate samples. It’s possible that some replicate samples may not contain any individuals of a very rare taxon, despite the fact that we mixed the water before filtering. In addition, the differences in microbial communities may also come from the amplicon sequencing-based detection method, as its influence on analyzing microbial community composition (Zhou et al., 2011). To be conservative, it is ideal to adopt a single DNA extraction method for increasing the comparability of different studies. Our results indicate that we can improve the community comparisons between studies that employed different extraction methods with an exception of RT and CRT for their large contributions to community differences among different DNA extraction kits.

Consistency of generating OTU profiles is highly desirable in order to facilitate assessments and comparisons of plankton species diversity and distribution across space and time. Reproducibility is a crucial issue for the study of microbial community ecology. Sample processing steps, from sampling to downstream data analysis, can introduce some biases; such biases can skew data sets by introducing changes in the relative abundances observed, and they can affect the variation among replicates (Zhou et al., 2011; Costea et al., 2017; Pollock et al., 2018; Velásquez-Mejía et al., 2018). The differences in the replicates for each DNA extraction kit are likely the result of

DNA extraction and sequencing depth, because other sample process steps should influence all our replicate samples in a similar way (Velásquez-Mejía et al., 2018). Similarly, Zhou et al. (2011) found a reliable diversity comparison across different samples by removing less-representative OTUs for bacteria (e.g., OTUs present only in 1 or 2 of the 14 tag sequence data sets). Our result supported this finding and further demonstrated that the taxa in different relative abundance showed distinct patterns of reproducibility. This could be explained by the high OTUs overlap between replicate samples when low-abundance OTUs were removed from the dataset (Figure 5). In addition, sequencing depth should also be considered, because the percentage of shared OTUs increased as the sequencing depth increase (Liu et al., 2017). However, sequencing depth only exhibited a significant and minor influence on the results of QD and QQ for bacteria, with no significant influence identified for other DNA extraction methods (Supplementary Figure S6). There are also some limitations in this research that need for further study. For example, the technical replicates should be increased to better estimate the background noise level (Zhou et al., 2011). Sample handling environment, and bioinformatics should also be considered for their important influence on the community composition (Costea et al., 2017; Sinha et al., 2017).

CONCLUSION

In conclusion, DNA extraction methods had an influence on the results of downstream microbial community analyses, including relative abundances of specific community members for both bacteria and eukaryotes at OTU and higher taxonomic levels. The rare plankton subcommunities are far more affected by DNA extraction kits than the abundant plankton. Every extraction kit was effective, though each showed its own strengths and weaknesses in observing special taxonomic community profiles. MPF produced higher α diversity for bacteria and high quality and yield of DNA, therefore it is the most suitable DNA methods for bacterial plankton diversity study. While QD and QQ methods could produce results that are easier to replicate for bacteria and eukaryotes, respectively. Moreover, our results highlight that the comparability between studies that employed different extraction methods can be improved by removing low-abundance OTUs for their larger contribution to community variation. There is no doubt that the efforts to assess bias within labs as proposed here, and guidelines for best practices

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. doi: 10.1371/journal.pone.0132783
- Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., and Huse, S. M. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* 4:e6372. doi: 10.1371/journal.pone.0006372
- Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* 26, 32–46. doi: 10.1111/j.1442-9993.2001.01070.pp.x

across labs, will facilitate comparative studies to characterize microbial communities from multiple sampling surveys and various environments.

DATA AVAILABILITY

The datasets generated for this study can be found in National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database, BioProject number PRJNA474064 and the accession number SRP149868.

AUTHOR CONTRIBUTIONS

JY designed the research. ML and YX performed the experiments. ML and JY analyzed the data and wrote the manuscript. All authors have significantly contributed to the writing of the manuscript.

FUNDING

This work was funded by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA23040302), the National Natural Science Foundation of China (91851104 and 31672312), and the Xiamen Municipal Bureau of Science and Technology (3502Z20172024).

ACKNOWLEDGMENTS

We are very grateful to Prof. David M. Wilkinson for his valuable comments on an earlier version of the manuscript. We thank our two reviewers for insightful comments who helped in improving the clarity of this paper. The research was also supported by the Wuhan Branch, Supercomputing Center, Chinese Academy of Sciences, China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.00454/full#supplementary-material>

- Bürgmann, H., Pesaro, M., Widmer, F., and Zeyer, J. (2001). A strategy for optimizing quality and quantity of DNA extracted from soil. *J. Microbiol. Methods* 45, 7–20. doi: 10.1016/S0167-7012(01)00213-5
- Cabeen, M. T., and Jacobs-Wagner, C. (2005). Bacterial cell shape. *Nat. Rev. Microbiol.* 3, 601–610. doi: 10.1038/nrmicro1205
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303
- Clarke, K. R., and Gorley, R. N. (2015). *PRIMER v7: User Manual/Tutorial*. Plymouth: PRIMER-E Ltd.

- Costea, P. I., Zeller, G., Sunagawa, S., Pelletier, E., Alberti, A., Levenez, F., et al. (2017). Towards standards for human fecal sample processing in metagenomic studies. *Nat. Biotechnol.* 35, 1069–1076. doi: 10.1038/nbt.3960
- Dai, T., Zhang, Y., Tang, Y., Bai, Y., Tao, Y., Huang, B., et al. (2016). Identifying the key taxonomic categories that characterize microbial community diversity using full-scale classification: a case study of microbial communities in the sediments of Hangzhou Bay. *FEMS Microbiol. Ecol.* 92:fw150. doi: 10.1093/femsec/fiw150
- Deiner, K., Walser, J. C., Mächler, E., and Altermatt, F. (2015). Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biol. Conserv.* 183, 53–63. doi: 10.1016/j.biocon.2014.11.018
- Dufrene, M., and Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol. Monogr.* 67, 345–366. doi: 10.2307/2963459
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381
- Eichmiller, J. J., Miller, L. M., and Sorensen, P. W. (2016). Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. *Mol. Ecol. Resour.* 16, 56–68. doi: 10.1111/1755-0998.12421
- Eland, L. E., Davenport, R., and Mota, C. R. (2012). Evaluation of DNA extraction methods for freshwater eukaryotic microalgae. *Water Res.* 46, 5355–5364. doi: 10.1016/j.watres.2012.07.023
- Fredricks, D. N., Smith, C., and Meier, A. (2005). Comparison of six DNA extraction methods for recovery of fungal DNA as assessed by quantitative PCR. *J. Clin. Microbiol.* 43, 5122–5128. doi: 10.1128/JCM.43.10.5122-5128.2005
- Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597–D604. doi: 10.1093/nar/gks1160
- Guo, Y. Y., Liu, M., Liu, L. M., Liu, X., Chen, H. H., and Yang, J. (2018). The antibiotic resistome of free-living and particle-attached bacteria under a reservoir cyanobacterial bloom. *Environ. Int.* 117, 107–115. doi: 10.1016/j.envint.2018.04.045
- Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 4, 1–9.
- Henderson, G., Cox, F., Kittelmann, S., Vahideh, H. M., Zethof, M., Noel, S. J., et al. (2013). Effect of DNA extraction methods and sampling techniques on the apparent structure of cow and sheep rumen microbial communities. *PLoS One* 8:e74787. doi: 10.1371/journal.pone.0074787
- Hermans, S. M., Buckley, H. L., and Lear, G. (2018). Optimal extraction methods for the simultaneous analysis of DNA from diverse organisms and sample types. *Mol. Ecol. Resour.* 18, 557–569. doi: 10.1111/1755-0998.12762
- Knight, R., Vrbanc, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., et al. (2018). Best practices for analyzing microbiomes. *Nat. Rev. Microbiol.* 16, 410–422. doi: 10.1038/s41579-018-0029-9
- Knudsen, B. E., Lasse, B., Munk, P., Oksana, L., Priemé, A., Aarestrup, F. M., et al. (2016). Impact of sample type and DNA isolation procedure on genomic inference of microbiome composition. *mSystems* 1:e00095-16. doi: 10.1128/mSystems.00095-16
- Kuhn, R., Böllmann, J., Krahl, K., Bryant, I. M., and Martienssen, M. (2017). Comparison of ten different DNA extraction procedures with respect to their suitability for environmental samples. *J. Microbiol. Methods* 143, 78–86. doi: 10.1016/j.mimet.2017.10.007
- Lear, G., Dickie, I., Banks, J., Boyer, S., Buckley, H. L., Buckley, T. R., et al. (2018). Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. *New Z. J. Ecol.* 42, 1A–50A. doi: 10.20417/nzjecol.42.9
- Levins, R. (1968). *Evolution in Changing Environments*. Princeton, NJ: Princeton University Press.
- Li, P., Yang, S. F., Lv, B. B., Zhao, K., Lin, M. F., Zhou, S., et al. (2015). Comparison of extraction methods of total microbial DNA from freshwater. *Genet. Mol. Res.* 14, 730–738. doi: 10.4238/2015.January.30.16
- Liu, L. M., Liu, M., Wilkinson, D. M., Chen, H. H., Yu, X. Q., and Yang, J. (2017). DNA metabarcoding reveals that 200- μ m-size-fractionated filtering is unable to discriminate between planktonic microbial and large eukaryotes. *Mol. Ecol. Resour.* 17, 991–1002. doi: 10.1111/1755-0998.12652
- Liu, L. M., Yang, J., Yu, Z., and Wilkinson, D. M. (2015). The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of china. *ISME J.* 9, 2068–2077. doi: 10.1038/ismej.2015.29
- Liu, M., Liu, L. M., Chen, H. H., Yu, Z., Yang, J. R., Xue, Y. Y., et al. (2019). Community dynamics of free-living and particle-attached bacteria following a reservoir *Microcystis* bloom. *Sci. Total Environ.* 660, 501–511. doi: 10.1016/j.scitotenv.2018.12.414
- Lynch, M. D., and Neufeld, J. D. (2015). Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* 13, 217–229. doi: 10.1038/nrmicro3400
- Magoč, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507
- McCarthy, A., Chiang, E., Schmidt, M. L., and Deneff, V. J. (2015). RNA preservation agents and nucleic acid extraction method bias perceived bacterial community composition. *PLoS One* 10:e0121659. doi: 10.1371/journal.pone.0121659
- Pedros-Alíó, C. (2012). The rare bacterial biosphere. *Annu. Rev. Mar. Sci.* 4, 449–466. doi: 10.1146/annurev-marine-120710-100948
- Pollock, J., Glendinning, L., Wisedchanwet, T., and Watson, M. (2018). The madness of microbiome: attempting to find consensus “best practice” for 16S microbiome studies. *Appl. Environ. Microb.* 84, e2627–e2617. doi: 10.1128/AEM.02627-17
- Purswani, J., Martín-Platero, A. M., Reboleiro-Rivas, P., González-López, J., and Pozo, C. (2011). Comparative analysis of microbial DNA extraction protocols for groundwater samples. *Anal. Biochem.* 416, 240–242. doi: 10.1016/j.ab.2011.05.024
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219
- Renshaw, M. A., Olds, B. P., Jerde, C. L., McVeigh, M. M., and Lodge, D. M. (2015). The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol-chloroform-isoamyl alcohol DNA extraction. *Mol. Ecol. Resour.* 15, 168–176. doi: 10.1111/1755-0998.12281
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., et al. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 12:87. doi: 10.1186/s12915-014-0087-z
- 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. *J. Microbiol. Methods* 115, 13–19. doi: 10.1016/j.mimet.2015.05.011
- Santos, S. S., Nunes, I., Nielsen, T. K., Jacquiod, S., Hansen, L. H., and Winding, A. (2017). Soil DNA extraction procedure influences protist 18S rRNA gene community profiling outcome. *Protist* 168, 283–293. doi: 10.1016/j.protis.2017.03.002
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microb.* 75, 7537–7541. doi: 10.1128/AEM.01541-09
- Shendure, J., Balasubramanian, S., Church, G. M., Gilbert, W., Rogers, J., Schloss, J. A., et al. (2017). DNA sequencing at 40: past, present and future. *Nature* 550, 345–353. doi: 10.1038/nature24286
- Singh, S. P., Rastogi, R. P., Häder, D. P., and Sinha, R. P. (2011). An improved method for genomic DNA extraction from cyanobacteria. *World J. Microb. Biotechnol.* 27, 1225–1230. doi: 10.1007/s1274-010-0571-8
- Sinha, R., Abu-Ali, G., Vogtmann, E., Fodor, A. A., Ren, B., Amir, A., et al. (2017). Assessment of variation in microbial community amplicon sequencing by the microbiome quality control (MBQC) project consortium. *Nat. Biotechnol.* 35, 1077–1086. doi: 10.1038/nbt.3981
- Sogin, M. L., Morrison, H. G., Huber, J. A., Mark Welch, D., Huse, S. M., Neal, P. R., et al. (2006). Microbial diversity in the deep sea and the under explored “rare biosphere”. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12115–12120. doi: 10.1073/pnas.0605127103
- Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., et al. (2015). Structure and function of the global ocean microbiome. *Science* 348:1261359. doi: 10.1126/science.1261359
- Sundberg, C., Al-Soud, W. A., Larsson, M., Alm, E., Yekta, S. S., Svensson, B. H., et al. (2013). 454 pyrosequencing analyses of bacterial and archaeal richness in

- 21 full-scale biogas digesters. *FEMS Microbiol. Ecol.* 85, 612–626. doi: 10.1111/1574-6941.12148
- Velásquez-Mejía, E. P., de la Cuesta-Zuluaga, J., and Escobar, J. S. (2018). Impact of DNA extraction, sample dilution, and reagent contamination on 16S rRNA gene sequencing of human feces. *Appl. Microbiol. Biotechnol.* 102, 403–411. doi: 10.1007/s00253-017-8583-z
- Walden, C., Carbonero, F., and Zhang, W. (2017). Assessing impacts of DNA extraction methods on next generation sequencing of water and wastewater samples. *J. Microbiol. Methods* 141, 10–16. doi: 10.1016/j.mimet.2017.07.007
- Weiss, S., Amir, A., Hyde, E. R., Metcalf, J. L., Song, S. J., and Knight, R. (2014). Tracking down the sources of experimental contamination in microbiome studies. *Genome Biol.* 15:564. doi: 10.1186/s13059-014-0564-2
- Xue, Y. Y., Chen, H. H., Yang, J. R., Liu, M., Huang, B. Q., and Yang, J. (2018). Distinct patterns and processes of abundant and rare eukaryotic plankton communities following a reservoir cyanobacterial bloom. *ISME J.* 12, 2263–2277. doi: 10.1038/s41396-018-0159-0
- Yu, Y., Lee, C., Kim, J., and Hwang, S. (2005). Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol. Bioeng.* 89, 670–678. doi: 10.1002/bit.20347
- Zhou, J., Wu, L., Deng, Y., Zhi, X., Jiang, Y., Tu, Q., et al. (2011). Reproducibility and quantitation of amplicon sequencing-based detection. *ISME J.* 5, 1303–1313. doi: 10.1038/ismej.2011.11

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Liu, Xue and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.