Molecular Standard Operating Procedure (MSOP)

for

Ocean Sampling Day (OSD)

2018 - 2019 campaigns

Version

1.0

Date

2022-06-30

Purpose

This document contains the Standard Operating Procedures for working with the molecular data of the OSD (https://www.assembleplus.eu/research/ocean-sampling-day) project. The samples containing the material are sent by each observatory to HCMR for processing (see the Handbook and the Protocol for details).

DNA Extraction

This protocol is used for extracting DNA from seawater filtered through 0.22 um pore size Sterivex™ Filter Units (Merck).

Materials:

- Sterivex™ Filter(s) stored at -80 C
- DNA-extraction kit (DNeasy PowerWater Kit, Cat. No. / ID: 14900-100-NF, Qiagen)
- Sterile pipettes and pipette tips
- DNA-decontaminating solution - 10% HCL
- TissueLyser II, Cat. No. / ID: 85300, Qiagen
- Centrifuge
- Nanodrop spectrophotometer
- PVC tube cutter tool, to cut open sterivex cartridge
- Sterile blades, to cut the filter
- Sterile forceps to handle the filter

Extraction protocol:

1. Wear gloves at all times. Carefully clean the bench station(s) and pipettes.
2. Proceed to DNA extraction using the DNA-extraction kit, as recommended by the manufacturer. Extract each replicate sample separately. Use two replicates (two Sterivex filters) per site.
3. Also do a blank extraction using the extraction buffer in place of sample, and process this as a regular sample when later constructing the amplicon libraries.
4. Quantify the extracted DNA using a spectrophotometer.
5. Store the DNA at -20 ºC until further processing.

Procedure for metagenome shotgun libraries and sequencing

Approximately 300 ng of extracted DNA was sent to Genoscope (Evry, France) on dry ice. About 250 ng of DNA was fragmented mechanically in a Covaris ultrasonicator. A homemade protocol with PCR, with enzymes and buffers from NEB, was used to create libraries with average insert sizes of 376 bp and a single 12 bp index. The libraries were quantified and pooled to be equimolar, before 2 x 150 bp sequencing on three lanes on a NovaSeq S4 flow cell. The generated data was around 19 Gbases per sample. The following trimming was performed by Genoscope: Adapters and primers were trimmed. Low quality nucleotides were trimmed from both ends (quality value lower than 20). The sequence after two consecutive unknown nucleotides (N) was trimmed. Reads shorter than 30 nucleotides after trimming were discarded. Reads and their mates that mapped onto run quality control sequences (PhiX genome) were removed.

Procedure for amplicon libraries and sequencing

The protocol for amplicon library construction is based on Illumina's two-step PCR protocol for 16s library construction "16S Metagenomic Sequencing Library Preparation", with a few minor changes incorporated.

Materials:

- Primers

Primer sequences and expected amplicon sizes. The sizes are without the 5'tail (below).

Primer Name	Sequence	Target gene	Target group	Amplicon size (bp)	Reference
All18 SF	5′-TGGTGCATGGCCGTTCTTAGT-3′	18S rRNA	metazoa, fungi, protozoa plants	180-500	Hardy et al. 2010
All18SR	5′-CATCTAAGGGCATCACAGACC-3′				
341F1	5′-CTACGGGNGGCWGCAG-3′	16s v3-v4	Bacteria/ Archaea	438-466	Klindworth et al, 2013
805RB1	5′-GACTACNVGGGTATCTAATCC-3′				Pavloudi et al 2017

- 5′ tails used for the first-step PCR

Primer Name	Sequence
1st_PCR_for	5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[locus-specific sequence] -3′
1st_PCR_rev	5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[locus-specific sequence] -3′

- Extracted DNA
- KAPA HiFi HotStart Readymix Kit (for 18s)
- KAPA HiFi Hotstart PCR kit (for 16s)
- Qubit® dsDNA BR Assay Kit or Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher/Invitrogen)
- NucleoMag® NGS Clean-up and Size Select (Macherey-Nagel)
- Thermal cycler
- Dedicated pipettes and pipette tips
- PCR reaction tubes and/or plates
- PCR grade water
- 1 M Trehalose
- 1.5 % agarose gel with Ethidium Bromide or other fluorescent dye, loading buffer (1x TAE), loading dye (6x Orange G)
- Agilent Bioanalyzer
- NEBNext library quant kit for Illumina
- MiSeq Reagent kit v3 (600 cycles)

1. Wear gloves at all times. Carefully clean the bench station(s) and pipettes.
2. PCR amplification is performed targeting two gene regions: 18S rRNA and 16s rRNA
3. Include PCR-blanks in each PCR, using water instead of extracted DNA as template.

4. The first-step PCR is performed with the aforementioned primers containing a universal 5′ tail as specified in the 16s library protocol from Illumina.
    - First-step PCR for 18S rRNA:
      Amplification reaction mix:
      12.5 ul 2x KAPA HiFi Hotstart Readymix
      0.75 ul from the forward primer (10 uM)
      0.75 ul from the reverse primer (10 uM)
      Template DNA plus water until a final volume of 25 ul per reaction.

      PCR conditions:
      95 °C for 3 min;
      30 cycles at 98 °C for 20 s, 58 °C for 15 s, 72 °C for 15 s;
      72 °C for 3 min

    - First-step PCR for 16s:
      Amplification reaction mix:
      5 ul Fidelity buffer
      0.75 ul dNTPs
      0.75 ul of the forward primer (10 uM)
      0.75 ul of the reverse primer (10 uM)
      0.5 ul KAPA HiFi Hotstart polymerase
      5 ul 1M Trehalose
      Template DNA plus water until a final volume of 25 ul per reaction.

      PCR conditions:
      95 °C for 5 min;
      26 cycles at 98 °C for 20 s, 57 °C for 120 s, 72 °C for 60 s;
      72 °C for 7 min

5. Purify the resulting PCR amplicons using Nucleomag magnetic beads, at a ratio 0.8:1 (magnetic beads: PCR product). After the second ethanol wash, make sure to completely remove all the remaining ethanol with a pipette. Elute in 25 ul 10 mM Tris-Cl, pH 8.5
6. Run 2 ul of the cleaned PCR product on a 1.5 % agarose gel with a suitable ladder. Even if no PCR product is visible, proceed with the second-step PCR.
7. Use the purified PCR amplicons as templates for the second-step PCR in order to include the indexes (barcodes), as well as the Illumina adaptors. A different set of indexed primers should be used for each sample, but the same indexes may be used for different loci.

Indexed forward primers for the second-step PCR. The index sequence is shown in blue.

Primer Name	Sequence	Index name
NGS_i5_S502	5′-AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTC-3′	S502
NGS_i5_S503	5′-AATGATACGGCGACCACCGAGATCTACACTATCCTCTTC GTCGGCAGCGTC-3′	S503
NGS_i5_S504	5′-AATGATACGGCGACCACCGAGATCTACACAGAGTAGATCGTCGGCAGCGTC-3′	S504
NGS_i5_S505	5′-AATGATACGGCGACCACCGAGATCTACACGTAAGGAGT CGTCGGCAGCGTC-3′	S505
NGS_i5_S506	5′-AATGATACGGCGACCACCGAGATCTACACACTGCATAT CGTCGGCAGCGTC-3′	S506
NGS_i5_S507	5′-AATGATACGGCGACCACCGAGATCTACACAAGGAGTAT CGTCGGCAGCGTC-3′	S507
NGS_i5_S508	5′-AATGATACGGCGACCACCGAGATCTACACCTAAGCCTT CGTCGGCAGCGTC-3′	S508
NGS_i5_S510	5′-AATGATACGGCGACCACCGAGATCTACACCGTCTAATTC GTCGGCAGCGTC-3	S510
NGS_i5_S511	5′-AATGATACGGCGACCACCGAGATCTACACTCTCTCCGTC GTCGGCAGCGTC-3′	S511
NGS_i5_S513	5′-AATGATACGGCGACCACCGAGATCTACACTCGACTAGT CGTCGGCAGCGTC-3′	S513
NGS_i5_S515	5′-AATGATACGGCGACCACCGAGATCTACACTTCTAGCTTC GTCGGCAGCGTC-3′	S515
NGS_i5_S516	5′-AATGATACGGCGACCACCGAGATCTACACCCTAGAGTT CGTCGGCAGCGTC-3′	S516
NGS_i5_S517	5′-AATGATACGGCGACCACCGAGATCTACACGCGTAAGAT CGTCGGCAGCGTC-3′	S517
NGS_i5_S518	5′-AATGATACGGCGACCACCGAGATCTACACCTATTAAGT CGTCGGCAGCGTC-3′	S518
NGS_i5_S520	5′-AATGATACGGCGACCACCGAGATCTACACAAGGCTATT CGTCGGCAGCGTC-3′	S520
NGS_i5_S521	5′-AATGATACGGCGACCACCGAGATCTACACGAGCCTTAT CGTCGGCAGCGTC-3′	S521
NGS_i5_S522	5′-AATGATACGGCGACCACCGAGATCTACACTTATGCGAT CGTCGGCAGCGTC-3′	S522

Indexed reverse primers for the second-step PCR. The index sequence is shown in blue.

Primer Name	Sequence	Index name
NGS_i7_N701	5′-CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCG TGGGCTCGG-3′	N701
NGS_i7_N702	5′-CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCG TGGGCTCGG-3′	N702
NGS_i7_N703	5′-CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCG TGGGCTCGG-3′	N703
NGS_i7_N704	5′-CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCG TGGGCTCGG-3′	N704
NGS_i7_N705	5′-CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCG TGGGCTCGG-3′	N705
NGS_i7_N706	5′-CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCG TGGGCTCGG-3′	N706
NGS_i7_N707	5′-CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTC GTGGGCTCGG-3′	N707
NGS_i7_N710	5′-CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCG TGGGCTCGG-3′	N710
NGS_i7_N711	5′-CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCG TGGGCTCGG-3′	N711
NGS_i7_N712	5′-CAAGCAGAAGACGGCATACGAGATTCCTCTACGTCTCG TGGGCTCGG-3′	N712
NGS_i7_N714	5′-CAAGCAGAAGACGGCATACGAGATTCATGAGCGTCTCG TGGGCTCGG-3′	N714

NGS_i7_N715	5′-CAAGCAGAAGACGGCATACGAGATCCTGAGATGTCTCG TGGGCTCGG-3′	N715
NGS_i7_N716	5′-CAAGCAGAAGACGGCATACGAGATTAGCGAGTGTCTCG TGGGCTCGG-3′	N716
NGS_i7_N718	5′-CAAGCAGAAGACGGCATACGAGATGTAGCTCCGTCTCG TGGGCTCGG-3′	N718
NGS_i7_N719	5′-CAAGCAGAAGACGGCATACGAGATTACTACGCGTCTCG TGGGCTCGG-3′	N719
NGS_i7_N720	5′-CAAGCAGAAGACGGCATACGAGATAGGCTCCGGTCTCG TGGGCTCGG-3′	N720
NGS_i7_N721	5′-CAAGCAGAAGACGGCATACGAGATGCAGCGTAGTCTCG TGGGCTCGG-3′	N721
NGS_i7_N722	5′-CAAGCAGAAGACGGCATACGAGATCTGCGCATGTCTCG TGGGCTCGG-3′	N722
NGS_i7_N723	5′-CAAGCAGAAGACGGCATACGAGATGAGCGCTAGTCTCG TGGGCTCGG-3′	N723
NGS_i7_N724	5′-CAAGCAGAAGACGGCATACGAGATCGCTCAGTGTCTCG TGGGCTCGG-3′	N724
NGS_i7_N726	5′-CAAGCAGAAGACGGCATACGAGATGTCTTAGGGTCTCG TGGGCTCGG-3′	N726
NGS_i7_N727	5′-CAAGCAGAAGACGGCATACGAGATACTGATCGGTCTCG TGGGCTCGG-3′	N727
NGS_i7_N728	5′-CAAGCAGAAGACGGCATACGAGATTAGCTGCAGTCTCG TGGGCTCGG-3′	N728
NGS_i7_N729	5′-CAAGCAGAAGACGGCATACGAGATGACGTCGAGTCTCG TGGGCTCGG-3′	N729

8. The amplification reaction mix of the second-step 18s PCR contains:
   10 ul 2x KAPA HiFi Hotstart Readymix
   2 ul indexed forward primer (10 uM)
   2 ul indexed reverse primer (10 uM)
   DNA plus water until a final volume of 20 ul per reaction.
9. The amplification reaction mix of the second-step 16s PCR contains:
   5 ul Fidelity Buffer
   0.75 ul dNTPs
   0.75 ul indexed forward primer (10 uM)
   0.75 ul indexed reverse primer (10 uM)
   0.5 ul KAPA HiFi Hotstart polymerase
   5 ul 1M Trehalose
   DNA plus water until a final volume of 25 ul per reaction.
10. PCR conditions:
    95 °C for 3 min
    8 cycles at 98 °C for 20 s, 55 °C for 30 s, 72 °C for 30 s
    72 °C for 5 min
11. Purify the resulting PCR amplicons using magnetic beads, at a ratio 0.75:1 (magnetic beads: PCR product). After the second ethanol wash, make sure to completely remove all the remaining ethanol with a pipette. Elute in 25 ul 10 mM Tris-Cl, pH 8.5
12. Run 2 ul of the cleaned PCR product on a 1.5 % agarose gel with a suitable ladder. If bands are visible below 200 bp it indicates that adapter dimers are present. Repeat the bead purification for samples with such bands.
13. Quantify the amplicons either with Qubit® dsDNA BR Assay Kit or Quant-iT PicoGreen dsDNA Assay Kit.
14. Convert the concentration of the PCR amplicons from ng/ul to nM using the equation (1000000*Concentration in ng/µl)/(Total amplicon length in bp*660).

15. Create an amplicon sequencing pool by mixing the PCR amplicons in equimolar amounts (at a desired final concentration of 10 nM). Include also the extraction and PCR blanks.
16. Run a 1:5 dilution of each pool on a DNA High Sensitivity Chip on an Agilent Bioanalyzer to check for residual adapter dimer and to estimate library size and molarity. If adapter dimers are present, do another magnetic bead purification on the pool.
17. Quantify the amplicon pools using the NEBNext library quant kit for Illumina and dilute to the desired concentration, according to the Illumina MiSeq sequencing protocol.
18. Denature the libraries according to Illumina's protocol and add around 20 % PhiX to increase the sequence diversity. Sequence the amplicon pool using a MiSeq Reagent Kit v3 (2×300-cycles).
19. Submit the raw sequence files to the European Nucleotide Archive (ENA) (Amid et al. 2019).

References

Amid, C., Alako, B.T.F., Balavenkataraman Kadhirvelu, V., Burdett, T., Burgin, J., Fan, J., Harrison, P.W., Holt, S., Hussein, A., Ivanov, E., Jayathilaka, S., Kay, S., Keane, T., Leinonen, R., Liu, X., Martinez-Villacorta, J., Milano, A., Pakseresht, A., Rahman, N., Rajan, J., Reddy, K., Richards, E., Smirnov, D., Sokolov, A.,Vijayaraja, S., and Cochrane, G. 2019. The European Nucleotide Archive in 2019. Nucleic Acids Res. doi:10.1093/nar/gkz1063.

Hardy, C.M., Krull, E.S., Hartley, D.M., and Oliver, R.L. 2010. Carbon source accounting for fish using combined DNA and stable isotope analyses in a regulated lowland river weir pool. Mol. Ecol. **19**(1): 197–212. doi:10.1111/j.1365-294X.2009.04411.x.

Illumina Inc 2013. 16S Metagenomic Sequencing Library Preparation. https://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 2013 Jan 7;41(1):e1. doi: 10.1093/nar/gks808. Epub 2012 Aug 28. PMID: 22933715; PMCID: PMC3592464.

Pavloudi, C, Kristoffersen, JB, Oulas, A, De Troch, M & Arvanitidis, C 2017. Sediment microbial taxonomic and functional diversity in a natural salinity gradient challenge Remane's "species minimum" concept. PeerJ, DOI 10.7717/peerj.3687