invitrogen

Metagenomic study of bacteria in lake water using Platinum Direct PCR Universal Master Mix

Introduction

Metagenomic studies help find new microorganisms in the environment. 16S ribosomal RNA (rRNA) gene analysis is a standard approach for the investigation of microbial diversity in environmental samples. Traditional experimental workflows of bacterial metagenomic studies include sample collection, genomic DNA (gDNA) extraction, PCR amplification with degenerate primers targeting 16S rRNA gene sequences, next-generation sequencing (NGS), and bioinformatic analysis of sequencing data.

Extraction of gDNA can be time-consuming and carries a risk of losing metagenomic material required for bacterial identification. Invitrogen™ Platinum™ Direct PCR Universal Master Mix eliminates gDNA extraction steps by enabling DNA amplification directly from biological samples. The master mix is designed to work with a variety of sample types and enables primer annealing at 60°C. In this application note, we present a method for studying bacterial genera in lake water samples without the need to purify gDNA, thereby reducing the workflow time and risk of sample loss, using Platinum Direct PCR Universal Master Mix. We also provide an experimental design for direct amplification of 16S rRNA gene sequences, followed by NGS of amplicons and data analysis.

Materials and methods

Materials

- Invitrogen™ Dynabeads™ DNA DIRECT™ Universal Kit (**Cat. No. 63006**)
- Platinum Direct PCR Universal Master Mix (**Cat. No. A44647100**)
- Forward primer: 5´-CCTACGGGNGGCWGCAG-3´
- Reverse primer: 5´-GACTACHVGGGTATCTAATCC-3´
- Thermo Scientific™ nuclease-free water (**Cat. No. R0581**)
- Purified *E. coli* gDNA
- Invitrogen™ Collibri™ PS DNA Library Prep Kit for Illumina™ Systems, with CD indexes (**Cat. No. A38612024**)

Primers

Primer sequences for the bacterial 16S rRNA gene are from Klindworth et al. [1], which describes testing of the primers for their coverage ability. The authors indicate that the primer pair covers 86% of all bacteria and exhibits the largest number of matched phyla. The size of the expected amplicon using the primer set is 450 bp.

ThermoFisher SCIENTIFIC

PCR

Platinum Direct PCR Universal Master Mix was used to amplify bacterial DNA from three samples of a freshwater lake to generate sequencing templates. Bacterial DNA was either amplified after extraction using the Dynabeads DNA DIRECT Universal Kit or amplified directly from water without purification (noted as gDNA and direct, respectively, in Figure 1). A no-template control (NTC) was included as a negative control, and an *E. coli* gDNA sample was included as a positive control. To verify that reagents were free of bacterial contamination, a nuclease-free water sample (extraction negative control) was also included. Reaction setup and thermal cycling conditions were similar to those recommended in the user guide for Platinum Direct PCR Universal Master Mix (Tables 1 and 2).

DNA library preparation and sequencing

DNA sequencing libraries were prepared from PCR amplicons using the Collibri PS DNA Library Prep Kit for Illumina Systems, according to the kit protocol. Prepared libraries were sequenced on the Illumina™ MiSeq™ platform following a 2 x 300 bp paired-end reads protocol. After adapter trimming, quality assessment, and read cleanup, sequences were classified to the genus level by the closest match in the SILVA 132 database [2-4]. Matching reads, abundance, and distribution of identified taxa were calculated using Bracken software v2.1.0 [5].

Results

PCR

Bacterial DNA in three lake water samples was successfully amplified using Platinum Direct PCR Universal Master Mix. The direct amplification protocol was followed, in which water samples were added directly to the master mix. Amplification was also performed with gDNA extracted from the water samples to compare the bacterial diversity detected between methods. Analysis of PCR products on an agarose gel shows specific amplification of the 450 bp target from the lake water and positive control sample (Figure 2).

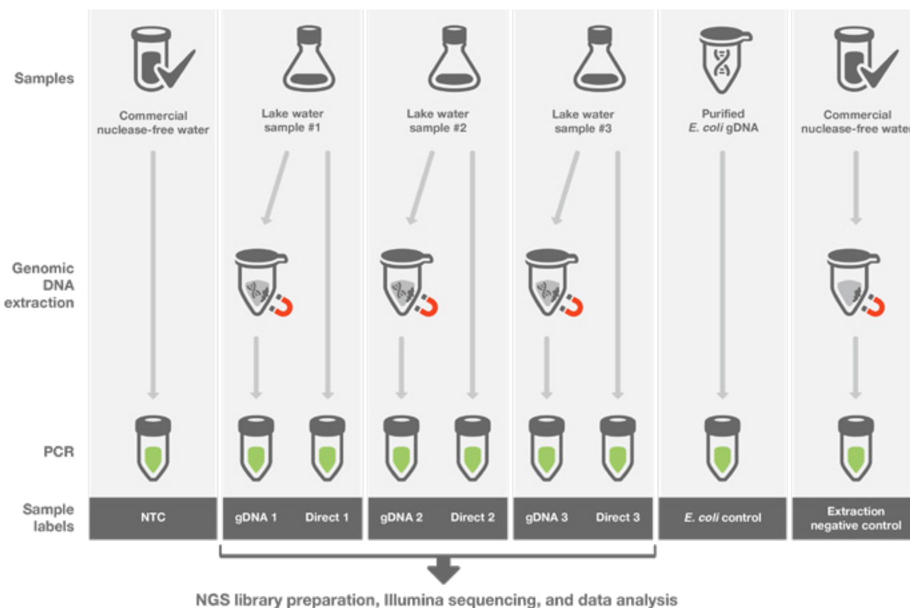

Figure 1. Experimental workflow for metagenomic study using PCR amplicons.

Table 1. PCR reaction setup.

Component	Per 50 µL reaction	Final concentration
Nuclease-free water	16 µL	–
2X Platinum Direct PCR Universal Master Mix	25 µL	1X
5 µM forward primer	2 µL	0.2 µM
5 µM reverse primer	2 µL	0.2 µM
Lake water sample (or 1 ng of gDNA)	5 µL	–

Table 2. PCR cycling protocol.

Step	Number of cycles	Temperature	Time
Initial denaturation	1	94°C	2 min
Denaturation	40	94°C	15 sec
Annealing		60°C	15 sec
Extension		68°C	20 sec
Hold	1	4°C	Indefinite

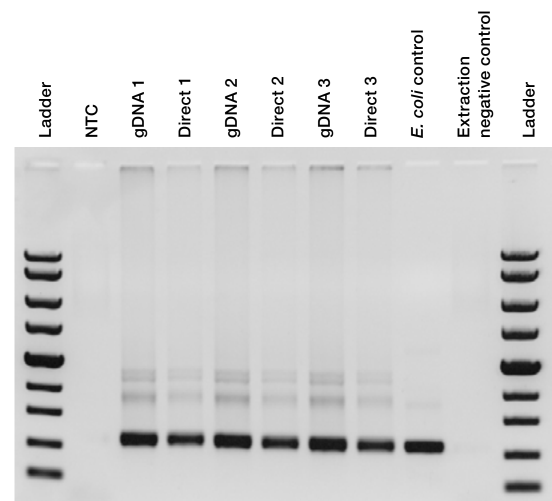

Figure 2. 16S rRNA gene amplification from water samples with Platinum Direct PCR Universal Master Mix. A 450 bp fragment was amplified from gDNA extracted from water samples (gDNA 1, 2, and 3) or amplified directly from water samples (direct 1, 2, and 3). Ladder: Thermo Scientific™ ZipRuler™ Express DNA Ladder Set.

NGS data analysis

PCR products of extracted gDNA and water samples were sequenced on the Illumina MiSeq platform, and data analysis was done based on the SILVA 132 database [2]. Shannon diversity indices (DIs) from the analysis indicate high quality of data from the prepared DNA libraries (Table 3), and a total of 52 bacterial genera were identified. The Shannon DI defines the diversity of species in a population relative to their abundance, where a larger value indicates higher diversity [6]. The index typically falls between 1.5 and 3.5, and rarely exceeds 4.5 [7].

The top 30 genera of bacteria identified in each sample are listed in Table 4. They are sorted by fraction of total reads in all of the samples, from highest to lowest. Their raw sequencing data are available in the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) database [8]. As expected, most of the bacteria identified are normally found in water and soil in the environment, and many of them are photosynthetic.

Table 3. Shannon DIs from computational analyses for the sequenced libraries.

Sample	Shannon DI
gDNA 1	3.3
Direct 1	4.1
gDNA 2	3.4
Direct 2	4.2
gDNA 3	3.5
Direct 3	4.2

Table 4. Bacteria identified in water samples, with their fraction of total sequencing reads.

SILVA 132 database match	gDNA 1	gDNA 2	gDNA 3	Direct 1	Direct 2	Direct 3
hgcI clade	0.279	0.274	0.229	0.096	0.117	0.093
Uncultured	0.055	0.065	0.083	0.132	0.131	0.166
Ramlibacter	0.128	0.131	0.126	0.083	0.075	0.069
Streptomyces	0.053	0.053	0.052	0.040	0.032	0.029
CL500-29 marine group	0.063	0.063	0.062	0.008	0.009	0.008
Pirellula	0.022	0.023	0.025	0.046	0.044	0.045
Burkholderia, *Caballeronia*, *Paraburkholderia*	0.039	0.038	0.034	0.021	0.025	0.028
ADurb.Bin063-1	0.005	0.006	0.006	0.051	0.041	0.057
Planktomarina	0.016	0.020	0.019	0.032	0.036	0.026
Nostoc PCC-7107				0.063	0.033	0.034
Polynucleobacter	0.033	0.028	0.026	0.006		0.006
Cyanobium PCC-6307	0.011	0.014	0.024	0.008	0.015	0.008
Luteolibacter	0.006		0.007	0.024	0.021	0.022
MWH-UniP1 aquatic group	0.017	0.014	0.025	0.005	0.006	0.011
Methyloparacoccus	0.022	0.022	0.020			
Pseudarcicella	0.024	0.020	0.015			
Fluviicola				0.019	0.019	0.020
Dinghuibacter	0.015	0.014	0.015	0.007	0.007	
Rhodobacter			0.009	0.021	0.015	0.010
Prosthecobacter				0.020	0.014	0.015
Aphanizomenon MDT14a				0.009	0.024	0.014
CL500-3				0.020	0.012	0.010
Roseburia	0.006	0.008		0.008	0.018	
Dolichospermum NIES41				0.007	0.016	0.016
Pedosphaera				0.012	0.017	0.007
Pseudorhodoplanes				0.007		0.027
Desulfobacca				0.009	0.015	0.006
OM27 clade				0.009	0.009	0.012
Candidatus Planktoluna	0.011	0.011	0.008			
Algoriphagus	0.012	0.008	0.008			

invitrogen

Conclusions

With Platinum Direct PCR Universal Master Mix, we successfully identified bacterial genera directly from lake water samples. Direct PCR and PCR of extracted DNA give similar results, indicating high sensitivity and robustness of the direct PCR method. NGS results show that read frequencies were more broadly distributed in directly amplified samples than in extracted DNA samples, indicating that higher diversity can be identified with direct amplification. In addition, more bacterial genera were identified by direct amplification (34, vs. 29 amplified from extracted gDNA), suggesting that some sample material may have been lost during DNA extraction. Taken together, using Platinum Direct PCR Universal Master Mix in metagenomic analysis of water samples helps simplify the process and improves the chances of covering a broader range of targets.

References

1. Klindworth A, Pruesse E, Schweer T et al. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41(1):e1.
2. Quast C, Pruesse E, Yilmaz P et al. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41 (Database issue):D590–D596.
3. Yilmaz P, Parfrey LW, Yarza P et al. (2014) The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res* 42 (Database issue):D643–D648.
4. Glöckner FO, Yilmaz P, Quast C et al. (2017) 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J Biotechnol* 261:169–176.
5. Lu J, Breitwieser FP, Thielen P et al. (2017) Bracken: estimating species abundance in metagenomics data. *PeerJ Comput Sci* 3:e104.
6. Ortiz-Burgos S (2016) Shannon-Weaver Diversity Index. In: Kennish MJ (editor), *Encyclopedia of Estuaries. Encyclopedia of Earth Sciences Series.* Springer, Dordrecht.
7. Gaines WL, Harrod RJ, Lehmkuhl JF (1999) *Monitoring Biodiversity: Quantification and Interpretation.* U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.
8. NCBI SRA database. **https://www.ncbi.nlm.nih.gov/sra**. Accession number PRJNA607875.

Find out more at **thermofisher.com/platinum-direct-pcr**

ThermoFisher SCIENTIFIC

For Research Use Only. Not for use in diagnostic procedures. © 2020 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. Illumina and MiSeq are trademarks of Illumina, Inc. **COL23968 0320**