A protocol for sequencing and analysing 16S-ITS-23S amplicons using Oxford Nanopore's Native Barcoding kit to profile prokaryotes on species-level in a mixed community [Under development]

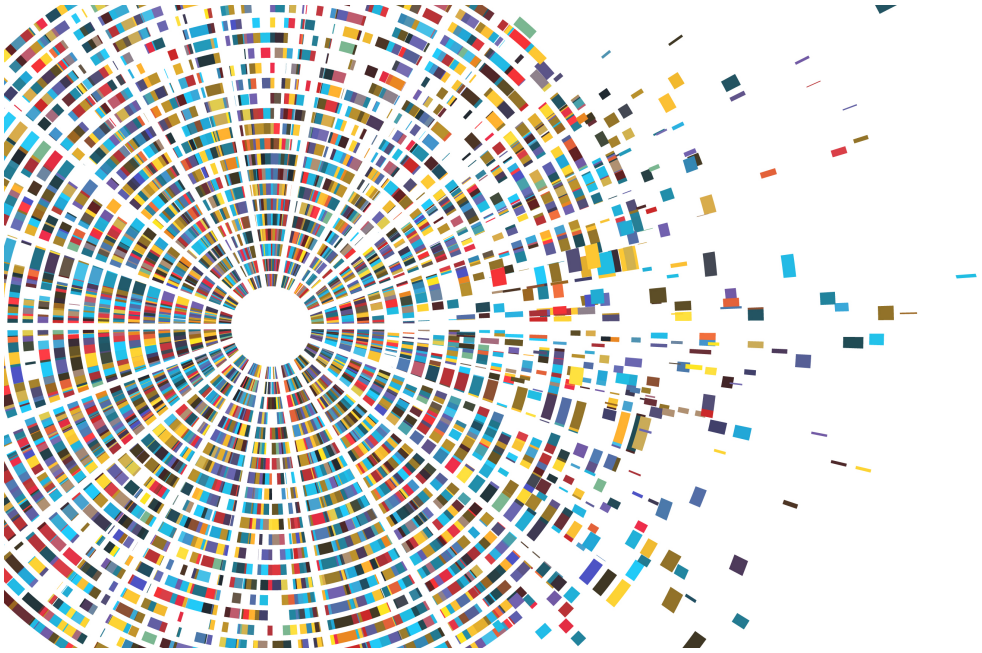
Christian Krohn, PhD, RMIT University

2024-06-17

Contents

Chapter 1

About this GitBook



This GitBook is under development (2024).

This GitBook provides a basic step-by-step protocol for students who want to get started with sequencing multiple samples of high quality, full-length microbial amplicons with an Oxford Nanopore instrument, and then process and visualise the data using various packages in R.

This protocol reflects the experiences we had in the lab and may help to get you started on your own journey of discovery. One of the biggest hurdles for students

that embark on sequencing environmental DNA for the first time, is the effort required to learn new technologies, various coding languages, file types and formats, packages or platforms that are involved (think Unix, Linux, Slurm, Qiime, R, RMarkdown, python, conda, ggplot, docker, GitHub, data instances…) before they even can start looking at exploring the data for biological meaning and producing publishable output. Hence, this protocol aims at lowering your 'activation' energy by providing some guidance for each step from DNA extraction, to amplification, library preparation, sequencing, data processing and finally some basic visualisations in R.

This protocol uses the DNeasy Powersoil Pro Kit (Qiagen, Hilden, Germany) to extract DNA from wastewater sludge, as well as the Native Barcoding Kit (SQK-NBD114.96) with a R10.4.1 flowcell (FLO-MIN114) to sequence long (~4-4.4 kb) amplicons from environmental DNA. The primer used was developed by (Martijn et al., 2019). Initial costs to purchase all consumables will be ~AUD$7,700.

Check out our workflows for short-read 16S sequencing at https://chrismitbiz.github.io/ABlab-workflows if you want to get started with Miseq 16S sequencing instead.

If you are interested in sequencing the living biomass via PMA treatment in combination with short-read 16S sequencing, check out our recent paper Dead in the water – Role of relic DNA and primer choice for targeted sequencing surveys of anaerobic sewage sludge intended for biological monitoring (Krohn et al., 2024)

Get in touch

We work at the Andy Ball lab, RMIT University, Bundoora, Melbourne and are part of the Industry-led Biosolids Training Centre. Email me or comment on the discussion section of the GitHub repository for this GitBook. You will need to get a GitHub account to join the discussion. Its free.

More about me and my PhD research can be found here: https://clean-dirt-digests.netlify.app.
Follow me on Twitter or LinkedIn.

Bio:
Dr Chris Krohn is an early career researcher whose interests could be summed up with: “Environmental sequencing, microbial ecology, chlorinated pollutants, organic matter, wastewater, anaerobic digestion, and how everything connects”.

In 2021 I joined the ARC Biosolids Training Centre at RMIT (www.transformingbiosolids.org.au), where we help water utilities to improve circular resource management by getting more renewable biogas and carbon/fertiliser values out of our municipal biosolids (essentially our poo). In project 1C of the Centre I develop metagenomic (DNA-based) methods to monitor the microbiome of anaerobic digestion, an important microbial treatment process for wastewater. I believe DNA-based diagnoses of wastewater sludges will help the water/biosolids sector improve resource recoveries and risk management.

Before that, after a career in one of the most fast-cycled and short-sighted manufacturing industries that took me from Germany to Vietnam and Hong Kong/Shenzhen, I decided to hit the switch and start thinking long-term and circular. Ten back-to-uni years later, in 2021 I finished a PhD in Soil Science at La Trobe Uni where I sequenced soil DNA and explored if and how soil biology was involved in the degradation of extremely persistent legacy pesticides that contaminate agricultural surface soils since several decades.

Chapter 2

Prerequisites and consumables

2.1 Prerequisites

- Experience with basic molecular laboratory techniques and DNA quality control (use of laminar flow hood, pipetting micro volumes, handling and preserving small volume reagents, sterilisation of consumables, doing gel electrophoresis)
- Experience with loading Nanopore flowcells (for example by doing a test run with Lambda DNA using the Control Expansion Kit)

- Confident with R, R Studio and package environment managers such as Conda and Docker. Knowledge in shell scripting and Linux syntax.

2.2 DNA extraction kit

We commonly extract DNA from soils or from wastewater sludges using the DNeasy Powersoil Pro Kit (Qiagen) for both, soil and sludge. It resulted in high quality DNA, suitable for library preparation and sequencing. Extraction with this kit includes a bead beating step to release DNA from cells out of difficult matrices such as soils and sludge (biofilms) or similar. The sheared DNA will be fragmented with a N50 of ~7 kb in some cases (Jensen et al., 2024). Shearing of DNA may be minimised by reducing/optimising bead-beating duration and speed. Other commonly used and commercially available kits for DNA extraction are available (Jensen et al., 2024; Gand et al., 2023).

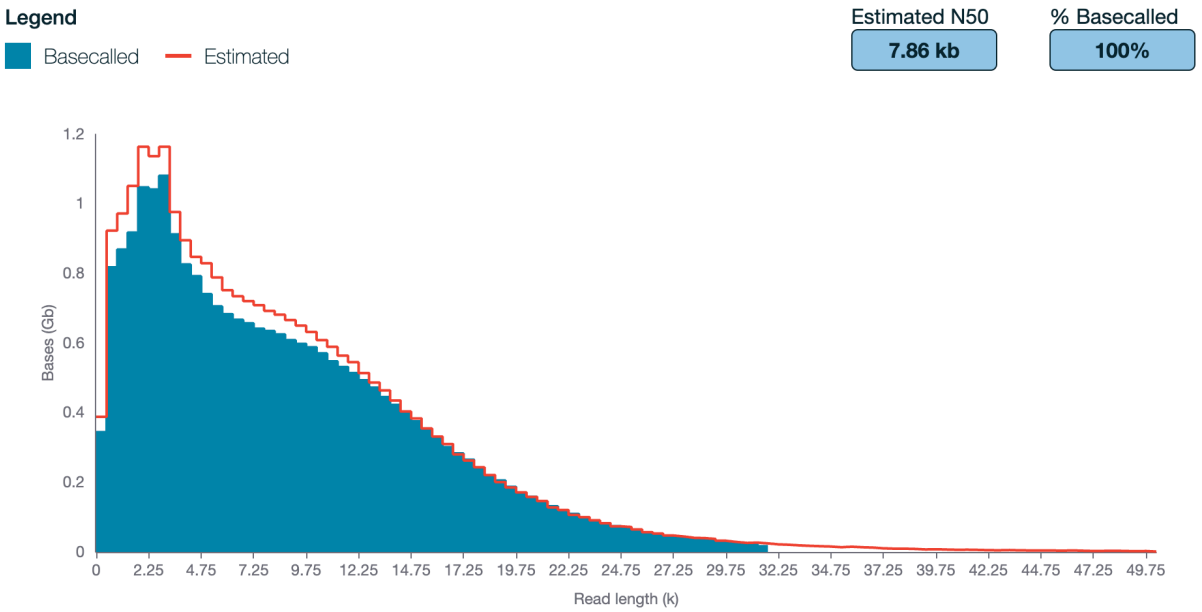


Figure 2.1: "Example of read size distribution of sludge DNA extracted with the DNeasy Powersoil Pro Kit in our lab"

2.3 Other consumables

- Native Barcoding Kit (Oxford Nanopore, SQK-NBD114.96)

- R10.4.1 flowcell (Oxford Nanopore, FLO-MIN114)
- NEBNext Quick Ligation Module (E6056, New England Biolab)
- NEBNext Ultra II End repair/dA-tailing Module (E7546S, New England Biolab)
- Blunt/TA Ligase Master Mix (M0367, New England Biolab)
- 10 mM dNTPs (N0447S, New England Biolab)
- Q5 Hot Start High-Fidelity DNA Polymerase (M0493, New England Biolab)
- DNeasy Powersoil Pro Kit (#47016, Qiagen)
- Qubit 1X dsDNA BR Assay Kit (Q33266, Thermo Fisher)
- Eppendorf LoBind tubes (Eppendorf)
- twin.tec® PCR plate 96 LoBind, semi skirted (0030129504, Eppendorf)
- JetSeq Clean Magnetic beads (MER-BIO-68031, Millenium Science) - to clean up PCR products
- 10 µM Forward Primer A519F (CAGCMGCCGCGGTAA) (Martijn et al., 2019)
- 10 µM Reverse Primer U2428R (CCRAMCTGTCTCACGACG) (Martijn et al., 2019)

- Nuclease-free water
- 10 mM Tris-HCl pH 8.0 with 50 mM NaCl (UltraPure™ 1M Tris-HCI, pH 8.0 #15568025 & NaCl (5 M), RNase-free #AM9760G, Thermo Fisher)
- Bovine Serum Albumin (BSA) (50 mg/ml) (AM2616, InvitrogenTM UltraPure)
- HyperLadder 1kb (BIO-33025, Bioline, Millenium Science)
- 80% ethanol, freshly prepared in nuclease-free water

2.4 Protocols

- A protocol for PCR amplification and clean-up (included in this Gitbook).
- 'Ligation sequencing amplicons - Native Barcoding Kit 96 V14' available on nanoporetech.com - Checklist and full protocol.
- Prepare spreadsheets to normalise DNA into 96-well plates to avoid pipetting errors.

2.5 Equipment

- Nanopore sequencer. This protocol is based on the MinION Mk1C but is not limited to this model.
- Nanodrop spectrophotometer to check DNA quality.
- Qubit fluorometer (Thermo Fisher).
- PCR thermal cycler.
- Gel electrophoresis equipment.
- Vortex. We use the Vortex-Genie 2, including the 24 x 2 ml tube adapter for bead beating.

2.6 Computational resources and database

- An instance or computer with 2 TB storage, 64 GB RAM and 32 CPUs
- 16S-ITS-23S Database. To be confirmed.

Chapter 3

Protocol

3.1 DNA extraction

3.1.1 Checklist

- Extraction kit - DNeasy Powersoil Pro, Qiagen, Hilden, Germany
- Alternatively the high throughput version of the kit used with the QIAcube DNeasy 96 PowerSoil Pro QIAcube HT Kit
- Vortex with 24 x 1.5 mL tube adapter (e.g. Vortex Genie 2 + adapter). Alternatively, a PowerLyzer Homogenizer.
- NanoDrop spectrophotometer to assess DNA quality.
- Qubit fluorometer for accurately measure DNA concentrations.

3.1.2 Process

- Follow the extraction kit's protocol with 15 mins bead beating with Genie vortex and 24-tube adapter. Reduce this to 10 mins if less than 24 samples are extracted or if shearing of DNA should be minimised.
- Measure DNA quality using DNA extract (1 µL) using a Nanodrop spectrophotometer
- Measure DNA concentration using a Qubit fluorometer.
- Into a 96-well plate, normalise extracted DNA to required PCR concentrations. For example, if 10 ng template is required for PCR, normalise DNA to 5 ng/µL.

- Store DNA at 4 ° C until library preparation - no more than 1 week.
- Store DNA at -20 ° C if sequencing is more than 1 week away.

3.2 Amplification of 16S-ITS-23 operon

3.2.1 Checklist

- 10 mM dNTPs (N0447S, New England Biolab)
- Q5 Hot Start High-Fidelity DNA Polymerase (M0493, New England Biolab)
- 10 µM Forward Primer A519F (CAGCMGCCGCGGTAA) (Martijn et al., 2019)
- 10 µM Reverse Primer U2428R (CCRAMCTGTCTCACGACG) (Martijn et al., 2019)
- JetSeq Clean Magnetic beads - or equivalent (MER-BIO-68031, Millenium Science) - to clean up PCR products
- twin.tec® PCR plate 96 LoBind, semi skirted (0030129504, Eppendorf)
- Nuclease-free water
- 10 mM Tris-HCl pH 8.0 with 50 mM NaCl (UltraPure™ 1M Tris-HCI, pH 8.0 #15568025 & NaCl (5 M), RNase-free #AM9760G, Thermo Fisher)
- HyperLadder 1kb (BIO-33025, Bioline, Millenium Science)
- 80% ethanol, freshly prepared in nuclease-free water
- Qubit 1X dsDNA BR Assay Kit (Q33266, Thermo Fisher)

Notes

- Do not vortex tubes during library preparation and use wide-bore tips to prevent DNA fragmentation. Fragmentation of amplicons may lead to incomplete reads.
- The primer amplifies the whole rrn operon.

Benefits of targeting the whole rrn operon:

- Superior species-level resolution and accuracy (Cuscó et al., 2019; Srinivas et al., 2024).
- Covers Bacteria and Archaea.

Risks of targeting the whole rrn operon:

- Not as representative of true abundances as full-length 16S amplicons.
- Species with unlinked 16S and 23S rrn DNA will be missed with this approach (for example $< 9\%$ of rRNA genes in wastewater sludge (Brewer et al., 2020)).

3.2.2 PCR

Time required ~4 hrs incl. 2hrs, 40mins PCR.

- Prepare a mastermix for required number of 50 µl reactions (Table 3.1).
- Add 3 µl of eDNA (5ng /µl) into a 96-well plate (e.g. Eppendorf twin.tec® PCR plate 96 LoBind, semi-skirted) using a multichannel pipette.

- Add 47 µl of mastermix using a multichannel pipette and carefully pipette up and down 10 x
- Run thermocycler (Table 3.2).
- Verify amplification length via 1% agarose gel electrophoresis @ 100V for 30 min including a 1 kb ladder.
- Store at 4 ˚ C overnight if needed.

Important:

- Use hot start polymerase for ease of use, with high fidelity/accuracy and one that is suitable for long amplicons. For example, Q5 High- Fidelity DNA Polymerase kit (New England Biolabs) with GC enhancer (Martijn et al., 2019).

- 200 fmol is required per sample for the Native Barcoding Kit from ONP. Based on 4.25 kb (4-4.5kb), the final DNA concentration after cleaning up PCR products should be no less than 48 ng/µL (at 11.5 µl input volume) - giving 552 ng of 4.25kb amplicons.

- It may require two–three PCR reactions to achieve the required DNA amount (200 fmol); e.g. pool 2–3 x 50µl PCR products, clean combined and elude in 32.5 µl Tris.

Table 3.1: List of components for each reaction (each tube) for PCR. See https://www.neb.com/en-au/protocols/2012/08/30/pcr-using-q5-hot-start-high-fidelity-dna-polymerase-m0493 for details.

Component	50 µl reaction	Final concentration
5X Q5 Reaction Buffer - M0493S NEB	10 µL	1X
10 mM dNTPs N0447S	1 µL	200 µM
10 µM Forward Primer	2.5 µL	0.5 µM
10 µM Reverse Primer	2.5 µl	0.5 µM
Template DNA - 15 ng	3 µL (5 ng/µL)	$<$ 1,000 ng
Q5 Hot Start High-Fidelity DNA Polymerase -M0493S NEB	1.0 µL	0.04 U/µL
5X Q5 High GC Enhancer M0493S NEB	10 µL	(1X)
Nuclease-Free Water	to 50 µL	

Table 3.2: Thermocycler conditions (Martijn et al 2019).

Cycle conditions
1 cycle:
30 s - Initial Denaturation 98 degree C
35 cycles:
10 s - 98 degree C
30 s - 64 degree C
210 s - 72 degree C
1 cycle:
10 min - Final Extension 72 degree C

Bibliography

Brewer, T. E., Albertsen, M., Edwards, A., Kirkegaard, R. H., Rocha, E. P., and Fierer, N. (2020). Unlinked rRNA genes are widespread among bacteria and archaea. *ISME Journal*, 14(2):597–608.

Cuscó, A., Catozzi, C., Viñes, J., Sanchez, A., and Francino, O. (2019). Microbiota profiling with long amplicons using Nanopore sequencing: Full-length 16S rRNA gene and the 16S-ITS-23S of the rrn operon. *F1000Research*, 7.

Gand, M., Bloemen, B., Vanneste, K., Roosens, N. H., and De Keersmaecker, S. C. (2023). Comparison of 6 DNA extraction methods for isolation of high yield of high molecular weight DNA suitable for shotgun metagenomics Nanopore sequencing to detect bacteria. *BMC Genomics*, 24(1):438.

Jensen, T. B. N., Dall, S. M., Knutsson, S., Karst, S. M., and Albertsen, M. (2024). High-throughput DNA extraction and cost-effective miniaturized metagenome and amplicon library preparation of soil samples for DNA sequencing. *PLoS ONE*, 19(4 April):e0301446.

Krohn, C., Jansriphibul, K., Dias, D. A., Rees, C. A., van den Akker, B., Boer, J. C., Plebanski, M., Surapaneni, A., O'Carroll, D., Stuetz, R., Batstone, D. J., and Ball, A. S. (2024). Dead in the water – Role of relic DNA and primer choice for targeted sequencing surveys of anaerobic sewage sludge intended for biological monitoring. *Water Research*, 253:121354.

Martijn, J., Lind, A. E., Schön, M. E., Spiertz, I., Juzokaite, L., Bunikis, I., Pettersson, O. V., and Ettema, T. J. (2019). Confident phylogenetic identification of uncultured prokaryotes through long read amplicon sequencing of the 16S-ITS-23S rRNA operon. *Environmental Microbiology*, 21(7):2485–2498.

Srinivas, M., Walsh, C. J., Crispie, F., O'Sullivan, O., Cotter, P. D., Sinderen, D., and Kenny, J. G. (2024). Evaluating the Efficiency of 16S-ITS-23S operon Sequencing: A Comparison of Primer Pairs, Sequencing Platforms, and Taxonomic Classifiers.