

Molecular Standard Operating Procedure (MSOP)

for

Marine Biodiversity Observation network for genetic monitoring of hard-bottom communities (ARMS-MBON)

Version

1.0

Date

2020-04-01

## Purpose

This document contains the Standard Operating Procedures for working with the molecular data of the ARMS-MBON (www.arms-mbon.eu) project. The samples containing the material are sent by each observatory to HCMR for processing (see the [Handbook](http://www.assembleplus.eu/sites/assembleplus.eu/files/public/ARMS/ARMS_Handbook.pdf) for details).

## DNA Extraction

This protocol is used for each of the three ARMS fractions (motile 100μm ‒ 500μm, motile 500m ‒ 2mm, and sessile).

**Materials:**

* Falcon tubes containing the samples stored in DMSO
* DNA-extraction kit (DNeasy PowerSoil Kit or DNeasy PowerSoil Pro Kit)
* Sterile pipettes and pipette tips
* DNA-decontaminating solution
* agarose/EtBr gel and loading buffer
* DNA size ladder

**Procedures:**

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. Use about 0.5 grams of wet material from each sample. Extract each replicate sample separately.
3. Evaluate the quality of the extracted DNA by gel electrophoresis and quantify it using a spectrophotometer.
4. Store the DNA at -20 oC until further processing.

## PCR amplification and sequencing

**Materials:**

* Primers

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Primer Name** | **Sequence** | **Target gene** | **Target group** | **Amplicon size (bp)** | **Reference** |
| All18SF | 5′‐TGGTGCATGGCCGTTCTTAGT‐3′ | 18S rRNA | metazoa, fungi, protozoa, plants | 200-500 | ( Hardy et al. 2010 |
| All18SR | 5′‐CATCTAAGGGCATCACAGACC‐3′ |
| mlCOIintF | 5′‐GGWACWGGWTGAACWGTWTAYCCYCC‐3′ | COI | metazoa | 313 | Leray et al. 2013 |
| jgHCO2198 | 5′‐TAIACYTCIGGRTGICCRAARAAYCA‐3′ | Geller et al. 2013 |
| ITS1f | 5′‐CTTGGTCATTTAGAGGAAGTAA‐3′ | ITS | fungi | 250--600 | Gardes & Bruns 1993 |
| ITS2 | 5′‐GCTGCGTTCTTCATCGATGC‐3′ | ( White et al. 1990 |

* 5′ tails used for the first-step PCR

|  |  |
| --- | --- |
| **Primer Name** | **S Sequence** |
| 1st\_PCR\_for | 5′‐TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[locus-specific sequence] ‐3′ |
| 1st\_PCR\_rev | 5′‐GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[locus-specific sequence] ‐3′ |

* Extracted DNA
* KAPA HiFi HotStart PCR Kit and KAPA Taq PCR Kit
* Qubit® dsDNA HS Assay Kit or Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher/Invitrogen)
* AMPure XP beads (Beckman Coulter) or NucleoMag® NGS Clean-up and Size Select (Macherey-Nagel)
* KAPA Illumina Library Quantification Kit and Illumina Library Quantification DNA Standards
* Thermal cycler
* Dedicated pipettes and pipette tips
* PCR reaction tubes and/or plates
* PCR grade water
* agarose/EtBr gel and loading buffer
* MiSeq Reagent kit v3 (600 cycles)

1. Wear gloves at all times. Carefully clean the bench station(s) and pipettes.
2. Mix the biological replicate samples in equimolar amounts, so that you end up with one tube for each sample (i.e. for each MaterialSample-ID).
3. PCR amplification is performed targeting three gene regions: COI (metazoa), 18S rRNA (metazoa) and ITS (fungi), using the Two-Step PCR Approach.
4. The first-step PCR is performed with the aforementioned primers containing a universal 5′ tail as specified in the Nextera library protocol from Illumina.

* The first-step PCR for the COI:

Amplification reaction mix:

* 3.0 μl 10x KAPA Taq buffer A,
* 0.6 μl MgCl2 (25 mM),
* 0.75 μl KAPA dNTP Mix (10 mM),
* 1.8 μl from each primer (10 μM),
* 0.9 μl KAPA Taq DNA polymerase (5.0 U/μl)
* The final volume was 30.0 μl per reaction.
* DNA template concentration is about 10.0 ng/μl.

First PCR protocol:

* + - 95 °C for 5 min;
    - 16 cycles at 95 °C for 10 s, 62 °C (-1 °C/cycle) for 30 s, 72 °C for 1 min;
    - 24 cycles at 95 °C for 10 s, 46 °C for 30 s, 72 °C for 1 min;
    - 72 °C for 7 min
* The first-step PCR for the 18S rRNA:

Amplification reaction mix:

* 6.0 μl 5x KAPA HiFi Fidelity buffer,
* 6.0 μl Trehalose (1 M),
* 0.9 μl KAPA dNTP Mix (10 mM),
* 1.8 μl from each primer (5 μM),
* 0.6 μl KAPA HiFi HotStart DNA polymerase (1.0 U/μl)
* The final volume was 30.0 μl per reaction.
* DNA template concentration is about 10.0 ng/μl.

First PCR protocol:

* + - 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
* The first-step PCR for the ITS:

Amplification reaction mix:

* 3.0 μl 10x KAPA Taq buffer A,
* 0.75 μl KAPA dNTP Mix (10 mM),
* 1.5 μl from each primer (10 μM),
* 0.9 μl KAPA Taq DNA polymerase (5.0 U/μl)
* The final volume was 30.0 μl per reaction.

DNA template concentration is about 10.0 ng/μl.

First PCR protocol:

* 95 °C for 5 min;
* 35 cycles at 95 °C for 30 s, 52 °C for 30 s, 68 °C for 30 s;
* 68 °C for 10 min

1. Purify 20 μl of the resulting PCR amplicons using magnetic beads, at a ratio 1:1 (magnetic beads: PCR product).
2. Quantify the purified PCR amplicons using fluorometric quantitation.
3. Use the purified and quantified 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.

*Indexed forward primers for the second-step PCR:*

|  |  |  |
| --- | --- | --- |
| **Primer Name** | **Sequence** | **Index name** |
| NGS\_i5\_S502 | 5′‐AATGATACGGCGACCACCGAGATCTACACCTCTCTATTCGTCGGCAGCGTC‐3′ | S502 |
| NGS\_i5\_S503 | 5′‐AATGATACGGCGACCACCGAGATCTACACTATCCTCTTCGTCGGCAGCGTC‐3′ | S503 |
| NGS\_i5\_S505 | 5′‐AATGATACGGCGACCACCGAGATCTACACGTAAGGAGTCGTCGGCAGCGTC‐3′ | S505 |
| NGS\_i5\_S506 | 5′‐AATGATACGGCGACCACCGAGATCTACACACTGCATATCGTCGGCAGCGTC‐3′ | S506 |
| NGS\_i5\_S507 | 5′‐AATGATACGGCGACCACCGAGATCTACACAAGGAGTATCGTCGGCAGCGTC‐3′ | S507 |
| NGS\_i5\_S508 | 5′‐AATGATACGGCGACCACCGAGATCTACACCTAAGCCTTCGTCGGCAGCGTC‐3′ | S508 |
| NGS\_i5\_S510 ′ | 5′‐AATGATACGGCGACCACCGAGATCTACACCGTCTAATTCGTCGGCAGCGTC‐3 | S510 |
| NGS\_i5\_S511 | 5′‐AATGATACGGCGACCACCGAGATCTACACTCTCTCCGTCGTCGGCAGCGTC‐3′ | S511 |
| NGS\_i5\_S513 | 5′‐AATGATACGGCGACCACCGAGATCTACACTCGACTAGTCGTCGGCAGCGTC‐3′ | S513 |
| NGS\_i5\_S515 | 5′‐AATGATACGGCGACCACCGAGATCTACACTTCTAGCTTCGTCGGCAGCGTC‐3′ | S515 |
| NGS\_i5\_S516 | 5′‐AATGATACGGCGACCACCGAGATCTACACCCTAGAGTTCGTCGGCAGCGTC‐3′ | S516 |
| NGS\_i5\_S517 | 5′‐AATGATACGGCGACCACCGAGATCTACACGCGTAAGATCGTCGGCAGCGTC‐3′ | S517 |
| NGS\_i5\_S518 | 5′‐AATGATACGGCGACCACCGAGATCTACACCTATTAAGTCGTCGGCAGCGTC‐3′ | S518 |
| NGS\_i5\_S520 | 5′‐AATGATACGGCGACCACCGAGATCTACACAAGGCTATTCGTCGGCAGCGTC‐3′ | S520 |
| NGS\_i5\_S521 | 5′‐AATGATACGGCGACCACCGAGATCTACACGAGCCTTATCGTCGGCAGCGTC‐3′ | S521 |
| NGS\_i5\_S522 | 5′‐AATGATACGGCGACCACCGAGATCTACACTTATGCGATCGTCGGCAGCGTC‐3′ | S522 |

*Indexed reverse primers for the second-step PCR:*

|  |  |  |
| --- | --- | --- |
| **Primer Name** | **Sequence** | **Index name** |
| NGS\_i7\_N701 | 5′‐CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGG‐3′ | N701 |
| NGS\_i7\_N702 | 5′‐CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGG‐3′ | N702 |
| NGS\_i7\_N703 | 5′‐CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGG‐3′ | N703 |
| NGS\_i7\_N704 | 5′‐CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGG‐3′ | N704 |
| NGS\_i7\_N705 | 5′‐CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGG‐3′ | N705 |
| NGS\_i7\_N706 | 5′‐CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGG‐3′ | N706 |
| NGS\_i7\_N707 | 5′‐CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGG‐3′ | N707 |
| NGS\_i7\_N710 | 5′‐CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGG‐3′ | N710 |
| NGS\_i7\_N711 | 5′‐CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGG‐3′ | N711 |
| NGS\_i7\_N712 | 5′‐CAAGCAGAAGACGGCATACGAGATTCCTCTACGTCTCGTGGGCTCGG‐3′ | N712 |
| NGS\_i7\_N714 | 5′‐CAAGCAGAAGACGGCATACGAGATTCATGAGCGTCTCGTGGGCTCGG‐3′ | N714 |
| NGS\_i7\_N715 | 5′‐CAAGCAGAAGACGGCATACGAGATCCTGAGATGTCTCGTGGGCTCGG‐3′ | N715 |
| NGS\_i7\_N716 | 5′‐CAAGCAGAAGACGGCATACGAGATTAGCGAGTGTCTCGTGGGCTCGG‐3′ | N716 |
| NGS\_i7\_N718 | 5′‐CAAGCAGAAGACGGCATACGAGATGTAGCTCCGTCTCGTGGGCTCGG‐3′ | N718 |
| NGS\_i7\_N719 | 5′‐CAAGCAGAAGACGGCATACGAGATTACTACGCGTCTCGTGGGCTCGG‐3′ | N719 |
| NGS\_i7\_N720 | 5′‐CAAGCAGAAGACGGCATACGAGATAGGCTCCGGTCTCGTGGGCTCGG‐3′ | N720 |
| NGS\_i7\_N721 | 5′‐CAAGCAGAAGACGGCATACGAGATGCAGCGTAGTCTCGTGGGCTCGG‐3′ | N721 |
| NGS\_i7\_N722 | 5′‐CAAGCAGAAGACGGCATACGAGATCTGCGCATGTCTCGTGGGCTCGG‐3′ | N722 |
| NGS\_i7\_N723 | 5′‐CAAGCAGAAGACGGCATACGAGATGAGCGCTAGTCTCGTGGGCTCGG‐3′ | N723 |
| NGS\_i7\_N724 | 5′‐CAAGCAGAAGACGGCATACGAGATCGCTCAGTGTCTCGTGGGCTCGG‐3′ | N724 |
| NGS\_i7\_N726 | 5′‐CAAGCAGAAGACGGCATACGAGATGTCTTAGGGTCTCGTGGGCTCGG‐3′ | N726 |
| NGS\_i7\_N727 | 5′‐CAAGCAGAAGACGGCATACGAGATACTGATCGGTCTCGTGGGCTCGG‐3′ | N727 |
| NGS\_i7\_N728 | 5′‐CAAGCAGAAGACGGCATACGAGATTAGCTGCAGTCTCGTGGGCTCGG‐3′ | N728 |
| NGS\_i7\_N729 | 5′‐CAAGCAGAAGACGGCATACGAGATGACGTCGAGTCTCGTGGGCTCGG‐3′ | N729 |

1. The amplification reaction mix of the second PCR contains:

6.0 μl 5x KAPA HiFi Fidelity buffer,

0.75 μl KAPA dNTP Mix (10 mM),

3.0 μl from each indexed primer (10 μM),

0.75 μl KAPA HiFi HotStart DNA polymerase (1.0 U/μl)

The final volume was 30 μl per reaction.

DNA template concentration is about 20.0 ng/ μl.

The second PCR protocol is:

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.

1. Purify 20 μl of the resulting PCR amplicons and quantify them.
2. Calculate the concentration of the PCR amplicons (nM) using the equation

(1000000\*Concentration in ng/ul)/(Total amplicom length in bp\*660).

1. Create an amplicon sequencing pool by mixing the PCR amplicons in equimolar amounts (at a desired final concentration of 10 nM).
2. Quantify the amplicon pool using the Illumina Library Quantification Kit and dilute it to the desired concentration, according to the Illumina sequencing protocol.
3. Sequence the amplicon pool using a MiSeq Reagent Kit v3 (2 × 300-cycles).
4. Submit the raw sequence files to the European Nucleotide Archive (ENA) (Amid et al. 2019).
5. Analyse the sequences using the PEMA pipeline (Zafeiropoulos et el. 2020). Examples of parameter values that could be a starting point for the analysis are shown below.

*Example parameters for the tools invoked by PEMA:*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tool** | **Parameter** | **Parameter Value** | | |
| **COI** | **18S rRNA** | **ITS** |
| trimmomatic (v.0.38) | maxInfo | Yes | Yes | Yes |
| targetLength | 200 | 200 | 200 |
| strictness | 0.3 | 0.3 or 0.5 | 0.3 |
| seedMismatches | 2 | 2 | 2 |
| palindromeClipThreshold | 30 | 30 | 30 |
| simpleClipThreshold | 15 | 15 | 15 |
| leading | 10 | 10 | 10 |
| trailing | 15 | 15 | 15 |
| minlen | 100 | 100 | 100 |
| PANDAseq (v. 2.11) | pandaseqAlgorithm | simple\_bayesian | simple\_bayesian | simple\_bayesian |
| pandaseqMinlen | 150 | 150 | 150 |
| minoverlap | 20 | 20 | 20 |
| threshold | 0.6 | 0.6 | 0.6 |
| SWARM  (v. 2) | d | 10 | 1 | 5 or 20 |
| removeSingletons | Yes | No | No |

Deviations from the MSOP:

In the 18S rRNA amplifications of the pilot 2018 samples, the protocol that was followed was slightly different: a) the 1st PCR primers included the barcodes and b) the ligation of the adaptors was performed with the TruSeq DNA PCR-free amplicon workflow.

## 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.

Gardes, M., and Bruns, T.D. 1993. ITS primers with enhanced specificity for basidiomycetes ‐ application to the identification of mycorrhizae and rusts. Mol. Ecol. **2**(2): 113–118. doi:10.1111/j.1365-294X.1993.tb00005.x.

Geller, J., Meyer, C., Parker, M., and Hawk, H. 2013. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Mol. Ecol. Resour. **13**(5): 851–61. doi:10.1111/1755-0998.12138.

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.

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T., and Machida, R.J. 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Front. Zool. **10**(1): 34. Frontiers in Zoology. doi:10.1186/1742-9994-10-34.

White, T.J., Bruns, T., Lee, S., and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* PCR protocols: a guide to methods and applications. pp. 315–322. doi:10.1016/b978-0-12-372180-8.50042-1.

Zafeiropoulos, H., Viet, H.Q., Vasileiadou, K., Potirakis, A., Arvanitidis, C., Topalis, P., Pavloudi, C., and Pafilis, E. 2019. PEMA: from the raw .fastq files of 16S rRNA and COI marker genes to the (M)OTU-table, a thorough metabarcoding analysis. bioRxiv: 709113. doi:10.1101/709113.