May 09, 2022

# 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing

Estelle Bigeard[1], Adriana Lopes Dos Santos[2], Catherine Ribeiro[3]

[1]AD2M, Station Biologique de Roscoff, CNRS, SU; [2]Asian School of Environment - NTU; [3]Station Biologique de Roscoff



dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

Ecology of Marine Plankton (ECOMAP) team - Roscoff | Roscoff Culture Collection 1

Estelle Bigeard
AD2M, Station Biologique de Roscoff, CNRS, SU

For metabarcoding purpose, the first step involves the amplification by PCR of a given gene region (for example V4 or V9 region of 18S rRNA gene) or gene itself if its size does not exceed 600bp (the longest fragment size that can be sequenced by Illumina technology).

The defined forward and reverse primers that are complementary upstream and downstream of the region of interest, needs to be designed with overhang adapters which will be used in a subsequent limited-cycle amplification step, in order to add the dual-index barcodes and Illumina flow cell adapters. To design illumina primers, it will be necessary to know the sequencing method, and therefore the adapters sequence.

The following protocol explains the generation of 18S-V4 PCR amplicons for eukaryote metabarcoding analyses from sea water samples.

At first we developed this protocol for different project using equipements of Genomer sequencing platform (Roscoff, France).
More recently, for the Peacetime project, this protocol was adapted to generate libraries with Get-PlaGE sequencing service (Toulouse, France).

DOI

dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

Estelle Bigeard, Adriana Lopes Dos Santos, Catherine Ribeiro 2022. 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing. **protocols.io** https://dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

protocol

Julie Dinasquet, Estelle Bigeard, Frédéric Gazeau, Farooq Azam, Cécile Guieu, et al.. Impact of dust addition on the microbial food web under present and future conditions of pH and temperature. 2021. ⟨hal-03260571⟩ DOI : 10.5194/bg-2021-143

gene amplification, metabarcoding, NGS, sequencing, eukaryote, genomic DNA

protocol ,

Aug 31, 2020

May 09, 2022

41260

Stoeck, T. et al. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Molecular Ecology, 19 (Suppl. 1), 21–31 doi: 10.1111/j.1365-294X.2009.04480.x (2010)
Piredda, R et al. Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. FEMS Microbiology Ecology 93:fiw200. DOI: 10.1093/femsec/fiw200 (2017)



**Citation:** Estelle Bigeard, Adriana Lopes Dos Santos, Catherine Ribeiro 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing https://dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

• **Equipments**

PCR hood (UV)
Thermocycler
Spark & Nanoquant Plate (Tecan)
Qubit 4 fluorometer (Invitrogen)
Microwave
Electrophoresis system & generator (BioRad)
Image Quant LAZ4000 (Ge Healthcare)

• **PCR amplification**

Primers Quality SePOP in water, stock 100µM - Eurogentec
Primers with Illumina tail Quality NGS Purification RP-HPLC, dried - Eurogentec

Nuclease-Free Water - Invitrogen™Ambion™ - Ref AM9937
Phusion® High-Fidelity PCR Master Mix with HF Buffer - New England Bioloabs - Ref M0531S
BSA - New England Biolabs - Ref B9000S
KAPA HiFi HotStart ReadyMix® - Roche Diagnostic France - Ref KK2602

nuclease free 96-wells PCR plates
1.5ml microtubes

• **Electrophoresis**

*50X TAE Buffer :*

- *Tris-Base (2M final) 242.5 g*
- *Acetic acid (>99.7%, 2M final) 7 ml*
- *0,5 M EDTA pH 8,0 (50 mM final) 100 ml*

*dH20 up to 1L*

Tris-Base - Sigma Aldrich - Ref T6791
Acetic acid - Thermo Fisher Scientific - Ref 11463473
0,5 M EDTA pH 8,0 - Fisher Scientific - Ref 15575-020
agarose - Interchim - Ref 31272L
Ethidium bromide - Sigma Aldrich - Ref E1510
SmartLadder - 200 to 10000 bp - Eurogentec - Ref MW-1700-10

## Samples : Acid nucleic extraction and quantification

1 Extract DNA/RNA from the samples using the method/kit appropriate to the specific samples.

Filtered environmental samples usually have more DNA than sorted cells samples, so the V4 region of the 18S rRNA gene can be amplified by a single PCR reaction (without the need of a nested protocol).

When working with RNA, another step to transform the RNA in cDNA is necessary.

Quantify and quality-check the final DNA via NanoDrop or Qubit/PicoGreen.



**Citation:** Estelle Bigeard, Adriana Lopes Dos Santos, Catherine Ribeiro 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing https://dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

The A280/260 ratios should be 1.8 or higher and once working with total community extracted DNAs, a concentration of at least 1 ng/µL is desirable to get consistent PCR results. Avoid secondary extractions or clean-ups for inhibitors, until PCRs have truly shown inhibition (example after trying the PCR by diluting the original DNA and/or adding BSA to the final PCR reaction) – many suboptimal DNAs will still work in PCR. A gel can be run to verify integrity of the extracted material, but it is generally unnecessary for PCR-only studies **(Figure 1)**. Some protocols suggest the normalization of the DNA samples concentration prior to the PCR. In our experience, this is not always possible given the low concentration of some samples. The volume of sample added to the PCR can be settled in function of a range of concentrations instead of normalizing all the samples to the same concentration.

Load in agarose gel 0.8%, 1µl DNA + loading buffer. The limit of DNA detection using EtBr gels is around 10ng.

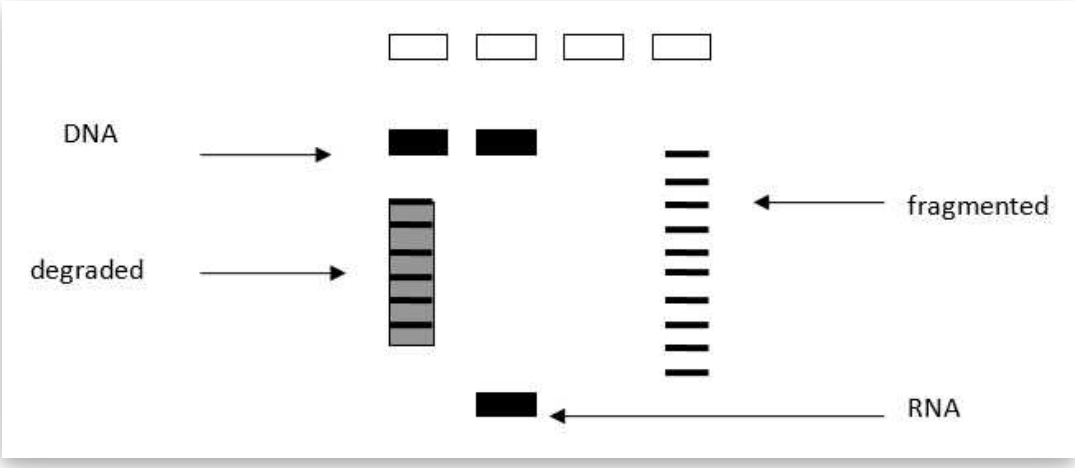


Figure 1: DNA/RNA detection in agarose gel.

Each DNA sample have to be normalized to 5ng/µl for PCR. For that, dilute DNA in nuclease free water.

It is not necessary to check the DNA concentration after dilution.

Samples with a DNA concentration less than 5ng/µl will be amplified as is in PCR.

## PCR amplification

2 Below is a suggestion of protocols for 18S amplification from total genomic DNA. This protocol has been tested and used with success.

However, before starting the library preparation, few samples from the batch (es) to be analyzed should be selected and the PCR conditions should be tested. The tests can be performed by using the primers without the illumina “tail” which are less expensive since it can be order by normal purification method, like Se-POP. By using also these primers, it avoids excessive manipulation of the primers with the illumina tail and therefore the chances of contamination. Low number of cycles, pool multiple (i.e., triplicate) PCRs for each sample, high initial time of denaturation are fundamental practices for the PCR when targeting diversity studies and should be considered for all reactions, including the tests in order to reduce or avoid PCR bias caused by



**Citation:** Estelle Bigeard, Adriana Lopes Dos Santos, Catherine Ribeiro 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing https://dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

PCR selection and PCR drift.

• **PCR selection** includes all mechanisms which inherently favor the amplification of certain templates due to properties of the genes and/or of their flanking sequences of the target region. Potentially important contributors to PCR selection are preferential denaturation due to overall low GC content, higher binding efficiency of GC-rich permutations of degenerate primers, differential accessibility of rRNA genes within genomes, and correlation between amplification probabilities and gene copy numbers within genomes.

• **PCR drift** is cause by stochastic variation in the early cycles of the reaction (when amplification still proceeds largely from the genomic templates), and its outcome should therefore not be reproducible in replicate PCR amplifications.

More details about these biases can be found in the following references:
-Aird, D. et al. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. Genome Biol. 12, R18 (2011).
-Wagner, A. et al. Surveys of Gene Families Using Polymerase Chain Reaction: PCR Selection and PCR Drift. Syst. Biol. 43, 250 (1994).-Polz, M. F. & Cavanaugh, C. M. Bias in template-to-product ratios in multitemplate PCR. Appl. Environ. Microbiol. 64, 3724–3730 (1998).

## Primers used for DNA amplification and sequencing

2.1 Use the existing primers (Table 1 & Table 2) or design your own custom gene primers with the proper Illumina indices and Index adaptor orientations. These adapters will be different in function of the libraries preparation protocol, and therefore in function of sequencing platform (Table 2).

PCR amplification of the V4 region of the 18S rDNA (~380 bp) using primers TAReuk454FWD1 (V4F) and V4 18S Next.Rev (TAReukREV3_modified or V4R_Zig) allow to amplify most eukaryotic groups and to solve the problem of the under-representation of Haptophytes.

Natural samples usually have more DNA than sorted cells samples, so the V4 region of the 18S rRNA gene can be amplified by a single PCR reaction (without the need of a nested protocol).

It is recommended to order the illumina primers with a good synthesis and purity parameters.

We have ordered at EUROGENTEC the Illumina primers with RP-HPLC purification method, 10 or 40nmol synthesis scale and at 100µM TE concentration. We could also ordered in dried and dissolve them with nuclease free water or TE 1x buffer to obtain 100µM stock solution (as described on the Eurogentec's technical data sheet).
Then we dilute them to 10 µM working concentration (1/10th the typical 100 µM working stock concentration for primers). This proceed should be performed at the PCR HEPA filter hood and with molecular grade water.
The primers (stock and working solutions) are stored at -20°C.



**Citation:** Estelle Bigeard, Adriana Lopes Dos Santos, Catherine Ribeiro 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing https://dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

| Target gene | Primer | Product size | Sequence (5'-3') | Reference |
|---|---|---|---|---|
| 18S-V4 | V4F | 380 | CCA GCA SCY GCG GTA ATT CC | Stoeck et al. 2010 |
| | V4R_Zig | | ACT TTC GTT CTT GAT YRA TGA | Piradda et al. 2017 |

Table 1: Primers used for DNA amplification and sequencing

| Primer | Product size with the tail | Sequencing by: | Final Primer (Illumina tail) |
|---|---|---|---|
| V4F_illumina | 527 | Genomer (Roscoff) | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG CCAGCASCYGCGGTAATTCC |
| V4R_Zig _illumina | | | GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG ACTTTCGTTCTTGATYRATGA |
| V4F_illumina | 527 | Genotoul | CTTTCCCTACACGACGCTCTTCCGATCT CCAGCASCYGCGGTAATTCC |
| V4R_Zig _illumina | | | GGGAGTTCAGACGTGTGCTCTTCCGATCT ACTTTCGTTCTTGATYRATGA |

Table 2: 18S final primers with Illumina tail

## Controles

2.2 Test a negative control for each series. The negative PCR product could be inserted in the library to have a control of air contaminations.

It is possible to test a positive control (mock communities) to check the sequencing efficiency and sequencing error. For that, you have to select axenic cultures, to know the starting cells concentration and the DNA concentration. We could buy commercial Mock communities (ex: Zymo).

## With Phusion Master mix (1 round)

2.3 Place plates, tubes and nuclease free water to prepare the PCR mix under UV.

PCR amplification is done in triplicate (3 x 20µl), as follows:

| Reagent | Initial concentration | Final concentration | Volume |
|---|---|---|---|
| Primer V4F_illumina | 10 µM | 0.2 µM | 1 µL |
| Primer V4R_Zig_illumina | 10 µM | 0.2 µM | 1 µL |
| Phusion Master Mix2x | 2x | 1x | 10 µL |
| DMSO | 100% | 3% | 0.6 µL |
| BSA | 20 mg/mL | 0.4 mg/mL | 0.4 µL |
| DNA | | | 1 µL |
| H2O up to a 20 µL reaction | | | 6 µL |

Table 3: PCR reaction using Phusion® High-Fidelity PCR Master Mix (high GC templates)



**Citation:** Estelle Bigeard, Adriana Lopes Dos Santos, Catherine Ribeiro 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing https://dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

You can modify the amount of DNA in each reaction in order to optimize amplification. After quantifying the extraction product (see step 1, “Samples : Acid nucleic extraction and quantification”), you must perform PCR tests in order to observe the best amount of DNA in the PCR reaction, which will depend of the type of your samples.

| DNA concentration | DNA volume |
|---|---|
| < 0.2 ng/mL | 1 - 3 µL |
| Between 0.2 and 4 ng/mL | 1 µL |
| > 4 ng/mL | 0.5 µL |

Table 4 : Suggested DNA amounts

| | | Temperature (°C) | Time | |
|---|---|---|---|---|
| **Initial denaturation** | | 98°C | 5 min | |
| **Amplification** | **Denaturation** | 98°C | 20 sec | 25 cycles |
| | **Annealing** | 52°C | 30 sec | |
| | **Elongation** | 72°C | 90 sec | |
| **Final elongation** | | 72°C | 5 min | |

Table 5: Thermal conditions using Phusion® High-Fidelity PCR Master Mix

## Gel electrophoresis

3 PCR products should be checked initially by gel electrophoresis for unspecific amplification and band size.

- Check PCR products on agarose gel to see if the bands are weak or strong.

Prepare an agarose gel 1-1.5% in 0.5X TAE buffer
Microwave until the solution is clear (2 minutes)
Stain with 20 µL ethidium bromide per 100ml (or SYBR-safe dye)

Ethidium bromide 10 mg/ml **Sigma Aldrich Catalog #E1510**

Be careful : Put labcoat, nitriles gloves!



**Citation:** Estelle Bigeard, Adriana Lopes Dos Santos, Catherine Ribeiro 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing https://dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

Prepare tray & comb(s)
Pour gel into tray

When the gel is polymerized, remove comb
Load gel in electrophesis cuve containing the same 0.5X TAE buffer
Gel must be covered

Cut one strip of parafilm.
Load 1-2 µL loading dye onto parafilm, one drop per sample
Mix 1 to 3 µl of each PCR product with loading dye (up & down 3 times)
Load it in the agarose gel
Load 3 µL of SmartLadder into first and/or last well

Close & Plug in electrodes to power source
Run the migration for ~120 V, 30-45 minutes

Observe results using a UV imager like Image Quant LAZ4000, Ge Healthcare
The **negative control must be negative** and the **positive control must be positive** to validate the PCR.
The size of amplicons is indicated in Table 1 and 2.

- Samples with failed PCRs (or spurious bands) are re-amplified by optimizing the PCR (further template dilution to 1:100 or using BSA/other additives) to produce correct bands, or purified using a purification kit like “Nuclesopin Gel & Clean up kit– Macherey-Nagel”.

Amplicons are stored at -20°C until treatment.

## PCR Purification

4 This step allows to pool triplicates and clean PCR products, in particular the band of interest.
For that, we use the purification kit from Macherey-Nagel™ : NucleoSpin™ Gel and PCR Clean-up Kit.

Follow manual instructions, Instruction-NucleoSpin-Gel-and-PCR-Clean-up.pdf (mn-net.com), with severals modifications:

- First elution : add 30µl TE preheated at 70°C onto the column, incubate at 70°C for 5 minutes
- Centrifuge at 11000g 1 min
- Second elution: add 30µl TE preheated at 70°C or the previous eluate onto the column
- Centrifuge at 11000g 1 min
- Aliquote 2.5µl for the quantification

The purified PCR products are stored at -20°C.

## Control of the quantity & quality



**Citation:** Estelle Bigeard, Adriana Lopes Dos Santos, Catherine Ribeiro 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing https://dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1

5 Purified PCR products are measured using Nanoquant Plate and Spark Tecan.
This method allow to measure the concentration of amplified products and the 260/280 and 260/230 ratios using only 2µl of PCR product.

A more precise concentration could be measured using Qubit 4 fluorometer.

If the PCR's product are not concentrated enough, you could purifie and concentrate them in a smaller volume using a purification and concentration kit like "Nuclesopin Gel & Clean up kit– Macherey-Nagel".

Preparation of libraries for Illumina sequencing (metabarcoding)

6 Send **to the sequencing platform** pictures of gel to check they are conform (purity, degradation, bands intensity, etc) and follow its protocol.

Once PCR products are correct, they are organized in appropriated PCR plates to send (dryice) to the sequencing platform and generate libraries.

We use a high throughput sequencing using paired-end 2x250bp Illumina MiSeq.

**If the illumina sequencing is realized on the Genomer Platform in Roscoff (France), see the protocols.io "Preparation of libraries (Metabarcoding) for Illumina sequencing - Genomer Platform".**



**Citation:** Estelle Bigeard, Adriana Lopes Dos Santos, Catherine Ribeiro 18S-V4 rRNA amplification from total genomic DNA for NGS Illumina sequencing https://dx.doi.org/10.17504/protocols.io.yxmvmxdmnl3p/v1