Rapid library preparation for 16S rRNA sequencing

Keywords: DNA sequencing, Metagenomics, Amplicon Sequencing, 16S rRNA, sparQ HiFi PCR Mastermix, AccuStart II PCR ToughMix®, repliQa HiFi ToughMix®, sparQ PureMag Beads

ABSTRACT

Metagenomic sequencing is useful for studying mixed populations of micro-organisms in a sample. In particular, 16S rRNA gene sequencing is a targeted amplicon sequencing method to analyze the taxonomic composition of microbial communities at the genus or even species level. The choice of enzyme for the PCR amplification steps can impact the sequencing error rate, the abundance of chimeric sequences and the species representation of the original sample. Therefore, PCR enzyme selection is critical to obtaining optimal data. Furthermore, the ease and speed of library preparation will vary with reagent choice, with potentially large impacts on sample throughput and costs. In this application note we detail the methodology for 16S rRNA library preparation with three Quantabio PCR master mixes, demonstrating fast, reliable workflows and high quality sequencing data suitable for a range of applications.

INTRODUCTION

Taxonomic identification of bacterial species is vital in many environmental and healthcare applications to inform on the appropriate course of action, for example in environmental surveillance of pathogens or for determination of antibiotic resistance for patient management. Regions of the 16S ribosomal RNA gene (16S rRNA) of bacteria are frequently used as a marker to identify phylogenetic composition (genus and species) of a mixed community. The bacterial 16S rRNA gene is approximately 1500 bp long and contains nine variable regions separated by conserved regions. Instead of sequencing the entire 16S rRNA gene, typically the variable V3 and V4 regions are used to identify bacterial species using next generation sequencing (NGS). A steep decline in sequencing costs has expanded the use of sequencing for the characterization of microbial communities from diverse environments, but library preparation protocols can still be laborious and time-consuming.

A typical amplicon-based 16S rRNA protocol takes 3 hours and involves two rounds of PCR and two bead cleanup steps. First, 460 bp sequences covering the variable V3 and V4 regions are amplified using locus-specific primers containing additional overhang sequences and products are purified. This PCR product then undergoes another round of amplification where the overhang sequences serve as priming sites for barcoded Illumina® sequencing adapters.[1] Following purification and normalization, libraries are sequenced and subsequent metagenomic data analysis carried out to identify the genus and species of the bacteria in the sample.

In this application note, we prepared 16S rRNA sequencing libraries from a test sample of three bacterial species using either AccuStart II PCR ToughMix, sparQ HiFi PCR Master Mix or repliQa HiFi ToughMix. The former two are widely used and well-cited in literature for 16S rRNA library preparation. The latter, repliQa HiFi ToughMix, is a new product in this field. We present a novel protocol using repliQa HiFi ToughMix for 16S rRNA library preparation in just 1.5 h with minimal hands-on steps, demonstrating a much faster method for amplicon sequencing.

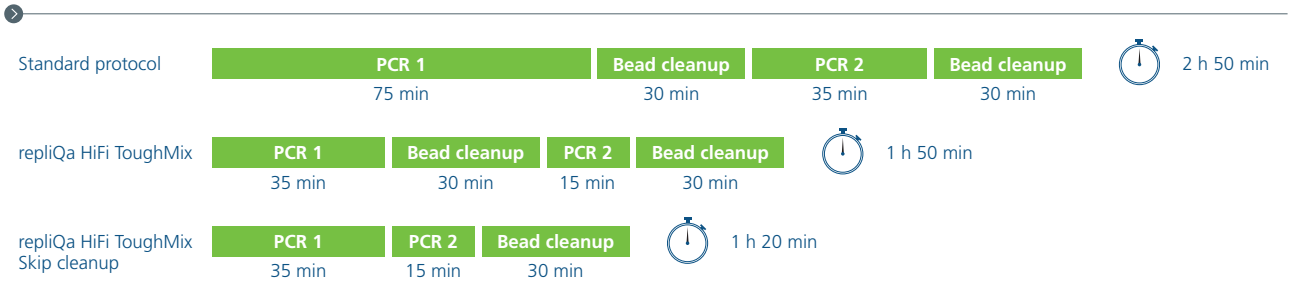

Figure 1 **Rapid protocol with repliQa HiFi ToughMix.** The standard protocol used for mixes sparQ HiFi PCR Master Mix and AccuStart II PCR ToughMix takes around 2 h 50 min. With repliQa HiFi ToughMix, this is reduced to only 1 h 50 min and further to 1 h 20 min without a bead cleanup step between PCR 1 and PCR 2.

Sample input

Libraries were prepared from mixed sample of 3 bacteria species in a 1:1:1 ratio by mass: *Fusobacterium nucleatum*, *Escherichia coli* and *Bordetella pertussis*. Species were chosen for their commonality in 16S studies and their variety of GC contents: *F. nucleatum* (27% GC), *E. coli* (50% GC) and *B. pertussis* (67% GC). Purified genomic DNA of *F. nucleatum* (# 25586D-5), *E. coli* (# 13706DX) and *B. pertussis* (#BAA-589DQ) were purchased from ATCC and diluted in TE (10 mM Tris, 0.1 mM EDTA) before use. For each reaction, 10 ng total input DNA was used.

Primers

Library preparation used a 2-step PCR method with region-specific PCR followed by barcoding PCR. Region-specific PCR for V3-V4 region of the 16S rRNA gene was carried out with V3-V4 Forward primer (5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG) and V3-V4 Reverse primer (5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC). The concentration of each primer was 1 µM.

Barcoding PCR was carried out with Nextera® XT Index Primers (N7XX) and Nextera XT Index 2 Primers (S5XX) from the Nextera XT Index Kit (Illumina, #FC-131-1001 or #FC-131-1002).

Library preparation

PCR 1 (Region Specific)

1. The Region Specific PCR master mix was prepared on ice with each PCR mix tested, according to Table 1. Then, 23 µl of master mix was transferred to a new thin-walled PCR tube for each reaction.

Component	Volume for 1 sample (µl)
Nuclease-free Water	5.5
2X PCR Mix	12.5
V3-V4 FP	2.5
V3-V4 RP	2.5
Total	23

Table 1 **Region-specific PCR master mix.** A master mix was prepared according to the volumes given in the table, scaled up for the number of samples with 10% overage. 2X PCR mix: AccuStart II PCR ToughMix, sparQ HiFi PCR Master Mix, repliQa HiFi PCR ToughMix or KAPA® HiFi HotStart Ready Mix. FP: forward primer, RP: reverse primer.

2. For each reaction, 2 µl DNA sample was added (10 ng total input DNA).
3. The PCR was run according to the cycling program for each PCR mix, detailed in table 2.

AccuStart II PCR ToughMix

Step	Temperature	Incubation time	Cycles
Initial denaturation	98°C	3 min	1
Denaturation	98°C	20 s	25
Anneal	55°C*	30 s	
Extend	72°C	30 s	
Final extension	72°C	5 min	1
Hold	4°C	Hold	1

sparQ HiFi PCR Master Mix/KAPA HotStart Ready Mix

Step	Temperature	Incubation time	Cycles
Initial denaturation	98°C	2 min	1
Denaturation	98°C	20 s	25
Anneal	55°C*	30 s	
Extend	72°C	30 s	
Final extension	72°C	5 min	1
Hold	4°C	Hold	1

* The annealing temperature should be adjusted for alternative primer sets.

repliQa HiFi ToughMix

Step	Temperature	Incubation time	Cycles
Initial denaturation	98°C	2 min	1
Denaturation	98°C	10 s	25
Anneal	55°C*	1 s	
Extend	68°C	1 s	
Hold	4°C	Hold	1

Table 2 PCR cycling conditions for region specific PCR.

* The annealing temperature should be adjusted for alternative primer sets.

Clean up of region-specific PCR products

For the rapid library preparation protocol using repliQa HiFi ToughMix with no bead clean up, steps 4-8 were excluded and undiluted PCR products from step 3 were used directly as input for barcoding PCR (Step 9 b).

4. 1X sparQ PureMag Beads (25 µl) was added to the completed PCR reaction and incubated at room temperature (RT) for 5 min. The beads were then pelleted on a magnet and the supernatant removed and discarded.

With repliQa HiFi ToughMix, diffuse pelleting of the beads may occur. If this hinders the ability to remove the supernatant, it can be overcome with the addition of 1 µl 10% Tween 20 to the PCR products before clean up. Alternatively, Proteinase K can be added to PCR products to a final concentration of 0.2 mg/ml and incubated for 1 min at RT before bead clean up. See repliQa HiFi ToughMix IFU for full details.

5. The tube was kept on the magnet, and 200 µl freshly prepared 80% ethanol added to wash the beads. The beads were allowed to re-pellet for 30 s before removing and discarding the ethanol. The wash was repeated for a total of two washes.

6. Keeping the tube on the magnet, the bead pellet was dried for 5 mins or until the beads appeared matte and no longer shiny.

7. To elute the PCR products, the pellet was resuspended in 52.5 µl 10 mM Tris-HCl pH 8.0 and incubated at RT for 5 min.

8. The beads were then pelleted on a magnet and 50 µl clean PCR products were collected.

Barcoding PCR

9 a. Following bead up: The Barcoding PCR reaction was set up for each sample according to table 3, using a different index for each sample. DNA input was 5 µl of purified eluted PCR products from step 8.

Component	Volume (µl)
Nuclease-free Water	10
2X PCR Mix	25
DNA	5
Nextera XT 1 (N7)	5
Nextera XT 1 (S5)	5
Total	50

Table 3 Barcoding PCR set up with purified PCR products input.
2X PCR Mix: AccuStart II PCR ToughMix, sparQ HiFi PCR Master Mix, repliQa HiFi PCR ToughMix or KAPA HiFi HotStart Ready Mix.

9 b. For the rapid library preparation protocol using repliQa HiFi ToughMix with no bead clean up: The Barcoding PCR reaction was set up for each sample according to Table 4, using a different index for each sample. DNA input was 2 µl of crude PCR products from step 3.

Component	Volume (µl)
Nuclease-free Water	13
2X PCR Mix	25
DNA	2
Nextera XT 1 (N7)	5
Nextera XT 1 (S5)	5
Total	50

Table 4 Barcoding PCR set up with crude PCR products input.

10. The PCR was run according to the cycling program for each PCR Mix, detailed in Table 5.

AccuStart II PCR ToughMix

Step	Temperature	Incubation time	Cycles
Initial denaturation	98°C	3 min	1
Denaturation	98°C	20 s	8
Anneal	55°C*	30 s	
Extend	72°C	30 s	
Final extension	72°C	5 min	1
Hold	4°C	Hold	1

sparQ HiFi PCR Master Mix/KAPA HotStart Ready Mix

Step	Temperature	Incubation time	Cycles
Initial denaturation	98°C	2 min	1
Denaturation	98°C	20 s	8
Anneal	55°C*	30 s	
Extend	72°C	30 s	
Final extension	72°C	5 min	1
Hold	4°C	Hold	1

repliQa HiFi ToughMix

Step	Temperature	Incubation time	Cycles
Initial denaturation	98°C	2 min	1
Denaturation	98°C	10 s	8
Anneal	55°C*	1 s	
Extend	68°C	1 s	
Hold	4°C	Hold	1

Table 5 PCR cycling conditions for barcoding PCR.

* The annealing temperature should be adjusted for alternative primer sets.

Final cleanup of barcoded libraries

11. 1X sparQ PureMag Beads (50 µl) was added to the completed PCR reaction and incubated at room temperature (RT) for 5 min. The beads were then pelleted on a magnet and the supernatant removed and discarded.

With repliQa HiFi ToughMix, diffuse pelleting of the beads may occur. If this hinders the ability to remove the supernatant, it can be overcome with the addition of 1 µl 10% Tween 20 to the PCR products before clean up. Alternatively, Proteinase K can be added to PCR products to a final concentration of 0.2 mg/ml and incubated for 1 min at RT before bead clean up. See repliQa HiFi ToughMix IFU for full details.

12. The tube was kept on the magnet, and 200 µl of freshly prepared 80% ethanol was added to wash the beads. The beads were allowed to re-pellet for 30 s before removing and discarding the ethanol. The wash was repeated for a total of two washes.

13. Keeping the tube on the magnet, the bead pellet was dried for 5 mins or until the beads appeared matte and no longer shiny.

14. To elute the PCR products, the pellet was resuspended in 27.5 µl 10 mM Tris-HCl pH 8.0 and incubated at RT for 5 min.

15. The beads were then pelleted on a magnet and 25 µl clean PCR products were collected.

Library validation and quantification

To verify the size of the library, 1 µl of the purified library was run on a D1000 ScreenTape on the TapeStation System (Agilent). The expected size of the library was ~595 bp.

Libraries were quantified using Qubit™ dsDNA BR Assay Kit (Thermo Fisher Scientific) using manufacturer's recommendations: 2 µl purified library was added to 198 µl of Qubit working solution and mixed by vortexing. The tubes were incubated for 2 min before loading the tubes in Qubit 3.0 Fluorometer.

Sequencing

Libraries were normalized to 4 nM using EB buffer (QIAGEN). Diluted DNA (10 µl at 4 nM) from each library was mixed together in a new tube for pooling libraries with unique indices. For this experiment, 14 unique libraries were pooled together for the MiSeq run. Up to 96 libraries can be pooled for a MiSeq V3 run.

Pooled library was sequenced using MiSeq Reagent Kit v3 (600 cycles) (# MS-102-3003, Illumina) using manufacturer's recommendations, which are briefly described below.

For denaturing the DNA and final dilution, these steps were followed:

1. 5 µl of pooled library (4 nM) and 5 µl of freshly prepared 0.2 N NaOH were added to a microcentrifuge tube, mixed by gentle vortexing and centrifuged briefly.
2. This mixture was incubated at room temperature for 5 min.
3. 990 µl pre-chilled HT1 buffer (supplied with the MiSeq Reagent Kit v3) was added to the 10 µl denatured DNA. This resulted in a 20 pM denatured library.
4. 300 µl of denatured diluted library (20 pM) was combined with 300 µl of HT1 buffer to make 10 pM diluted library.
5. 5% PhiX control was added to the library mixture, 570 µl of the diluted library (10 pM) and 30 µl of diluted denatured PhiX control (10 pM) were mixed together.
6. The combined library and PhiX control tube was incubated at 96°C for 2 min and then immediately transferred into an ice block and incubated on ice for 5 min.
7. The entire mix was loaded in the designated well of the MiSeq v3 reagent cartridge.

Data analysis

Sequencing data was analyzed with CLC Genomics Workbench 20.0.4 software using the Microbial Genomics Module workflow (QIAGEN). Following filtering and trimming of data for quality, reads were mapped to the 16S rRNA database (rRNA SILVA) in the software and clustered by operational taxonomic unit (OTU). The microbial composition of each sample was analyzed as a function of the reads assigned to the sample for comparison (alpha diversity).

RESULTS

Standard protocol

The 16S Metagenomic Sequencing Library Preparation guide from Illumina is widely used as a basic protocol for 16S rRNA library preparation and sequencing on a MiSeq instrument.[1] This protocol involves first region-specific PCR, then bead cleanup, followed by barcoding PCR and a final bead cleanup. The workflow takes about 2 h 50 mins, including >1 h hands-on time (Figure 1). We first tested this protocol with three Quantabio PCR Master Mixes: AccuStart II PCR ToughMix, sparQ HiFi PCR Master Mix and repliQa HiFi ToughMix. repliQa HiFi ToughMix contains a highly processive enzyme that can achieve up to 1 kb extension of template in 1 second. Therefore, the workflow time was reduced to just 1 h 50 min with repliQa HiFi ToughMix as a result of faster PCR cycling. The example enzyme used in the Illumina protocol is KAPA HiFi HotStart ReadyMix, therefore this enzyme was also tested to benchmark results.

Preparation of libraries with all mixes tested resulted in high yield of PCR product at the expected size, as visualized by TapeStation electrophoresis (Figure 2). Libraries were sequenced on an Illumina MiSeq instrument and the data analyzed using CLC Genomics Workbench (QIAGEN). The most accurate representation of the three species in the sample was found from libraries prepared with sparQ HiFi PCR Master Mix, with close to 1:1:1 representation for each species (Figure 3). In addition, use of sparQ HiFi Master Mix for amplification resulted in the lowest proportion of chimeric reads, only 11.7% compared with 33.3% with the KAPA enzyme (Figure 4). In comparison, use of KAPA HiFi HotStart ReadyMix for library amplification resulted in over-representation of one species (*F. nucleatum*) and under-representation of another (*B. pertussis*). Therefore, libraries prepared with sparQ HiFi PCR Master Mix gave far more useable reads from the same depth of sequencing.

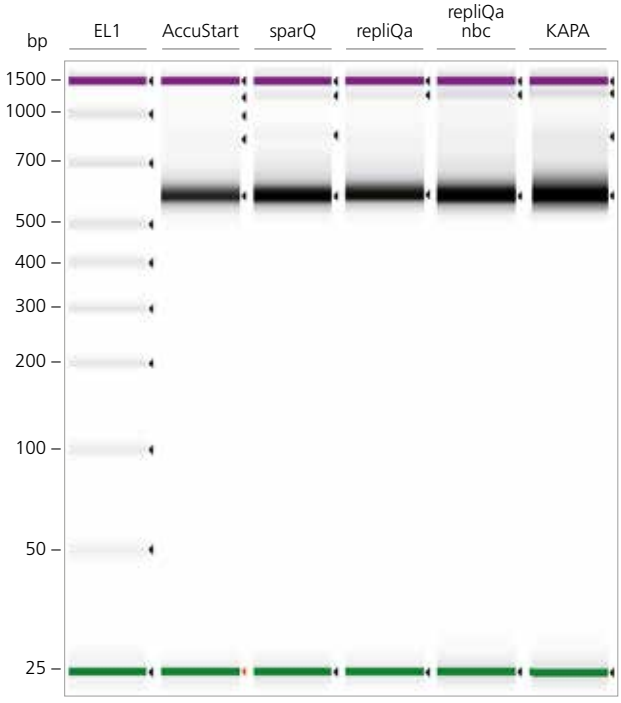

Figure 2 **High yield of specific PCR product with all mixes.** PCR products prepared with AccuStart II PCR ToughMix (AccuStart), sparQ HiFi PCR Master Mix (sparQ), repliQa HiFi ToughMix (repliQa), repliQa HiFi ToughMix without the first bead cleanup step (repliQa nbc) or KAPA HotStart Ready Mix (KAPA) were analyzed by TapeStation on a D1000 tape. The representative gel image demonstrates a high yield of PCR product at the expected size with all mixes tested.

These results suggest that sparQ HiFi PCR Master Mix gives very high performance for 16S rRNA library preparation with accurate representation of species in the sample and minimal chimeric reads. This has also been demonstrated for a range of amplicon sequencing methods with more complex input samples for example: 16S rRNA gene sequencing from plant roots and manure, mitochondrial 12S gene sequencing from lake and river water samples and ITS gene sequencing on DNA extracted from commercial plant products.[2–4]

AccuStart II PCR ToughMix also gave libraries with over-representation of *F. nucleatum* and under-representation of *B. pertussis*, and the highest rate of chimeric reads of the master mixes tested in this study. This result was surprising as previous studies have suggested that using a proofreading (high-fidelity) Taq polymerase can increase the rates of PCR artifacts and chimeras at high cycle numbers.[5,6] In addition, AccuStart II PCR ToughMix has been widely cited in literature for amplicon sequencing studies from a range of sample inputs including stool, urine, leaves, roots, soil and seawater.[7–11] The ToughMix formulation neutralizes PCR inhibitors, enabling exceptional performance even with crude input samples. This is especially useful for 16S metagenomic studies as they often involve environmental or clinical DNA samples that are rich in PCR inhibitors, such as polysaccharides, hemoglobin and humic acid.

For example, Mallot et al. compared amplification methods for gut microbial community profiling and found that AccuStart II PCR ToughMix gave favorable results for representation of individual taxa and the number of OTUs identified.[12] Along with the multiple references available for AccuStart II PCR ToughMix, this suggests that this mix is more suited for use with inhibitor-rich samples than the clean sample used in this study. AccuStart II PCR SuperMix may be a useful alternative for cleaner samples.[13,14]

Sequencing of libraries prepared using repliQa HiFi ToughMix with the faster cycling protocol gave more favorable metrics. All three bacteria species in the sample were identified in similar proportions to sparQ HiFi PCR Master Mix. There was an increase in the rate of chimeric sequences with repliQa HiFi ToughMix compared to sparQ HiFi PCR Master Mix (27.8% compared to 11.7%), but this was still less than that with KAPA HiFi HotStart ReadyMix (33.3%).

Accelerated 16S rRNA library preparation

Next, we investigated whether the ToughMix formula of repliQa HiFi ToughMix could overcome the need for bead cleanup after the first PCR step. The 16S rRNA region was amplified from the mixed bacterial sample as before, then 10% of this reaction (5 µl) was directly used as input for the barcoding PCR step. The PCR products were then purified by magnetic bead cleanup to produce the final sequencing library with a total workflow time of just 1 h 20 min. Sequencing of this library suggested that removing the first bead cleanup step had no impact on species identification or proportion of chimeric sequences read. Together, these data demonstrate that repliQa HiFi ToughMix is an excellent option for 16S rRNA library preparation, accelerating the workflow to just 1 h 20 min and removing a labor intensive hands-on cleanup step while delivering high quality sequencing data.

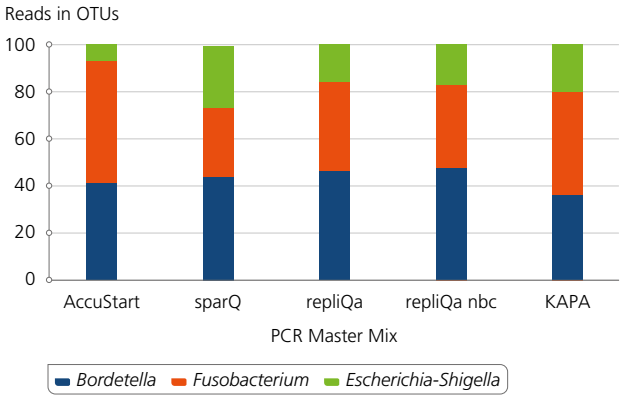

Figure 3 **Identification of bacteria in test sample.** Libraries prepared with AccuStart II PCR ToughMix (AccuStart), sparQ HiFi PCR Master Mix (sparQ), repliQa HiFi ToughMix (repliQa), repliQa HiFi ToughMix without the first bead cleanup step (repliQa nbc) or KAPA HotStart Ready Mix (KAPA) were sequenced on an Illumina MiSeq instrument and the reads in operational taxonomic units (OTUs) identified. All three species in the test sample were identified regardless of library preparation method.

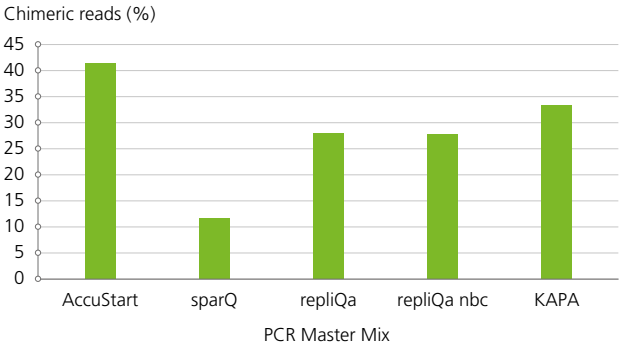

Figure 4 **Percentage of chimeric reads in each sample.** Libraries prepared with AccuStart II PCR ToughMix (AccuStart), sparQ HiFi PCR Master Mix (sparQ), repliQa HiFi ToughMix (repliQa), repliQa HiFi ToughMix without the first bead cleanup step (repliQa nbc) or KAPA HotStart Ready Mix (KAPA) were sequenced on an Illumina MiSeq instrument and the percentage of chimeric reads identified in each sample.

CONCLUSIONS

In this application note we demonstrate preparation of 16S rRNA libraries with three Quantabio PCR master mixes and demonstrate high quality sequencing data with each. AccuStart II PCR ToughMix is widely cited for a range of amplicon library preparation methods from crude samples, as the ToughMix formula overcomes PCR inhibitors, to produce high yielding libraries from clinical and environmental sample inputs. sparQ HiFi PCR Master Mix is a high-fidelity option for library preparation, delivering accurate representation of the species in the sample and a low rate of chimeric sequences. repliQa HiFi ToughMix is a highly processive enzyme that reduces protocol time from 2 h 50 min to just 1 h 50 min. We demonstrate that with repliQa HiFi ToughMix, the first bead clean up step can be excluded with no detriment to library yield or sequencing data, further reducing workflow time to 1 h 20 min. Together, these results provide options for rapid preparation of 16 rRNA libraries from even difficult input samples, which can be applied to various amplicon sequencing methods.

References

1. PCR Amplicon, PCR Clean-up, and Index PCR. 16S Metagenomic Sequencing Library Preparation; 2013.
2. Euclide PT, Lor Y, Spear MJ, et al. Environmental DNA metabarcoding as a tool for biodiversity assessment and monitoring: reconstructing established fish communities of north-temperate lakes and rivers. Divers Distrib. 2021.
3. Urumarudappa SKJ, Tungphatthong C, Prombutara P, Sukrong S. DNA metabarcoding to unravel plant species composition in selected herbal medicines on the National List of Essential Medicines (NLEM) of Thailand. Scientific reports. 2020;10(1):18259.
4. Wongkiew S, Koottatep T, Polprasert C, Prombutara P, Jinsart W, Khanal SK. Bioponic system for nitrogen and phosphorus recovery from chicken manure: Evaluation of manure loading and microbial communities. Waste management (New York, N.Y.). 2021;125:67-76.
5. Sze MA, Schloss PD. The Impact of DNA Polymerase and Number of Rounds of Amplification in PCR on 16S rRNA Gene Sequence Data. mSphere. 2019;4(3).
6. Ahn J-H, Kim B-Y, Song J, Weon H-Y. Effects of PCR cycle number and DNA polymerase type on the 16S rRNA gene pyrosequencing analysis of bacterial communities. Journal of microbiology (Seoul, Korea). 2012;50(6):1071-1074.
7. Marquart KA, Haller BR, Paper JM, et al. Influence of pH on the balance between methanogenesis and iron reduction. Geobiology. 2019;17(2):185-198.
8. Ugarelli K, Laas P, Stingl U. The Microbial Communities of Leaves and Roots Associated with Turtle Grass (Thalassia testudinum) and Manatee Grass (Syringodium filliforme) are Distinct from Seawater and Sediment Communities, but Are Similar between Species and Sampling Sites. Microorganisms. 2018;7(1).
9. Kaplan I, Bokulich NA, Caporaso JG, Enders LS, Ghanem W, Ingerslew KS. Phylogenetic farming: Can evolutionary history predict crop rotation via the soil microbiome? Evolutionary applications. 2020;13(8):1984-1999.
10. Peters BA, Wilson M, Moran U, et al. Relating the gut metagenome and metatranscriptome to immunotherapy responses in melanoma patients. Genome medicine. 2019;11(1):61.
11. Pinto D, Calabrese FM, Angelis M de, et al. Predictive Metagenomic Profiling, Urine Metabolomics, and Human Marker Gene Expression as an Integrated Approach to Study Alopecia Areata. Frontiers in cellular and infection microbiology. 2020;10:146.
12. Mallott EK, Malhi RS, Amato KR. Assessing the comparability of different DNA extraction and amplification methods in gut microbial community profiling. Access microbiology. 2019;1(7):e000060.
13. Stamps BW, Lyon WJ, Irvin AP, Kelley-Loughnane N, Goodson MS. A Pilot Study of the Effect of Deployment on the Gut Microbiome and Traveler's Diarrhea Susceptibility. Frontiers in cellular and infection microbiology. 2020;10:589297.
14. Ul-Hasan S, Bowers RM, Figueroa-Montiel A, et al. Community ecology across bacteria, archaea and microbial eukaryotes in the sediment and seawater of coastal Puerto Nuevo, Baja California. PloS one. 2019;14(2):e0212355.

MK-AN-0012 REV 02 16S Metagenomics 0122