



Sep 29, 2020

Diatom DNA library preparation for Illumina Miseq Sequencing using the *rbcL* marker gene

DOI

dx.doi.org/10.17504/protocols.io.bd94i98w

Cecile Chardon¹, Valentin Vasselon², Marine Vautier¹, Sonia Lacroix¹, Agnès Bouchez¹, Frédéric Rimet¹, Isabelle Domaizon¹

¹INRAE, CARRETEL, Pole R&D ECLA, Thonon les bains, France;

²Office Français pour la Biodiversité, Pole R&D ECLA, Thonon les bains, France

EcoALpsWater



Cecile Chardon

UMR CARRETEL

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN  ACCESS



DOI: <https://dx.doi.org/10.17504/protocols.io.bd94i98w>

Protocol Citation: Cecile Chardon, Valentin Vasselon, Marine Vautier, Sonia Lacroix, Agnès Bouchez, Frédéric Rimet, Isabelle Domaizon 2020. Diatom DNA library preparation for Illumina Miseq Sequencing using the *rbcL* marker gene . **protocols.io** <https://dx.doi.org/10.17504/protocols.io.bd94i98w>

Manuscript citation:

Bruder, K., & Medlin, L. K. (2007). Molecular assessment of phylogenetic relationships in selected species/genera in the naviculoid diatoms (Bacillariophyta). I. The genus *Placoneis*. *Nova Hedwigia*, 85(3), 331–352. <https://doi.org/10.1127/0029-5035/2007/0085-0331>

Keck, F., Vasselon, V., Rimet, F., Bouchez, A., & Kahlert, M. (2018). Boosting DNA metabarcoding for biomonitoring with phylogenetic estimation of operational taxonomic units' ecological profiles. *Molecular Ecology Resources*, 18(6), 1299–1309. <https://doi.org/10.1111/1755-0998.12919>

Kermarrec, L., Franc, A., Rimet, F., Chaumeil, P., Humbert, J. F., & Bouchez, A. (2013). Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: a test for freshwater diatoms. *Molecular Ecology Resources*, 13(4), 607–619. <https://doi.org/10.1111/1755-0998.12105>

Kermarrec, L., Franc, A., Rimet, F., Chaumeil, P., Frigerio, J.-M., Humbert, J., & Bouchez, A. (2014). A next-generation sequencing approach to river biomonitoring using benthic diatoms. *Freshwater Science*, 33(1), 349–363. <https://doi.org/10.1086/675079>

Rimet, F., Chaumeil, P., Keck, F., Kermarrec, L., Vasselon, V., Kahlert, M., ... Bouchez, A. (2016). R-Syst::diatom: an open-access and curated barcode database for diatoms and freshwater monitoring. *Database*, 2016, baw016. <https://doi.org/10.1093/database/baw016>

Stoof-Leichsenring, K. R., Epp, L. S., Trauth, M. H., & Tiedemann, R. (2012). Hidden diversity in diatoms of Kenyan Lake Naivasha: a genetic approach detects temporal variation. *Molecular Ecology*, 21(8), 1918–1930. <https://doi.org/10.1111/j.1365-294X.2011.05412.x>

Vasselon, V., Domaizon, I., Rimet, F., Kahlert, M., & Bouchez, A. (2017). Application of high-throughput sequencing (HTS) metabarcoding to diatom biomonitoring: Do DNA extraction methods matter? *Freshwater Science*, 36(1), 162–177. <https://doi.org/10.1086/690649>

Vasselon, V., Rimet, F., Tapolczai, K., & Bouchez, A. (2017). Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). *Ecological Indicators*, 82, 1–12. <https://doi.org/10.1016/j.ecolind.2017.06.024>

Vasselon, V., Bouchez, A., Rimet, F., Jacquet, S., Trobajo, R., Corniquel, M., ... Domaizon, I. (2018). Avoiding quantification bias in metabarcoding: Application of a cell biovolume correction factor in diatom molecular biomonitoring. *Methods in Ecology and Evolution*, 9(4), 1060–1069. <https://doi.org/10.1111/2041-210X.12960>

Visco, J. A., Apothélos-Perret-Gentil, L., Cordonier, A., Esling, P., Pillet, L., & Pawlowski, J. (2015). Environmental Monitoring: Inferring the Diatom Index from Next-Generation Sequencing Data. *Environmental Science & Technology*, 49(13), 7597–7605. <https://doi.org/10.1021/es506158m>

License: This is an open access protocol distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

We use this protocol and it's working

Created: March 26, 2020

Last Modified: September 29, 2020

Protocol Integer ID: 34844

Keywords: Diatom, rbcL, Illumina Miseq, diatom communities from freshwater sample, diatom communities through environmental dna metabarcoding approach, diatom dna library preparation for illumina miseq sequencing, diversity of diatom assemblage, benthic diatom, appropriate taxonomic marker for biomonitoring, diatom dna library preparation, species names to rbcL sequence, diatom assemblage, freshwater sample, precise taxonomic assignment, diatom community, appropriate taxonomic marker, diatoms metabarcoding, diatom, diatom metabarcoding application, environmental dna metabarcoding approach, potential of diatom metabarcoding application, illumina miseq sequencing, lake plankton sample collection, freshwater quality assessment, good taxonomical resolution, dna library for downstream miseq sequencing, ecological assessment of river, dna library, plankton sample, biofilms sample collection, biofilms dna extraction, rbcL marker gene, species name, downstream miseq sequencing, rbcL gene, rbcL sequence, dna extraction, alpswa,

Abstract

This protocol is part of the DNA workflow applied in the Eco-ALpsWater Project, here in particular to characterize the diversity of diatom assemblage in biofilms or plankton samples.

Different studies have already revealed the potential of diatom metabarcoding applications for freshwater quality assessment (Kermarrec et al. 2014; Vasselon et al. 2017ab; Visco et al. 2015). The choice of the marker gene and barcode region is key for obtaining relevant inventories of diversity and precise taxonomic assignment. For benthic diatoms, the *rbcL* gene has proved to be an appropriate taxonomic marker for biomonitoring (Kermarrec et al. 2013, 2014; Vasselon et al. 2017a,b) and a well-curated barcode reference library is already available to assign species names to *rbcL* sequences (R-Syst::diatom, Rimet et al. 2016).

For the Eco-AlpsWater project, biofilms sampled in rivers and lakeshores are collected as described in the dedicated protocols ("Lake plankton sample collection ..." and "Biofilms sample collection ...") and DNA is extracted as described in the protocol "Biofilms DNA extraction"; all these protocols are part of the Deliverable D.T1.1.2. We present here the following step in the DNA workflow (i.e. PCR amplification of selected barcodes, and wet lab methods to prepare DNA library for downstream MiSeq Sequencing). This protocol has been used in recent studies (e.g. Keck et al 2018 ; Vasselon et al 2018) where diatoms metabarcoding has been used for ecological assessment of rivers.

Several primers were proposed in the literature to characterize Diatom communities through environmental DNA metabarcoding approaches, including the 18S, COI and *rbcL* barcodes. Following the recommendation provided by Kermarrec et al. 2014, who compared the efficiency of those 3 barcodes to accurately characterize diatom communities from freshwater samples (lakes and rivers), the *rbcL* barcode will be used within the Eco-AlpsWater project as he provides a good taxonomical resolution.

Guidelines

Within the EcoAlps-Water project, a two steps PCR will be applied in order to add dual-index to each individual samples following "GenoToul Genomics and Transcriptomics" facility (GeT-PlaGe, Auzeville, France) requirement for MiSeq Illumina sequencing :

- **PCR 1** - realized by INRAE - CARRTEL, Thonon les Bains, France - described specifically in this protocol

Each DNA extracts are amplified in triplicate using the equimolar mixtures of Diat_rbcL_708F_1, Diat_rbcL_708F__2, Diat_rbcL_708F__3 and R3_1, R3_2 as forward and reverse primers respectively.

Half of the P5 (5' CTTTCCCTACACGACGCTCTTCCGATCT 3') and P7 (5' GGAGTTCAGACGTGTGCTCTTCCGATCT) Illumina adapters are included to the 5' part of the rbcL forward and reverse primers respectively. Those region will serve as primer fixation site for the second PCR.

After validation of PCR quality, the 3 PCR replicates obtained for each sample are pooled together and sent to the sequencing platform.

- **PCR 2** - realized by GeT-PlaGe, Auzeville, France - not described in this protocol

The PCR2 amplification using PCR1 purified amplicon as template and Illumina-tailed primers allowing to add dual-index specific to each sample

- **Final library preparation and sequencing** - realized by GeT-PlaGe, Auzeville, France - not described in this protocol

After the PCR2, all the indexed samples are mixed together in a final pool used to perform the Miseq sequencing.

Materials

- Samples
 - If DNA concentration is < 25ng/μL, use DNA extract
 - If DNA concentration is > 25ng/μL, diluted DNA at 25 ng/μL
- Reagents
 - TaKaRa LA Taq® DNA Polymerase (Mg²⁺ plus buffer) (Takara Bio)
 - forward and reverse *rbcl* primers at 10μM - see step 1 for the preparation
 - BSA at 10 mg/mL
 - molecular grade water
- Materials
 - 2 specific DNA-work stations (sterile area equipped with air filtration and UV systems) : 1 for steps before PCR amplification and 1 for steps after PCR amplification
 - thermal cycler
 - microcentrifuge for 1.5 to 2 mL tubes and for strips of 8 PCR tubes
 - pipettes : 1000 μL - 100 μL - 10μL (for steps before PCR amplification) and 100μL - 10μL (for steps after PCR amplification)
 - 2 refrigerated racks : 1 for 2mL tubes and 1 for strips of 8 PCR tubes
 - trash cans : 1 for liquid and 1 for solid
 - materials needed to control the PCR product
- Consumables
 - tips with filter :
 - > 1000μL
 - > 100μL : minimum 2 tips per samples
 - > 10μL : 2 tips per samples
 - 1.5 or 2 mL sterile microcentrifuge tube : minimum 1 per samples
 - strips of 8 PCR tubes
 - 96 wells plates (plate reference AB1400L or 4titude 0770/C)
 - transparent adhesive films for PCR (4titude)
 - gloves

Troubleshooting

Before start

- The following precautions must be applied to avoid contaminations:
 - PCR amplification and amplicon manipulation must be performed in different location (to avoid DNA contamination by amplicons)
 - **All steps have to be performed under a specific DNA-work station** (sterile area equipped with air filtration and UV systems).
 - Clean the bench with DNA off
 - Wear gloves throughout the process
 - Use tips with filters
 - Use fresh reagents and when you open a new reagent prepare aliquots
- Important recommendations:
 - Always include positive and negative controls during each PCR experiment to detect any contamination or reagents problems
 - Taq DNA Polymerase must be high fidelity and high accuracy to limit errors during PCR amplification (e.g. Takara LA DNA polymerase)
- Materials preparation:
 - Clean a specific DNA - work station and apply the UV for 15min
 - Check that refrigerated racks are at -20°C
- Solution preparation:
 - Place DNA and reagents (except Taq DNA Polymerase) at +4°C to thaw

Primers preparation

1

At the start of a new DNA library preparation, you must prepare "fresh" primers at 10 μ M, the volume required to process all the samples and make aliquots to avoid freezing / thawing of primers.

1.1 Global informations

- The original diatom *rbcl* primers include the Diat_*rbcl*_708F as forward primer (5' AGG TGA AGT TAA AGG TTC ATA CTT DAA 3', Stoof-Leichsenring et al. 2012) and the R3 as reverse primer (5' CCT TCT AAT TTA CCA ACA ACT G 3', Bruder et al. 2007) for an expected amplicon length of 312 bp.

An updated version of those primers was proposed in Vasselon et al. 2017 and will be used within the Eco-AlpsWater project. The forward and the reverse *rbcl* primers are respectively composed by an equimolar mixture of 3 forward *rbcl* primers (Diat_*rbcl*_708F_1, Diat_*rbcl*_708F_2, Diat_*rbcl*_708F_3) and 2 reverse *rbcl* primers (R3_1, R3_2).

Type of primer	Name	Sequence
Forward _1	Diat_ <i>rbcl</i> _708F_1	AGGTGAAGTAAAAGGT TCWTA CT TAAA
Forward _2	Diat_ <i>rbcl</i> _708F_2	AGGTGAAGTTAAAGGTT CWTAYTTAAA
Forward _3	Diat_ <i>rbcl</i> _708F_3	AGGTGAAACTAAAGGT TCWTA CT TAAA
Reverse _1	R3_1	CCTTCTAATTTACCWA CWACTG
Reverse _2	R3_2	CCTTCTAATTTACCWA CAACAG

Basic primers allowing to target the *rbcl* barcode

- Several protocols can be proposed to perform HTS library preparation, which may vary according to the strategy

applied to add index on PCR products allowing to multiplex several samples within one HTS run (tagged PCR, ligation,...) , the sequencing technology used (PGM, MiSeq,...) and the sequencing platform requirements.

Within the EcoAlps-Water project, a two steps PCR will be applied in order to add dual-index to each individual samples following the sequencing platform requirement for MiSeq Illumina sequencing.

For the first PCR, we use the equimolar mixes of Diat_rbcL_708F_1_P5, Diat_rbcL_708F_2_P5, Diat_rbcL_708F_3_P5 and R3_1_P7, R3_2_P7 as forward and reverse primers respectively.

Half of the P5 (5' CTTTCCCTACACGACGCTCTTCCGATCT 3') and P7 (5' GGAGTTCAGACGTGTGCTCTTCCGATCT 3') Illumina adapters are included to the 5' part of the rbcL forward and reverse primers respectively. Those region will serve as primer fixation site for the second PCR

	Type of primer	Name	Sequence
	Forward _1	Diat_rbcL_708 F_1_P5	CTTTCCCTACACGACGCT CTTCCGATCTAGGTGAAGT AAAAGGTTTCWTAATAA
	Forward _2	Diat_rbcL_708 F_2_P5	CTTTCCCTACACGACGCT CTTCCGATCTAGGTGAAGT TAAAGGTTTCWTAYTTAA
	Forward _3	Diat_rbcL_708 F_3_P5	CTTTCCCTACACGACGCT CTTCCGATCTAGGTGAAAC TAAAGGTTTCWTAATAA
	Reverse _1	R3_1_P7	GGAGTTCAGACGTGTGCTC TTCCGATCTCCTTCTAATT TACCWACWACTG
	Reverse _2	R3_2_P7	GGAGTTCAGACGTGTGCTC TTCCGATCTCCTTCTAATT TACCWACAACAG

Primers used during the PCR1 step for MiSeq Library preparation - *rbcL* maker gene

1.2 Preparation

- Equimolar mixture of *rbcL* primers at 100µM
 - for foward *rbcL* primers : mix 1volume of Diat_rbcL_708F_1_P5 (at 100µM), 1volume of Diat_rbcL_708F_2_P5 (at 100µM) and 1volume of Diat_rbcL_708F_3_P5 (at 100µM)



- for reverse *rbcl* primers : mix 1volume of R3_1_P7 (at 100 μ M), 1volume of R3_2_P7 (at 100 μ M)

Note : ensure that the final volume of the equimolar mix at 100 μ M allows to prepare the necessary volume of equimolar mixture at 10 μ M

- 1/10th dilution to have equimolar mixture of *rbcl* primers at 10 μ M
- for both mix : add 1volume of equimolar mix of *rbcl* primers at 100 μ M and 9volume of molecular grade water

Notes :

- *ensure that the final volume of the equimolar mix at 10 μ M allows to amplify all the samples*
- *make aliquots to avoid freezing / thawing of primers*
- *stored aliquots at -20°C*

PCR 1

2 Important informations :

- Each DNA extracts are amplified in triplicate
- PCR1 amplifications are performed using 1 μ L of extracted DNA (max [25ng/ μ L]) in a final volume of 25 μ L following mixture and reaction conditions as follow

- 2.1 ■ Under the specific DNA-work station, prepare one PCR mixture for all samples to be treated, see table below

Notes :

- *include in the PCR mixture calculations, 1 negative control, 1 positive control, 1 security sample*
- *before using a reagent or sample, mix it and make a brief centrifugation*
- *use refrigerated racks to keep reagents and PCR mixture in cool conditions*
- *add Taq DNA Polymerase at the last moment, pipetting slowly because the Taq DNA Polymerase solution is viscous*

	Reagents	[Initial]	[Final]	Volume per sample (μ L)
	Molecular grade water	-	-	15.6
	Buffer	10 X	1 X	2.5
	dNTP	2.5 mM	0.2 mM	2



BSA	10 mg/mL	0.5 mg/mL	1.25
Primer F	10 μ M	0.5 μ M	1.25
Primer R	10 μ M	0.5 μ M	1.25
TaKaRa LA Taq® DNA Polymerase	5 U/ μ L	0.75 U	0.15

PCR 1 - *rbcl* maker gene - Mixture composition for one sample with 25 μ L of final volume

- Vortex slowly the PCR mixture, make a brief centrifugation
- Distribute the PCR mixture, in strips of 8 PCR tubes, 24 μ L per PCR tube

Note :

- *You can use the same tip, to distribute the mixture in few PCR tubes*

- Add samples : 1 μ L of extracted DNA (max [25ng/ μ L])

Notes :

- *for negative control, add 1 μ L of solution without diatoms (e.g. molecular grade water)*
- *for positive control, add 1 μ L of DNA extract which must be amplified (e.g. 1 μ L of DNA extracted from diatom culture or 1 μ L of DNA extracted from environmental sample which be amplified during a previous PCR)*

- *volume of DNA used per PCR reaction can be adjusted if DNA concentration is low and water volume adjusted in consequence (e.g. if PCR performed with 3 μ L of DNA, water volume will be set to 13.6 μ L); if the concentration of extracted DNA is > 25ng/ μ L, make a dilution to have a final concentration at 25ng/ μ L and for the PCR add 1 μ L of diluted DNA*

- Close strips of 8 PCR tubes
- Vortex slowly the mixture of PCR mixture / DNA, make a brief centrifugation

- 2.2
- Place strips of 8 PCR tubes in thermal cycler
 - Start the program below

	Temperature	Time	Cycles
	95 °C	5 min	X1
	95 °C	1 min	X33
	54 °C	1 min	



72 °C	1 min	
72 °C	5 min	X1
10 °C	pause	

PCR 1 - *rbcl* maker gene - Reaction conditions

Note : cycle number can be increased to a maximum of 40 cycles but must be avoided to limit aspecific amplification and to limit biases linked to PCR amplification.

PCR validation, PCR1 product pool preparation, plate preparation and shipment

3 Important informations :

- Reminders :

- Each DNA extracts are amplified in triplicate
- Final PCR volume is 25µL

- Get-PlaGe facility requirements :

- Volume final of PCR1 product needed is 50µL and concentration minimum needed is 30ng/µL
- PCR1 products must be placed in 96 wells plates sealed by adhesive film
- PCR1 products must be one after the other in column, without empty wells

- Recommendations :

- To limit the risk of contaminations, work in a room dedicated to handling amplicons with dedicated materials (pipettes, specific DNA-work station,...)

- You need to reduce at maximum freezing / thawing of PCR1 products

we therefore recommend that you :

1) carry out the PCR reaction, validate the PCR and prepare the PCR pool over a week and during this period store the PCR1 products at + 4 ° C (in the fridge)

2) place the PCR1 product pools at -20°C while you process the other samples

3) thawing the PCR1 product pools to prepare the 96 wellsplates 2-3 days before plates shipment

4) realize shipment in cold conditions (0-4°C), in 1 - 2 days

3.1 PCR validation

Notes :

- *details of this protocol are not describe here*

- *to realize a quality control of PCR product (lenght of fragment and quantity), you can use electrophoresis on agarose gel or other techniques (e.g : TapeStation (Agilent Technologies, Santa Clara, California))*

- *use maximum of 4µL of PCR product per PCR reaction*

- For the quality control of PCR product, check :
 - for negative control, no amplification
 - for positive control, expected amplicon length \approx 376 bp
 - for samples, expected amplicon length \approx 376 bp
 - for one sample, amplification identical between triplicate of PCR product

If these 4 points are validated, you can pool triplicate of PCR product, one pool per sample

3.2 PCR1 product pools preparation

Notes :

- One pool per sample with a final volume at 57 μ L
- For one sample, use same tip
- To reduce the risk of errors, write the name of sample and "rbcl product" on the tube

- Clean a specific DNA - work station (dedicated to handling amplicons) and apply the UV for 15min
- Annotate the 1.5 or 2mL tubes
- Vortex slowly the PCR1 products, make a brief centrifugation
- Take 19 μ L of PCR1 product per replicate and place its in corresponding annotated tube

(so in total you have 3 X 19 μ L = 57 μ L)

- Store PCR1 product pools (if the plate will be prepared under 2-3 days, store at +4°C otherwise store at -20°C)

Note : we recommend you to realize and sequence negative control pools.

e.g if you have 9 sets of PCR, you can realize 3 negative control pools. The 1st with negative controls of PCR sets 1, 2, 3; The 2nd with negative controls of PCR sets 4, 5, 6; The 3rd with negative controls of PCR sets 7, 8, 9.

3.3 Plate preparation

- Before plate preparation, you need :
 - prepare map(s) of plate(s)
 - clean a specific DNA - work station (dedicated to handling amplicons) and apply the UV for 15min
 - thaw the PCR1 product pools
 - vortex them slowly and make a brief centrifugation
 - place the PCR1 product pools in a rack, in the right order

- Under specific DNA - work station, on 96 wells plate (use reference recommended by Get-PlaGe facility) write the plate name
- Pipetting 52µL of PCR1 product pool and put them in the right well
- Change the tip, and start again with the second sample
- Repeat this last step for all PCR1 product pools

Notes : to limit risk of errors,

- *fill in the plate per column*
- *use one new tips box per plate and use tips by column too*
- *move or discard the PCR1 product pool tube after each use*
- When your 96 wells plate is completed, sealed it correctly with adhesive films for PCR recommended by Get-PlaGe facility
- Place 96 wells plate in a bag and stored at + 4°C (maximum 2-3 days) before the shipment

3.4 Shipment

- Place bag containing sealed plate in a box.
- This box must contain everything necessary to ensure a temperature between 0°C and + 4°C
- The shipment must be made within 1 to 2 days

Note :

- *it is preferable to send the box before Wednesday*
- Send an email to the platform with the Excel document required for sample processing