

# Assessing high-throughput environmental DNA extraction methods for meta-barcode characterization of aquatic microbial communities

Abdolrazagh Hashemi Shahraki, Subba Rao Chaganti and Daniel Heath

## ABSTRACT

The characterization of microbial community dynamics using genomic methods is rapidly expanding, impacting many fields including medical, ecological, and environmental research and applications. One of the biggest challenges for such studies is the isolation of environmental DNA (eDNA) from a variety of samples, diverse microbes, and widely variable community compositions. The current study developed environmentally friendly, user safe, economical, and high throughput eDNA extraction methods for mixed aquatic microbial communities and tested them using 16 s rRNA gene meta-barcoding. Five different lysis buffers including (1) cetyltrimethylammonium bromide (CTAB), (2) digestion buffer (DB), (3) guanidinium isothiocyanate (GITC), (4) sucrose lysis (SL), and (5) SL-CTAB, coupled with four different purification methods: (1) phenol-chloroform-isoamyl alcohol (PCI), (2) magnetic Bead-Robotic, (3) magnetic Bead-Manual, and (4) membrane-filtration were tested for their efficacy in extracting eDNA from recreational freshwater samples. Results indicated that the CTAB-PCI and SL-Bead-Robotic methods yielded the highest genomic eDNA concentrations and succeeded in detecting the core microbial community including the rare microbes. However, our study recommends the SL-Bead-Robotic eDNA extraction protocol because this method is safe, environmentally friendly, rapid, high-throughput and inexpensive.

**Key words** | health, lysis buffer, meta-barcode, microbial community, pathogens, water

**Abdolrazagh Hashemi Shahraki**  
**Subba Rao Chaganti**  
**Daniel Heath** (corresponding author)  
Great Lakes Institute for Environmental Research,  
University of Windsor,  
Windsor, Ontario N9B 3P4,  
Canada  
E-mail: [dheath@uwindsor.ca](mailto:dheath@uwindsor.ca)

**Daniel Heath**  
Department of Biological Sciences,  
University of Windsor,  
Windsor, Ontario,  
Canada

## INTRODUCTION

Microbial community profiling using massively parallel sequencing of the 16S rRNA gene ('meta-barcoding') is widely utilized due to recent advances in high-throughput amplicon sequencing technologies that facilitate large-scale, culture-independent studies of environmental microbiota (Tringe & Hugenholtz 2008). More generally, microbiology is in the process of adopting high-throughput DNA sequencing methods for the detection of pathogens (Tagliani *et al.* 2017) and drug-resistant genes (Biswas *et al.* 2017) from clinical settings to environmental settings (Mohiuddin *et al.* 2017) and the study of microbial community dynamics in diverse ecosystems ranging from soil (Ramirez *et al.* 2018) to aquatic

environments (Currie *et al.* 2017). Meta-barcoding using environmental DNA (eDNA; mixed DNA obtained directly from environmental samples) is a new and promising approach to studying microbial communities in diverse ecosystems (Taberlet *et al.* 2012). The aquatic environment contains a diverse group of microbes with various cell wall compositions (Wang *et al.* 2012) and, consequently, different sensitivity to lysis solutions. Due to the expected low concentrations of bacterial pathogens, fecal indicators, and microbial source tracking markers in fresh water (Cabral 2010; Cloutier & McLellan 2017), optimally effective lysis buffer and purification methods must be developed and

doi: 10.2166/wh.2018.108

tested to guarantee effective and unbiased genomic eDNA extraction from both the common and rare species of the microbial community. On the other hand, a few recent studies have highlighted the selection of lysis buffer and purification methods as sources of variability in the characterization of microbial community composition (Yuan *et al.* 2012; Kennedy *et al.* 2014). Although phenol-chloroform-isoamyl alcohol (PCI) extraction is widely used as a DNA extraction method (Chaganti *et al.* 2012), this method is not safe for the laboratory worker, nor environmentally friendly, and is highly time-consuming for large sample sizes. The objectives of this study were to (i) evaluate different lysis buffer and eDNA purification protocol combinations for efficiency in eDNA extraction from aquatic samples and (ii) evaluate the best lysis buffer and extraction protocol combinations for bias in microbial community composition. The completion of those two objectives allows us to identify and recommend the most efficient, safe, cost-effective, and unbiased eDNA extraction method for aquatic microbial community meta-barcoding characterization, including rare and possibly pathogenic microbial taxa.

## MATERIALS AND METHODS

### Sampling site and sample processing

Water samples (15 L) were collected and delivered to the laboratory on 3 different days (11, 18, and 25 August 2016) from Holiday Beach, Windsor, Ontario, Canada. Each 15 L water sample was separated into 60 aliquots of 250 mL. Each aliquot was filtered through 0.22-micron pore size, 47 mm diameter polycarbonate filters (Isopore™, Millipore, MA). After filtration, each filter was cut in half, and each half was placed in a 2 mL sterile tube containing 50 g glass beads (0.1 mm diameter, Bio-Spec Products, Bartlesville, USA), for further lysis and extraction tests. One half was used for immediate DNA extraction (the other half kept for other applications). The resulting 60 samples (half-filters) from each sample date were divided into 20 groups of triplicates ( $20 \times 3 = 60$  samples). Total nucleic acids were extracted from the filters (all sample dates and replicates) using combinations of mechanical, chemical, and enzymatic cell lysis as shown in Figure 1. In this study, all combinations

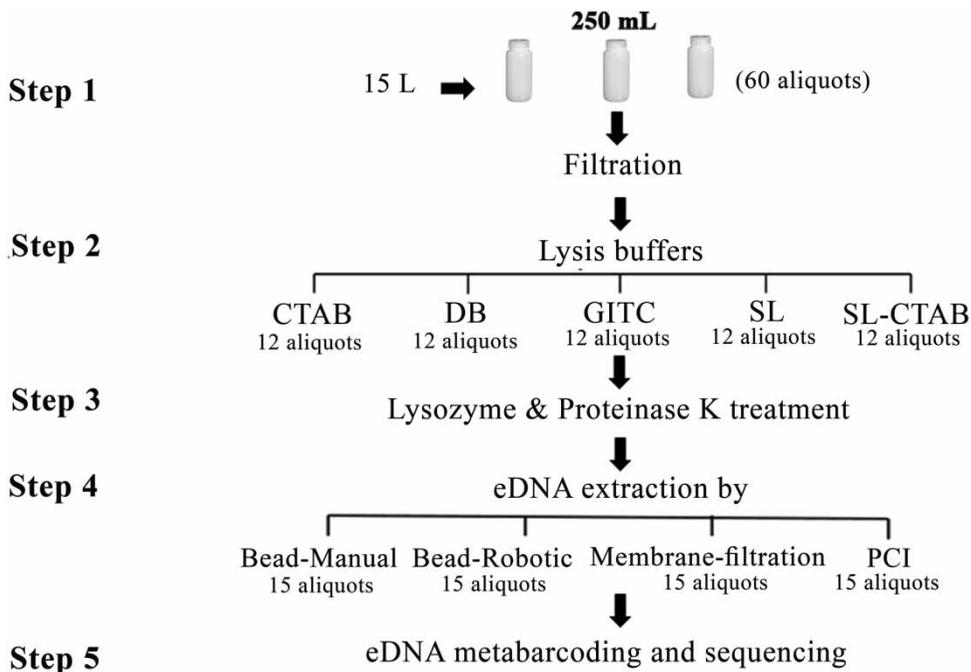
of five different lysis buffers and four different purification methods were evaluated on the three samples collected over three weeks in August 2016. eDNA purification success was evaluated by measuring DNA concentration. The DNAs that resulted from extraction protocols that yielded high quality and quantity of DNA were further evaluated by meta-barcoding the V5–V6 region of 16S rRNA fragment using massively parallel (Next Generation) sequencing or ‘NGS’.

### Cell lysis buffers

Five different lysis buffers, cetyltrimethylammonium bromide (CTAB), digestion buffer (DB) or sodium dodecyl sulfate (SDS), guanidinium isothiocyanate (GITC), sucrose (SL), and sucrose-CTAB (SL-CTAB) were used to lyse the filtered samples (Figure 1). CTAB, GITC, and SDS are chemical detergents which frequently are used to lyse cells (Nishiguchi *et al.* 2002). However, some of them work better for some organisms/cells such as GITC which more effectively lysis mycobacteria cells (Kotlowski *et al.* 2004). By using five different lysis buffers, we tried to test and select the best ones for eDNA extraction from freshwater samples.

- (1) CTAB lysis buffer: Cell lysis was carried out by adding 400  $\mu\text{L}$  of 2% CTAB; 20 g CTAB  $\text{L}^{-1}$ , 81 g NaCl  $\text{L}^{-1}$ , 0.5 M EDTA pH 7.5, 1 M Tris-HCl pH: 8 to each tube (Chaganti *et al.* 2012).
- (2) DB lysis buffer: Cell lysis was carried out by adding 400  $\mu\text{L}$  of DB buffer (1 M NaCl, 1 M Tris-HCl pH: 8, 0.5 EDTA, 10% SDS) to each tube.
- (3) GITC lysis buffer: Cell lysis was carried out by adding 400  $\mu\text{L}$  of GITC solution (4 M GITC pH: 7.5, 50 g Sarcosyl  $\text{L}^{-1}$ , 5 mM EDTA pH: 8, 5 g sodium citrate  $\text{L}^{-1}$ , and 5 g Triton X-100  $\text{L}^{-1}$ ) (Kotlowski *et al.* 2004) to each tube.
- (4) SL buffer: Cell lysis was accomplished by adding 400  $\mu\text{L}$  of SL buffer (400 mM NaCl, 750 mM sucrose, 20 mM ethylenediaminetetraacetic acid (EDTA), 50 mM Tris-HCl pH 9.0) (Boström *et al.* 2004) to each tube.
- (5) SL-CTAB lysis buffer: Cell lysis was carried out by adding 200  $\mu\text{L}$  of SL buffer and 200  $\mu\text{L}$  of 2% CTAB to each tube.

After adding each lysis buffer to all aliquots (Figure 1), they were subjected to bead-beating by a Mini-beadbeater-16 (Lab Services BV, The Netherlands) for 1 min (three times) with intensity of 3,450 oscillations/ $\text{min}^{-1}$ . Then, each



**Figure 1** | Schematic diagram showing the study protocol designed to evaluate efficacy of eDNA extraction from freshwater samples for microbial community characterization. This protocol used water samples taken on three different days. In Step 1, 60 250 mL aliquots of sampled water taken on a single date were filtered, grouped into 20 groups with three sample replicates. The samples (filters) were treated with five different lysis buffers (Step 2) and eDNA was extracted with four different extraction methods (Step 4). In Step 5, selected combinations of lysis buffers and extraction methods (CTAB-PCI, DB-Bead-Robotic, and SL-Bead-Robotic) with high eDNA yield were subjected to 16S rRNA PCR and bacterial community meta-barcoding.

sample was treated with 50  $\mu\text{L}$  lysozyme (Sigma-Aldrich, USA); 10  $\text{mg}/\text{mL}^{-1}$  for 1 h (37  $^{\circ}\text{C}$ ) and then 2  $\mu\text{L}$  proteinase K (Thermo Scientific, USA); 20  $\text{mg}/\text{mL}^{-1}$  for 1 h (50  $^{\circ}\text{C}$ ). eDNA was extracted from the lysate from all lysis buffer treatments by the following purification methods.

### Environmental DNA purification

We selected three DNA purification methods commonly used for eDNA purification: (1) phase separation and precipitation (PCI), (2) magnetic bead capture, and (3) membrane-filtration capture. We developed specific protocols for each approach and included an automated liquid handling platform as a high-throughput option for the magnetic bead capture approach (Figure 1). Below, we describe each purification protocol.

#### PCI method

Environmental DNA from one set of the treated samples from each lysis buffer (Figure 1) was purified manually

with PCI as described previously (Chaganti *et al.* 2012). Extracted eDNA was eluted in 50  $\mu\text{L}$  of 10 mM Tris-EDTA (TE) buffer (pH 7.4) and stored at  $-80^{\circ}\text{C}$  until use.

#### Membrane-filtration method

We used AcroPrep™ 96-well filter plates (Pall Corporation, USA) with a vacuum manifold system (MAVM0960R; Millipore) to purify eDNA from 150  $\mu\text{L}$  of the lysate from all lysis buffers (Elphinstone *et al.* 2003). Extracted eDNA was eluted in 50  $\mu\text{L}$  of TE buffer and kept at  $-80^{\circ}\text{C}$  until use.

#### Magnetic bead method (Bead-Robotic)

Bead-Robotic purification was carried out using the solid phase reversible immobilization (SPRI) paramagnetic bead-based method on an automated liquid handling platform (Tecan Freedom Evo150 Liquid Handling Platform, Perkin Elmer, USA; robot protocol script provided in Supplementary material, File S1, available with the online version of this paper) (Vo & Jedlicka 2014). To extract eDNA, 150  $\mu\text{L}$

of the lysate from all lysis buffers was mixed with 225  $\mu\text{L}$  of the SPRI bead solution and incubated at room temperature for 5 min to bind the DNA. The beads were separated via magnetic plate for 5 min and the cleared solution was removed. The beads were washed twice with 70% ethanol and air dried for 5 min. TE buffer (50  $\mu\text{L}$ ) was used as an elution buffer and extracted eDNA was transferred to a fresh tube and kept at  $-80^\circ\text{C}$  until use.

### Magnetic bead method (Bead-Manual)

Environmental DNA was purified manually from the same lysate of treated samples (Figure 1) as for the other extraction trials. We manually mixed 150  $\mu\text{L}$  of the lysate aliquot with 225  $\mu\text{L}$  of the SPRI-bead solution, and the mixture was incubated at room temperature for 5 min. The same steps were taken to purify eDNA as in the robotic protocol, but all steps were carried out manually. TE buffer (50  $\mu\text{L}$ ) was used as an elution buffer and extracted eDNA kept at  $-80^\circ\text{C}$  until use. The extracted eDNA (5  $\mu\text{L}$ ) from each extraction was loaded on the agarose gel for visualization of crude DNA extracts. A Nanodrop ND-1000 UV spectrophotometer (Nanodrop, DE) was used to measure eDNA concentration and purity. Absorbance at 260 nm was used to measure the concentration of nucleic acids. DNA purity was assessed via 260/280 and 260/230 absorbance ratios (Lucena-Aguilar et al. 2016), as determined via spectrophotometry.

### PCR amplification, sequencing, and data processing

Three combinations of lysis buffer and purification methods (CTAB-PCI, SL-Bead-Robotic, and DB-Bead-Robotic) were selected as the best candidates for meta-barcoding of 16S rRNA gene and microbial community analysis based on

high eDNA yield. The V5–V6 region of the 16S rRNA (350-bp) gene was amplified in 25  $\mu\text{L}$  of PCR master mix consisting of 1 $\times$  buffer (including  $\text{Mg}^{2+}$ ), 0.2 mM dNTPs, 0.16  $\mu\text{M}$  primers UniA-V5F and UniB-V6R (Table 1), 1 unit Taq and 1  $\mu\text{L}$  DNA (He et al. 2017). Polymerase chain reaction (PCR) products were cleaned using Sera-Mag Magnetic Beads (GE, Healthcare Life Science, UK) and used as a template for a second, short-cycle, PCR to ligate the adaptor and barcode sequences for NGS. The second PCR was conducted in 25  $\mu\text{L}$  with the same master mix as for the first PCR but with the UniA and UniB primers (Table 1) (He et al. 2017). The second-round PCR products from all the samples (now barcoded to code for eDNA origin) were pooled and purified using QIAquick Gel Extraction Kit (QIAGEN, Toronto, ON, Canada) following the manufacturer's instructions. The concentration of purified PCR product mix (library) was measured on an Agilent 2100 Bioanalyzer with a High Sensitivity DNA chip (Agilent Technologies, Mississauga, ON, Canada). The library was then diluted to 60  $\text{pmol}/\mu\text{L}^{-1}$  and sequenced on an Ion PGM™ System using the Ion PGM™ Sequencing 400 kit and an Ion 318™ Chip (Thermo Fisher Scientific, Burlington, ON, Canada).

### Bioinformatics and statistical analyses

#### eDNA yield

The impact of different lysis buffers, different purification methods and sampling date as independent variables on eDNA yield was assessed using SPSS software (version 15.0) with a two-way univariate general linear model (ANOVA). Partial Eta squared was used to define the effect size of each independent variable and their interaction on eDNA yield. For those methods which yielded more than

**Table 1** | 16S rRNA primers and PCR conditions for the first and second round PCRs used to develop the NGS library using eDNA extracted using the three most promising methods

	Primers (5'–3')	PCR cycles	Annealing temperature
First PCR	UniA-V5F; acctgcctgccg-ATTAGATACCCNGGTAG UniB-V6R; acgccaccgagc-CGACAGCCATGCANCACT	25	55 °C
Second PCR	UniA; CATCTCATCCCTGCGTGTCTCCGACTCAGXXXXXXXXXXGATacctgcctgccg UniB; CCTCTCTATGGGCAGTCGGTGATacgccaccgagc	8	60 °C

The UniA primer in the second PCR (ligation of barcode and adaptor sequences for NGS) includes individually unique 10–12 bp (indicated by XXXXXXXXXXX) 'barcode' sequences that allow sorting of final sequence reads to the original sample after multiplexed sequencing (He et al. 2017).

a minimum quantity of eDNA ( $\geq 9 \text{ ng}/\mu\text{L}^{-1}$ ) Tukey's HSD was performed to detect significant differences in means of eDNA yield ( $p < 0.05$  was considered to be statistically significant).

### Microbial community

The raw NGS sequence data were de-multiplexed, quality filtered, and trimmed of the adaptor, barcode, and primer sequences using QIIME. Chimeras were also removed using ChimeraSlayer in QIIME (version 1.9.0-dev) (Caporaso *et al.* 2010). The operational taxonomic units (OTUs) were assembled based on sequence similarity (97%) among all sequence reads combined and then taxonomically assigned using RDP Classifier software with a minimum 80% confidence level (Edgar 2010). The identified bacteria (genus level) were deposited in GenBank (accession numbers MH547121–MH547202 and MB547000–MB547333). Alpha diversity indexes were calculated on the OTU table which was rarefied to 3,110 reads using QIIME (version 1.9.0-dev) (Caporaso *et al.* 2010). The singleton and doubleton OTUs were removed from the OTU table. The Kruskal–Wallis test (non-parametric) was used to evaluate the variation between replicates. We defined an OTU (also genera) as 'abundant' when it had a relative abundance  $\geq 1\%$  of the community, 'moderate' when it had a relative abundance between 0.1 and 1% of the community, and 'rare' when its abundance was below 0.1% (Logares *et al.* 2014). Principal coordinate analysis (PCoA) based on Bray–Curtis dissimilarity was used to investigate the variation in the microbial community structures among the extracted samples from different dates using PAST v3.12 software (Hammer *et al.* 2001). Pearson correlation coefficient was applied to measure correlation of the microbial community between the OTUs and genera with high relative abundance and the high abundant genera and those genera with potential harmful species for CTAB-PCI, SL-Bead-Robotic, and DB-Bead-Robotic extraction methods. Friedman non-parametric two-way ANOVA was used to compare alpha diversity indexes between date replicates and the microbial community of each extraction method. To visualize the overlap in microbial taxonomic diversity and OTU diversity, we prepared Venn diagrams in Venn Diagram Plotter v. 1.3.3250.34910 (Pacific Northwest

National Laboratory, <http://www.pnl.gov/>; <http://omics.pnl.gov/>) to show shared versus unique OTUs and taxa detected among the extraction methods. The relative abundance of the phyla/class, genera with potential pathogens, fecal indicators and microbial source tracking markers (Shannon *et al.* 2007) were compared between extraction methods.

### Cost-effectiveness and safety analysis

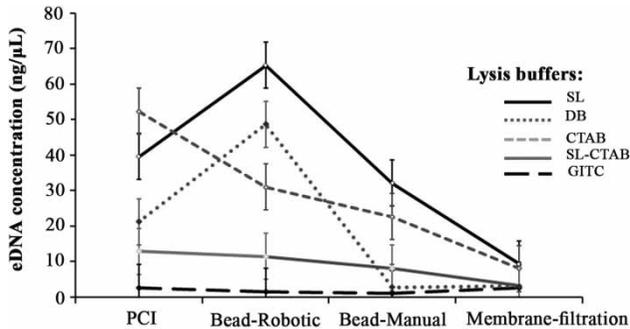
To evaluate which eDNA extraction method is best based on not only yield, but also on processing time, cost, and safety for users and the environment, we estimated the time required for eDNA extraction, cost of the chemicals, other supplies per sample, and hazard levels of the chemicals used for each method. Note, however, that equipment cost was not included in our analyses. Methods compared include (1) CTAB-PCI as a manual method, (2) SL-Bead-Robotic, and (3) DB-Bead-Robotic as robotic methods.

## RESULTS

### ANOVA for eDNA concentration

An ANOVA test was conducted on eDNA concentrations for all five lysis buffers and four purification methods with three replicates at three different dates. No significant variation ( $p > 0.05$ ) was detected among replicates and eDNA concentration on different dates. ANOVA analysis revealed that lysis buffers ( $F = 44.8$ ,  $p < 0.0001$ ), purification methods ( $F = 179.3$ ,  $p < 0.0001$ ), and the interaction between lysis buffer and extraction methods ( $F = 26.7$ ,  $p < 0.0001$ ) had highly significant effects on eDNA yield. The effect size of purification methods on eDNA yield was almost double (partial Eta squared = 0.85) compared to the effect size of lysis buffers (partial Eta squared = 0.48). The effect size of the interaction effect was intermediate (partial Eta squared = 0.61).

Among the five tested lysis buffers, averaged across their biological and technical replicates (nine samples for each method), SL, CTAB, and DB yielded the highest eDNA concentrations, while SL-CTAB and GITC yielded the lowest concentrations of eDNA (Figure 2). Among the four purifications methods, Bead-Robotic and PCI yielded



**Figure 2** | Yield of total eDNA (concentration  $\pm 1$  SD  $\text{ng}/\mu\text{L}^{-1}$ ) from 250 mL freshwater filtered samples using 20 combinations of five different lysis buffers and four different extraction methods. The values shown in the figure represent mean values of nine samples (three sampling times altogether) from each DNA extraction method and lysis buffer.

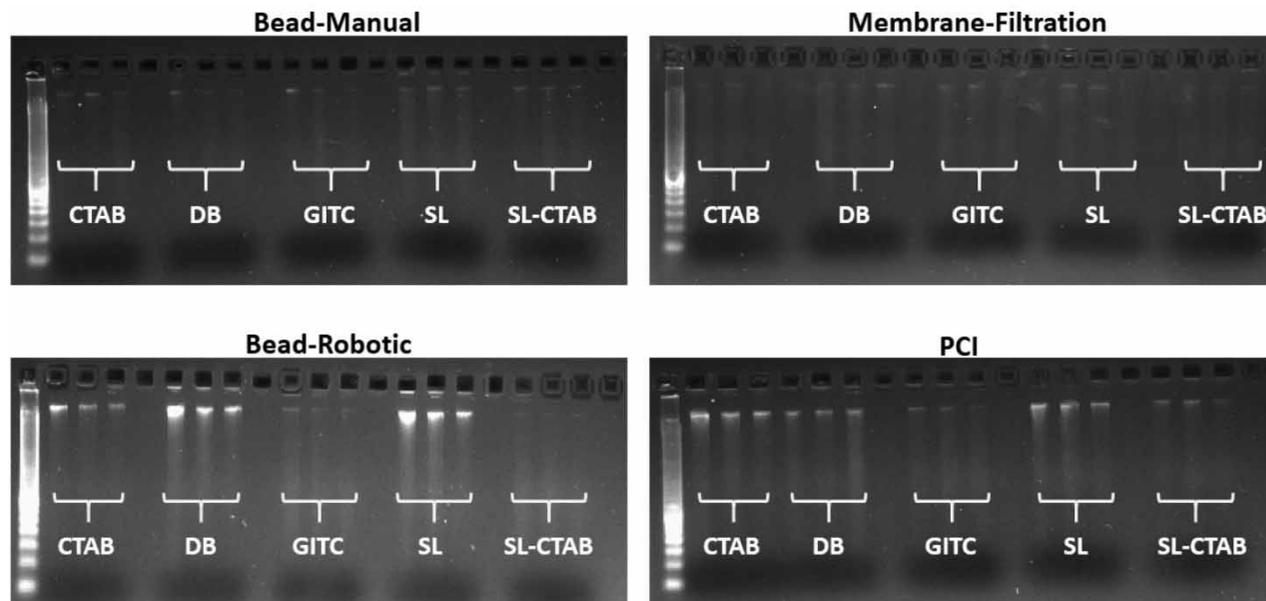
significantly higher eDNA concentrations. Although Bead-Manual and Bead-Robotic methods used the same beads, results indicated two- to three-fold higher concentrations were obtained for the robotic method than the manual method (Figure 3 and Supplementary material, Table S1, available with the online version of this paper). We defined a minimum threshold for eDNA quantity ( $9 \text{ ng}/\mu\text{L}^{-1}$ ) to select lysis buffers and their interactions with four purification methods (Bead-Robotic, Bead-Manual, PCI, and membrane-filtration) were thus included in Tukey's

post-hoc analysis (12 acceptable combinations). Post-hoc comparisons using Tukey's test indicated that eDNA yield by SL-Bead-Robotic (mean:  $65.33 \pm 12.9 \text{ ng}/\mu\text{L}$ ), CTAB-PCI (mean:  $55.33 \pm 5.4 \text{ ng}/\mu\text{L}^{-1}$ ), and DB-Bead-Robotic (mean:  $48.66 \pm 6.56 \text{ ng}/\mu\text{L}^{-1}$ ) were significantly higher ( $p < 0.05$ ) than other methods (CTAB-Bead-Manual, SL-Bead-Manual, CTAB-Bead-Robotic, CTAB-membrane-filtration, SL-membrane-filtration, SL-PCI, SL-CTAB-PCI, DB-PCI) (Figure 2 and Supplementary material, Table S1, available online). Hence, those samples were selected for meta-barcoding of 16S rRNA.

### Microbial meta-barcode community analysis

Nine PCR samples (which include samples collected at three sample dates and their three replicates) were sequenced for each of the selected high yield lysis – purification combination methods: CTAB-PCI, SL-Bead-Robotic and DB-Bead-Robotic methods (in total, 27 samples were meta-barcoded and sequenced).

After demultiplexing and filtering out poor-quality sequences, we obtained 104,731 high-quality sequences for the 27 samples. The number of sequences per sample ranged from 3,110 to 5,516 with an average of 3,978



**Figure 3** | Total eDNA ( $5 \mu\text{L}$ ) extracted by different lysis buffers (CTAB, DB, GITC, SL, and SL-CTAB) and purification methods (Bead-Manual, Bead-Robotic, membrane-filtration, and PCI), loaded on 2% agarose gel with GeneRuler 50 bp DNA Ladder (Thermo Scientific, USA) as size marker (first left lanes in each gel). The gels were visualized by Gel Doc XR system (BIORAD, Canada). eDNA extracted by DB and SL with Bead-Robotic method and CTAB with PCI method were selected for community analysis by meta-barcoding of 16S rRNA.

sequences per sample. The Kruskal–Wallis test results indicated that there was no significant variation ( $p > 0.05$ ) among the triplicates within each sample. So the triplicates were combined to enhance the read depth for further statistical analysis. Further, all OTUs data for each extraction method (CTAB-PCI, SL-Bead-Robotic, and DB-Bead-Robotic) were combined over the three dates of sampling. The purpose of the meta-barcoding analyses was to test for eDNA extraction protocol effects on bacterial community composition. Hence, by combining the data for the three sample dates, we achieved the maximum community diversity to test for eDNA extraction method effects (Supplementary material, Figure S1, available online). Further, combined data were used to test alpha diversity indexes, and evaluate the sensitivity of methods (CTAB-PCI, DB-Bead-Robotic, and SL-Bead-Robotic) to detect the microbial community at phyla/class level and microbial genera with potential pathogens. Also, abundant, moderately abundant, rare OTUs and genera, and method-specific OTUs and genera were analyzed.

## Operational taxonomic units (OTUs), community composition and diversity indexes

### Alpha diversity indexes

Shannon-H index comparison between CTAB-PCI ( $6.95 \pm 0.02$ ), SL-Bead-Robotic ( $6.82 \pm 0.22$ ), and DB-Bead-Robotic ( $5.55 \pm 0.18$ ) indicated that there was no significant difference between the CTAB-PCI and SL-Bead-Robotic. However, DB-Bead-Robotic showed a significant difference ( $p < 0.05$ ) with both CTAB-PCI and SL-Bead-Robotic. Moreover, the Chao-1 index showed the same pattern with a significant difference ( $p < 0.05$ ) between SL-

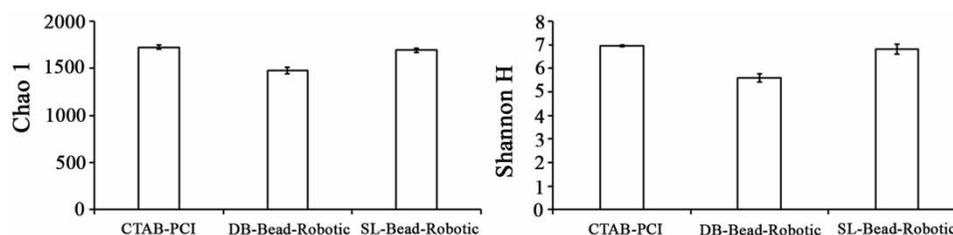
Bead-Robotic/CTAB-PCI methods and DB-Bead-Robotic ( $1,478 \pm 35$ ), while no significant difference ( $p > 0.05$ ) was found between SL-Bead-Robotic ( $1,692 \pm 19$ ) and CTAB-PCI ( $1,725 \pm 20$ ) (Figure 4).

### Microbial community at phyla/class level

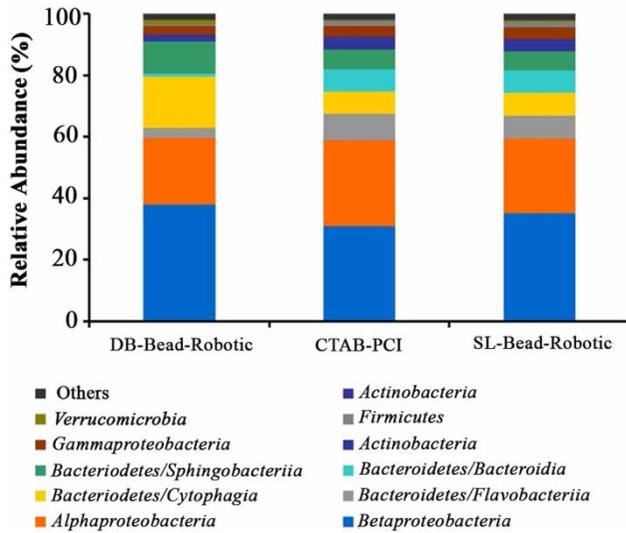
Overall, the relative contributions of *Betaproteobacteria* and *Alphaproteobacteria* were similar (no significant variation,  $p > 0.05$ ) among CTAB-PCI, SL-Bead-Robotic, and DB-Bead-Robotic. CTAB-PCI and SL-Bead-Robotic were significantly more sensitive ( $p < 0.05$ ) in detecting *Flavobacteriia* and *Bacteroidia* in the samples compared to DB-Bead-Robotic. However, DB-Bead-Robotic was significantly more sensitive ( $p < 0.05$ ) in detecting *Cytophagia* and *Sphingobacteriia* than the two other extraction methods (Figure 5).

### Microbial genera with potential pathogens

Comparison of assigned OTUs showed no significant variation ( $p > 0.05$ ) in the relative abundance of genera with potential harmful species (*Pseudomonas* and *Staphylococcus*), microbial source tracking marker species (*Bacteroides*), and fecal indicator species (*Escherichia*) between CTAB-PCI and SL-Bead-Robotic methods (Figure 6). CTAB-PCI was significantly more sensitive ( $p < 0.05$ ) in detecting *Acinetobacter*, *Burkholderia* and *Enterococcus* compared to SL-Bead-Robotic. However, compared to CTAB-PCI and SL-Bead-Robotic methods, only *Acinetobacter*, *Burkholderia*, *Legionella* (genus with potential pathogens), and *Escherichia* (fecal indicator) were detected by DB-Bead-Robotic (Figure 6), and their relative abundance was significantly lower ( $p < 0.05$ ) than observed in the CTAB-PCI and SL-Bead-Robotic method eDNAs.



**Figure 4** | Alpha-diversity indexes (left; Chao-1 and right; Shannon-H) based on microbial community meta-barcoding and OTU clustering at 97% sequence similarity for three eDNA extraction methods. Bars show mean values across all three sample dates and error bars are  $\pm 1$  SD.



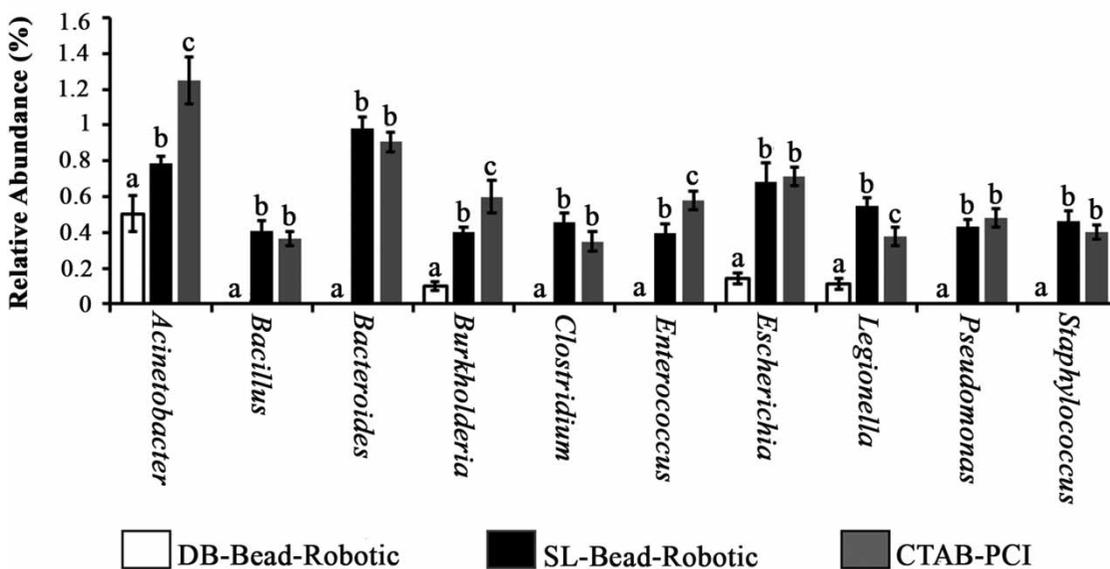
**Figure 5** | Relative abundance of bacterial phyla/class for the combined water samples ( $N = 9$ ) for each eDNA extraction method (DB: DB-Bead-Robotic, SL: SL-Bead-Robotic, and CTAB: CTAB-PCI). The bars show phyla/class with at least 0.1% relative abundance in combined water samples. The 'others' category includes the sum of all phyla that occurred at less than 0.1% relative abundance.

### Abundant OTUs and genera comparison

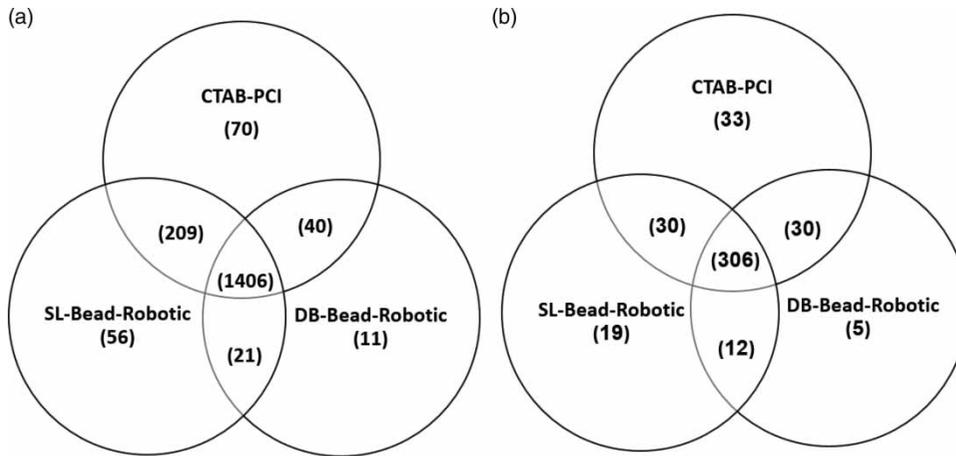
In total, 1,813 OTUs were detected by CTAB-PCI, SL-Bead-Robotic, and DB-Bead-Robotic eDNA extraction methods and, of these, 1,406 OTUs were shared by all three methods

(Figure 7(a)). Out of 1,406 shared OTUs, 14 OTUs reached the  $\geq 1\%$  threshold for abundant OTUs in all the sample sets from each extraction method. Among the high abundance OTUs, significant correlations were observed between the relative abundance of CTAB-PCI and SL-Bead-Robotic OTUs ( $r = 0.91$ ,  $p < 0.0001$ ), CTAB-PCI and DB-Bead-Robotic OTUs ( $r = 0.81$ ,  $p < 0.0001$ ), and DB-Bead-Robotic and SL-Bead-Robotic OTUs ( $r = 0.86$ ,  $p < 0.0001$ ). This indicates good agreement among the eDNA extraction combinations for abundant, or 'core' microbial community composition.

Out of 435 bacterial genera, 306 common genera were detected for all the eDNA samples (CTAB-PCI, SL-Bead-Robotic, and DB-Bead-Robotic) (Figure 7(b)). Among those, the high abundance ( $\geq 1\%$ ) genera identified by all three extraction methods were *Algoriphagus*, *Dechloromonas*, *Limnohabitans*, *Paludibacter*, *Pelomonas*, *Polynucleobacter*, *Rhodoferrax*, *Sphingorhabdus*, *Tabrizicola*, and *Terrimonas*. Significant correlations were observed between the relative abundance of these genera by CTAB-PCI and SL-Bead-Robotic ( $r = 0.93$ ,  $p < 0.0001$ ), CTAB-PCI and DB-Bead-Robotic ( $r = 0.79$ ,  $p < 0.0001$ ), and SL-Bead-Robotic and DB-Bead-Robotic OTUs ( $r = 0.80$ ,  $p < 0.0001$ ).



**Figure 6** | Relative abundance of genera known to include pathogenic species, microbial source tracking species, and fecal indicator species generated using meta-barcoding and eDNA produced using three extraction methods. Error bars show  $\pm 1$  SD. Different letters above the bars indicate significant differences ( $p < 0.05$ ) assessed using Tukey's post hoc multiple comparisons.



**Figure 7** | Venn diagrams showing the distribution and number (in parentheses) of the shared and method-specific (unique) microbial taxa identified by 16S rRNA meta-barcoding in water samples taken on three dates in August 2016: (a) OTUs and (b) genera. Note fewer taxa are shown for genus-level assigned groups due to unassigned OTUs and combined OTUs.

### Moderately abundant OTUs and genera composition

Out of 1,406 shared OTUs, 122 OTUs had relative abundance between 0.1 and 1% (moderately abundant OTUs) in the microbial community of all extraction methods. Significant correlations were observed between the relative abundance of CTAB-PCI and SL-Bead-Robotic among the moderately abundant OTUs ( $r = 0.93$ ,  $p < 0.0001$ ), CTAB-PCI and DB-Bead-Robotic OTUs ( $r = 0.73$ ,  $p < 0.0001$ ), and DB-Bead-Robotic and SL-Bead-Robotic OTUs ( $r = 0.82$ ,  $p < 0.0001$ ). Out of the 306 shared bacterial genera, 93 genera were moderately abundant among three different extraction methods. Significant correlations were observed between the relative abundance of the moderately abundant genera ( $r = 0.95$ ,  $p < 0.0001$ ) among CTAB-PCI and SL-Bead-Robotic, among CTAB-PCI and DB-Bead-Robotic OTUs ( $r = 0.75$ ,  $p < 0.0001$ ), and between DB-Bead-Robotic and SL-Bead-Robotic OTUs ( $r = 0.88$ ,  $p < 0.0001$ ).

### Rare OTUs and genera comparison

Out of 1,406 shared OTUs, 1,270 OTUs appeared in the microbial community of all extraction methods as rare ( $>0.1\%$ ). Significant correlations were observed between the relative abundance of CTAB-PCI and SL-Bead-Robotic among the rare OTUs ( $r = 0.89$ ,  $p < 0.0001$ ), CTAB-PCI and DB-Bead-Robotic OTUs ( $r = 0.63$ ,  $p < 0.0001$ ), and

DB-Bead-Robotic and SL-Bead-Robotic OTUs ( $r = 0.67$ ,  $p < 0.0001$ ). Out of the 306 shared bacterial genera, 203 genera were identified as rare genera among the three different extraction methods. Significant correlations were observed between the relative abundance of the rare genera ( $r = 0.83$ ,  $p < 0.0001$ ) among CTAB-PCI and SL-Bead-Robotic, among CTAB-PCI and DB-Bead-Robotic OTUs ( $r = 0.69$ ,  $p < 0.0001$ ), and between DB-Bead-Robotic and SL-Bead-Robotic OTUs ( $r = 0.74$ ,  $p < 0.0001$ ).

### Method-specific OTUs and genera comparison

Out of 1,813 detected OTUs, 1,406 OTUs were detected among the microbial community of all extraction methods (shared OTUs), while 407 OTUs were only detected by one or two extraction methods (method-specific OTUs). Out of 407 method-specific OTUs, 382 were classified as rare OTUs while 25 OTUs were grouped as moderately abundant OTUs. CTAB-PCI (with 319 OTUs) followed by SL-Bead-Robotic (with 209 OTUs) were more effective methods than the DB-Bead-Robotic method (with 72 OTUs) in detecting method-specific OTUs (Figure 7(a)). The detailed comparison showed 70 and 56 method-specific OTUs for CTAB-PCI and SL-Bead-Robotic methods, respectively, while DB-Bead-Robotic identified only 11 method-specific OTUs (Figure 7(a)). At the genus level, out of 435 bacterial genera, 306 genera were shared among all three extraction methods while 129 genera were method-specific

(112 rare genera and 17 moderately abundant genera). Out of 129 method-specific genera, 30, 19, and 5 genera were detected by CTAB-PCI, SL-Bead-Robotic, and DB-Bead-Robotic, respectively (Figure 7(b)). These findings together show that CTAB-PCI and SL-Bead-Robotic are more effective in the extraction of eDNA from specific genera than DB-Bead-Robotic. For example, hard-to-lyse Gram-positive bacteria including *Bacillus* and *Clostridium* (spore-forming bacteria) and *Staphylococcus* were detected as method-specific genera for both CTAB-PCI and SL-Bead-Robotic methods but not by DB-Bead-Robotic method (Figure 6).

### Cost-effectiveness and safety analysis

Our assessment of the best method to extract eDNA from aquatic samples to characterize the microbial community includes not only technical parameters (such as DNA yield and meta-barcoding bias), but also the costs of the methods: those costs include financial, human safety, and environmental impact costs. The financial analysis shows that extraction of genomic eDNA using magnetic beads in an automated liquid handling platform is more cost-effective than PCI or membrane-filtration for large sample sizes (Table 2). Additionally, magnetic beads are safe for the operator and environmentally friendly, while phenol and chloroform in the PCI method and sodium iodide (NaI) in the membrane-filtration method are neither safe for the user nor environmentally sustainable.

## DISCUSSION

The effect of the choice of lysis buffers on the detected microbial community composition has been addressed by

other researchers using fecal, clinical, and soil samples (Carrigg *et al.* 2007; Salonen *et al.* 2010; Willner *et al.* 2012; Gupta *et al.* 2016). However, the impact of different lysis buffers and DNA purification methods on aquatic (fresh-water) eDNA extraction, and consequently on the efficacy of aquatic microbial community composition characterization by eDNA meta-barcoding is poorly addressed. The criteria employed in this study for the evaluation of the five different lysis buffers and four different purification techniques included high cell lysis efficiency to obtain the maximum eDNA yield coupled with the most detailed meta-barcode detection of the microbial community composition. Out of five different lysis buffers, SL, CTAB, and DB yielded the highest eDNA concentrations and among the four purification methods, significantly higher eDNA concentrations were extracted by Bead-Robotic and PCI (Supplementary material, Table S1). The differences observed between eDNA quantity and quality by Bead-Manual in comparison to Bead-Robotic would be related to the time required to process the samples in the manual format (in the bead-based DNA purification protocol it is crucial to let the beads dry without any ethanol; however, too much drying will also impact the resuspension of the DNA into the solution). Diversity indexes (Figure 4) and analysis of method-specific OTUs (Figure 7) show that CTAB and SL lysis buffers extracted eDNA more efficiently from diverse taxa than DB lysis buffer. CTAB selectively precipitates the nucleic acids (Kim *et al.* 1990), and a high concentration of sucrose creates osmotic pressure resulting in microbial cell wall bursts (Boström *et al.* 2004). SDS as a key component of DB lysis buffer breaks up cell membrane structures (detergent); however, SDS is more effective in cell lysis of Gram-negative bacteria than Gram-positive and needs to be carefully optimized to extract eDNA from

**Table 2** | Assessment of consumables cost per sample and process duration of four eDNA extraction methods

Extraction methods	Time for sample preparation	Time DNA extraction/plate (96 samples)	Cost (\$)/sample	Safe	High throughput
PCI	–	6 h	2.5	No	– <sup>a</sup>
Bead-Robotic	3 h	20 min	1	Yes	Yes
Bead-Manual	3 h	2 h	1	Yes	–
Membrane-filtration	3 h	20 min	2	No	Yes

<sup>a</sup>Hard to be high-throughput.

environmental samples (Nishiguchi *et al.* 2002). Different microbial community composition detected by CTAB-PCI and SL-Bead-Robotic compared to DB-Bead-Robotic shows that DNA extraction might introduce bias in the detected microbial community. This bias was also observed at OTUs and genera level (Figures 6 and 7). In our study, two extraction methods including CTAB-PCI and SL-Bead-Robotic showed no significant variation in the global microbial community, OTUs, and genera with potential pathogens; however, the microbial community generated by DB-Bead-Robotic was different at different levels (global microbial community, OTUs, etc.) to CTAB-PCI and SL-Bead-Robotic. The detection of rare taxa in the freshwater microbial community was deemed to be of high importance, as these taxa often represent the groups of greatest concern for human health and for ecological microbial community health. Psifidi *et al.* (2015) found that out of 11 different DNA extraction methods, including three different commercial kits, the PCI method and bead extraction (in-house developed magnetic beads) were most appropriate for extracting high quality and quantity DNA from blood samples suitable for microarray analysis and NGS. Our findings also show that the bead purification method coupled with the optimal lysis buffer yielded high quality eDNA for massively parallel sequencing (NGS and meta-barcoding). In another study, Kennedy *et al.* (2014) reported that commercial kits have high variability in their DNA extraction yield (228–561 ng/ $\mu\text{L}^{-1}$  and 9–36 ng/ $\mu\text{L}^{-1}$  using the FastDNA kit and the MoBio kit, respectively) from human fecal samples, which may lead to not only loss of rare taxa but also divergent core microbiome results. In our study, eDNA yield by SL-Bead-Robotic ( $65.33 \pm 12.9$  ng/ $\mu\text{L}^{-1}$ ), CTAB-PCI ( $55.33 \pm 5.4$  ng/ $\mu\text{L}^{-1}$ ), and DB-Bead-Robotic ( $48.66 \pm 6.56$  ng/ $\mu\text{L}^{-1}$ ) was much higher than the MoBio kit (9–36 ng/ $\mu\text{L}^{-1}$ ) but lower than DNA yield using the FastDNA kit in the last study (Kennedy *et al.* 2014). However, we did not use any commercial kit in the current study and direct comparison of the extraction methods which we employed in this study with those kits is not possible. Our work shows that the core microbiome structures detected using CTAB-PCI and SL-Bead-Robotic were similar but had some variation compared to DB-Bead-Robotic (Figure 5). CTAB-PCI detected more rare taxa and both CTAB-PCI and SL-Bead-Robotic were more

effective in the detection of eDNA from bacteria associated with human health issues. More recently, magnetic bead purification coupled with automated liquid handling platforms for DNA purification from microbial communities have been recommended (Marotz *et al.* 2017); however, they used a commercial kit (MoBio PowerMag Soil DNA isolation kit, Qiagen, CA) to lyse the samples, which is not cost-effective for large sample size in research or monitoring laboratories. In that study, even by using commercial kit for lysis of the samples, the average concentration of extracted DNA across all robotic purification platforms was 10–20 ng/ $\mu\text{L}^{-1}$  for high biomass samples (Marotz *et al.* 2017), while in our study, depending on lysis buffer, the robotic platform purified much more eDNA concentration by SL lysis buffer ( $65.33 \pm 12.90$  ng/ $\mu\text{L}^{-1}$ ), DB lysis buffer ( $48.66 \pm 6.56$  ng/ $\mu\text{L}^{-1}$ ), and CTAB lysis buffer ( $30.49 \pm 6.82$  ng/ $\mu\text{L}^{-1}$ ) and low eDNA concentration by GITC lysis buffer ( $2.02 \pm 0.65$  ng/ $\mu\text{L}^{-1}$ ) (Supplementary material, Table S1). All these results show that selecting the appropriate lysis buffer and eDNA purification is an important and critical step in microbial community analysis studies.

## CONCLUSION

CTAB-PCI and SL-Bead-Robotic methods worked much better than the DB-Bead-Robotic method to extract high quantity eDNA from aquatic environmental samples and ensure the recovery of not only the core microbiome but also the rare critical taxa in the community. However, compared to other tested methods, the SL-Bead-Robotic is more rapid and user/environmentally safe. This, coupled with relatively low cost (does not include the infrastructure cost) and high throughput makes it the optimal solution for eDNA extraction for microbial community applications.

## ACKNOWLEDGEMENTS

This project was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) post-graduate funding to AHS and NSERC Strategic Partnership Grants for Projects funding to DDH for laboratory and field costs. We thank Russell Hepburn and

Shelby Mackie (Environmental Genomics Facility, Great Lakes Institute for Environmental Research, Windsor University) for their support and assistance in eDNA purification and sequencing of the samples.

## REFERENCES

- Biswas, C., Chen, S.-A., Halliday, C., Kennedy, K., Playford, E., Marriott, D., Slavin, M., Sorrell, T. & Sintchenko, V. 2017 Identification of genetic markers of resistance to echinocandins, azoles 5-fluorocytosine in *Candida glabrata* by next-generation sequencing: a feasibility study. *Clinical Microbiology and Infection* **23** (9), 677–676.
- Boström, K. H., Simu, K., Hagström, Å. & Riemann, L. 2004 Optimization of DNA extraction for quantitative marine bacterioplankton community analysis. *Limnology and Oceanography: Methods* **2** (11), 365–373.
- Cabral, J. P. 2010 Water microbiology. Bacterial pathogens and water. *International Journal of Environmental Research and Public Health* **7** (10), 3657–3703.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K. & Gordon, J. I. 2010 QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7** (5), 335–336.
- Carrigg, C., Rice, O., Kavanagh, S., Collins, G. & O'Flaherty, V. 2007 DNA extraction method affects microbial community profiles from soils and sediment. *Applied Microbiology and Biotechnology* **77** (4), 955–964.
- Chaganti, S. R., Lalman, J. A. & Heath, D. D. 2012 16S rRNA gene-based analysis of the microbial diversity and hydrogen production in three mixed anaerobic cultures. *International Journal of Hydrogen Energy* **37** (11), 9002–9017.
- Cloutier, D. D. & McLellan, S. L. 2017 Distribution and differential survival of traditional and alternative indicators of fecal pollution at freshwater beaches. *Applied and Environmental Microbiology* **83** (4), e02881–e02816.
- Currie, A. R., Tait, K., Parry, H., De Francisco-Mora, B., Hicks, N., Osborn, A. M., Widdicombe, S. & Stahl, H. 2017 Marine microbial gene abundance and community composition in response to ocean acidification and elevated temperature in two contrasting coastal marine sediments. *Frontiers in Microbiology* **22** (8), 1599.
- Edgar, R. C. 2010 Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26** (19), 2460–2461.
- Elphinstone, M. S., Hinten, G. N., Anderson, M. J. & Nock, C. J. 2003 An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Molecular Ecology Resources* **3** (2), 317–320.
- Gupta, P., Manjula, A., Rajendhran, J., Gunasekaran, P. & Vakhlu, J. 2016 Comparison of metagenomic DNA extraction methods for soil sediments of high elevation Puga hot spring in Ladakh, India to explore bacterial diversity. *Geomicrobiology Journal* **34** (4), 289–299.
- Hammer, Ø., Harper, D. A. T. & Ryan, P. D. 2001 Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4** (1), 9–18.
- He, X., Chaganti, S. R. & Heath, D. D. 2017 Population-specific responses to interspecific competition in the gut microbiota of two Atlantic salmon (*Salmo salar*) populations. *Microbial Ecology* **75** (1), 140–151.
- Kennedy, N. A., Walker, A. W., Berry, S. H., Duncan, S. H., Farquarson, F. M., Louis, P. & Thomson, J. M. 2014 The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. *PLoS One* **24** (9), e88982.
- Kim, W. K., Mauthe, W., Hausner, G. & Klassen, G. R. 1990 Isolation of high molecular weight DNA and double-stranded RNAs from fungi. *Canadian Journal of Botany* **68** (2), 1898–1902.
- Kotlowski, R., Martin, A., Ablordey, A., Chemlal, K., Fonteyne, P.-A. & Portaels, F. 2004 One-tube cell lysis and DNA extraction procedure for PCR-based detection of *Mycobacterium ulcerans* in aquatic insects, molluscs and fish. *Journal of Medical Microbiology* **53** (9), 927–933.
- Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., Claverie, J.-M., Decelle, J., Dolan, J. R. & Dunthorn, M. 2014 Patterns of rare and abundant marine microbial eukaryotes. *Current Biology* **24** (8), 813–821.
- Lucena-Aguilar, G., Sánchez-López, A. M., Barberán-Aceituno, C., Carrillo-Avila, J. A., López-Guerrero, J. A. & Aguilar-Quesada, R. 2016 DNA source selection for downstream applications based on DNA quality indicators analysis. *Biopreservation and Biobanking* **14** (2), 264–270.
- Marotz, C., Amir, A., Humphrey, G., Gaffney, J., Gogul, G. & Knight, R. 2017 DNA extraction for streamlined metagenomics of diverse environmental samples. *Biotechniques* **62** (6), 290–293.
- Mohiuddin, M. M., Salama, Y., Schellhorn, H. E. & Golding, G. B. 2017 Shotgun metagenomic sequencing reveals freshwater beach sands as reservoir of bacterial pathogens. *Water Research* **115** (15), 360–369.
- Nishiguchi, M. K., Doukakis, P., Egan, M., Kizirian, D., Phillips, A., Prendini, L., Rosenbaum, H. C., Torres, E., Wyner, Y. & DeSalle, R. 2002 DNA isolation procedures. In: *Techniques in Molecular Systematics and Evolution* (D. S. Rob, G. Gonzalo & W. Ward, eds). Springer Basel AG, New York, USA, pp. 249–287.
- Psifidi, A., Dovas, C. I., Bramis, G., Lazou, T., Russel, C. L., Arsenos, G. & Banos, G. 2015 Comparison of eleven methods for genomic DNA extraction suitable for large-scale whole-genome genotyping and long-term DNA banking using blood samples. *PLoS One* **10** (1), e0115960.
- Ramirez, K. S., Knight, C. G., De Hollander, M., Brearley, F. Q., Constantinides, B., Cotton, A., Creer, S., Crowther, T. W., Davison, J. & Delgado-Baquerizo, M. 2018 Detecting macroecological patterns in bacterial communities across

- independent studies of global soils. *Nature Microbiology* **3** (2), 189–196.
- Salonen, A., Nikkilä, J., Jalanka-Tuovinen, J., Immonen, O., Rajilić-Stojanović, M., Kekkonen, R. A., Palva, A. & De Vos, W. M. 2010 Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *Journal of Microbiological Methods* **81** (2), 127–134.
- Shannon, K., Lee, D.-Y., Trevors, J. & Beaudette, L. 2007 Application of real-time quantitative PCR for the detection of selected bacterial pathogens during municipal wastewater treatment. *Science of the Total Environment* **382** (1), 121–129.
- Taberlet, P., Coissac, E., Hajibabaei, M. & Rieseberg, L. H. 2012 Environmental DNA. *Molecular Ecology* **21** (8), 1789–1793.
- Tagliani, E., Hassan, M. O., Waberi, Y., De Filippo, M. R., Falzon, D., Dean, A., Zignol, M., Supply, P., Abdoukader, M. A. & Hassangue, H. 2017 Culture and Next-generation sequencing-based drug susceptibility testing unveil high levels of drug-resistant-TB in Djibouti: results from the first national survey. *Scientific Reports* **7** (1), 17672.
- Tringe, S. G. & Hugenholtz, P. 2008 A renaissance for the pioneering 16S rRNA gene. *Current Opinion in Microbiology* **11** (5), 442–446.
- Vo, A. T. & Jedlicka, J. 2014 Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for faecal and swab samples. *Molecular Ecology Resources* **14** (6), 1183–1197.
- Wang, Y., Sheng, H.-F., He, Y., Wu, J.-Y., Jiang, Y.-X., Tam, N. F.-Y. & Zhou, H.-W. 2012 Comparison of the levels of bacterial diversity in freshwater, intertidal wetland, and marine sediments by using millions of illumina tags. *Applied and Environmental Microbiology* **78** (23), 8264–8271.
- Willner, D., Daly, J., Whiley, D., Grimwood, K., Wainwright, C. E. & Hugenholtz, P. 2012 Comparison of DNA extraction methods for microbial community profiling with an application to pediatric bronchoalveolar lavage samples. *PloS One* **7** (4), e34605.
- Yuan, S., Cohen, D. B., Ravel, J., Abdo, Z. & Forney, L. J. 2012 Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PloS One* **7** (3), e33865.

First received 24 April 2018; accepted in revised form 7 October 2018. Available online 30 October 2018