Chapter 14

Preparation of Amplicon Libraries for Metabarcoding of Marine Eukaryotes Using Illumina MiSeq: The Adapter Ligation Method

Matthieu Leray, Quiterie Haenel, and Sarah J. Bourlat

Abstract

Amplicon-based studies of marine microscopic eukaryotes, also referred to as metabarcoding studies, can be performed to analyze patterns of biodiversity or predator–prey interactions targeting the mitochondrial cytochrome oxidase 1 (CO1) or the small ribosomal subunit (SSU) markers. Because high-throughput sequencing (HTS) Illumina platforms provide millions of reads per run, hundreds of samples may be sequenced simultaneously. This protocol details the preparation of multiplexed amplicon libraries for Illumina MiSeq sequencing. We describe a strategy for sample multiplexing using a combination of tailed PCR primers and ligation of indexed adapters.

Key words Metabarcoding, Eukaryotes, Illumina MiSeq, Multiplexing, Tags, TruSeq

1 Introduction

High Throughput Sequencing (HTS) technologies have revolutionized the field of community ecology in recent years. Deep sequencing of PCR amplicons provides cost-effective estimates of species diversity and taxonomic composition from samples that were traditionally sorted by hand [1, 2] and characterized at coarse taxonomic levels [3]. Moreover, community profiles generated using HTS data are independent of taxonomic expertise and therefore more comparable between studies for environmental monitoring and environmental status assessment [4–6]. Because the number of reads produced in a single HTS run is far greater than what is necessary for characterizing most community samples, hundreds of samples may be pooled into a single run to be sequenced simultaneously [7, 8]. Prior to pooling, sequences belonging to each sample must be tagged with a unique identifier to be recognizable in downstream data processing. Several alternative tagging approaches have been used in the literature for sample

Sarah J. Bourlat (ed.), *Marine Genomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1452,
DOI 10.1007/978-1-4939-3774-5_14,

multiplexing (e.g., the dual PCR method, *see* Chapter 13 by Bourlat et al. and Chapter 12 by Fonseca and Lallias). In this chapter, we present a hierarchical tagging method for sequencing amplicons on the Illumina MiSeq that combines the use of tailed PCR primers followed by the ligation of indexed adapters (*see* Fig. 1). This method uses two sets of tags for maximizing the number of samples pooled in a MiSeq run. The first tag is added to the target fragment during PCR amplification using tailed locus-specific primers. The second tag is added to the product of amplification via ligation in the form of a single-indexed Illumina Y-adapter. An example of the pooling strategy is presented in Fig. 2. Marker sequences amplified using this method may be up to 500 bp using the Illumina V3 reagent kit (2 × 300 bp paired end), including a 100 bp overlap for paired end reads, which is suitable for sequencing fragments of the cytochrome oxidase 1 (CO1) or the small ribosomal subunit (SSU) markers.

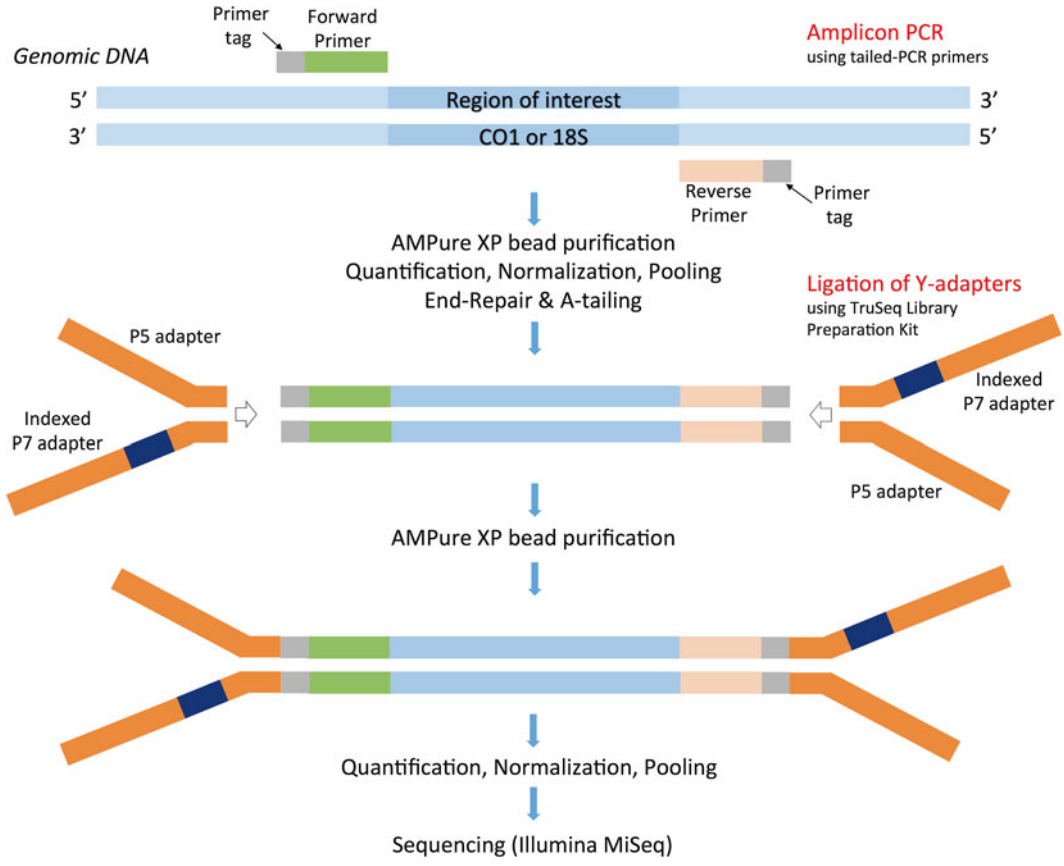

Fig. 1 Scheme for Illumina MiSeq multiplex library preparation using the tailed PCR primers and ligation of single-indexed Y-adapters. The first PCR step uses amplicon-specific primers with a 6-nucleotide tag. The second step uses a ligation to add a single-indexed Y-adapter to pooled PCR amplicons

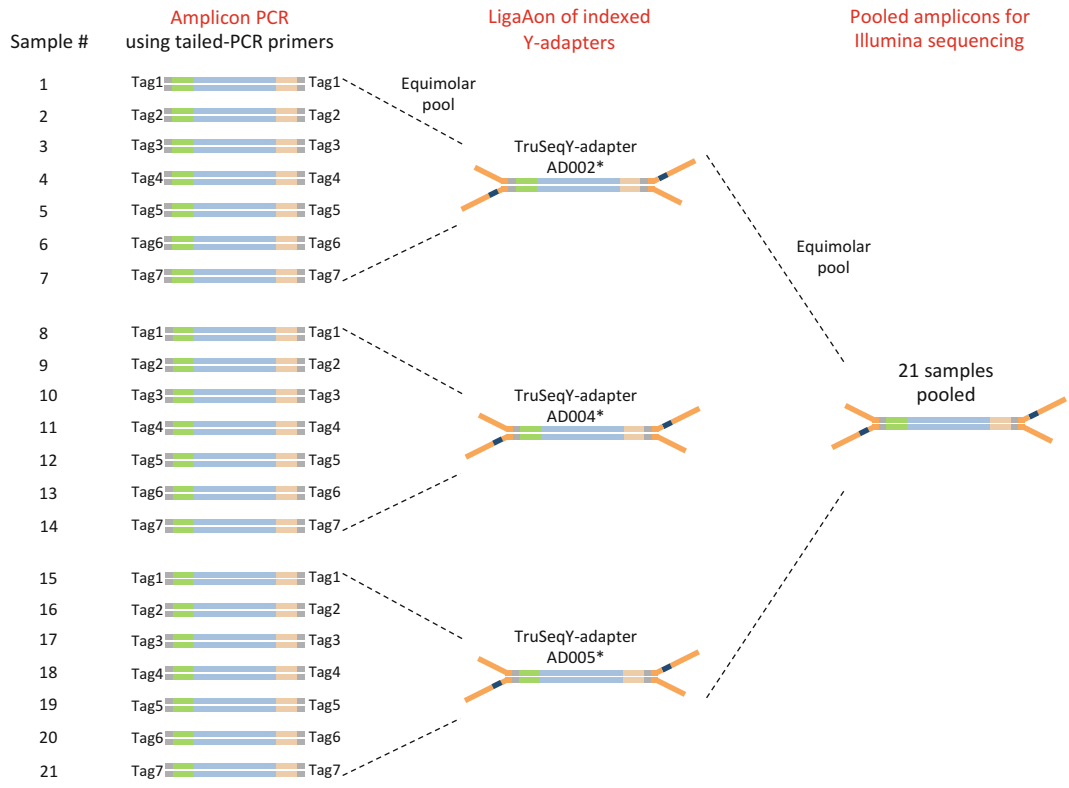

Fig. 2 Illustration of the pooling strategy used for the adaptor ligation library preparation

2 Materials

2.1 DNA Extraction

1. For sediment samples, use the PowerSoil DNA Isolation Kit (MoBio). For tissue samples, use the UltraClean Tissue & Cells DNA Isolation Kit (MoBio) or equivalent. For environmental samples with large biomass (e.g. plankton), use PowerMax Soil DNA Isolation Kit (MoBio).
2. DNA cleanup kit such as PowerClean Pro DNA Clean-Up Kit (MoBio).

2.2 PCR Amplification Using Tailed Locus-Specific PCR

1. 96-well PCR plates and microseal film.
2. 50× Clontech Advantage2 Polymerase Mix (Takara).
3. 10× Clontech Advantage2 PCR buffer (Takara).
4. dntp mix containing sodium salts of dATP, dCTP, dGTP, and dTTP, each at 2.5 mM (total concentration 10 mM).
5. Nuclease-free water.
6. Thermocycler.
7. Gel electrophoresis apparatus and reagents.

8. 100 bp DNA ladder.
9. Tailed locus-specific PCR primers, as detailed below:

   For this protocol, we designed six nucleotide tags that are appended to the locus-specific forward and reverse primers (tailed primers) to amplify a 313 bp region of CO1 [9]. Tags differ from each other by at least three nucleotides. A total of 49 primer combinations can be used for sample multiplexing using the seven unique tailed forward and seven unique tailed reverse PCR primers presented below (e.g., sample1, mlCOIint_Tag1—dgHCO_Tag1; sample 2, mlCOIint_Tag1—dgHCO_Tag2; sample 3, mlCOIint_Tag2—dgHCO_Tag1; sample 4, mlCOIint_Tag2—dgHCO_Tag2…). These same tags can be appended to other PCR primers to amplify CO1 (e.g., Lobo_R1, based on [10]) or 18S (e.g., SSU_FO4, SSU_R22, based on [1]). *See* Chapter 13 by Bourlat et al. for further information on the 18S LOBO primers and chapter 12 by Fonseca and Lallias for further information on SSU primer pairs (SSU_FO4 and SSU_R22).

   All CO1-specific tailed forward and reverse primers are detailed below:

Primer name	Tailed primer sequence (5′–3′)
mlCOIintF_Tag1	AGACGC**GGWACWGGWTGAACWGTWTAYCCYCC**
mlCOIintF_Tag2	AGTGTA**GGWACWGGWTGAACWGTWTAYCCYCC**
mlCOIintF_Tag3	ACTAGC**GGWACWGGWTGAACWGTWTAYCCYCC**
mlCOIintF_Tag4	ACAGTC**GGWACWGGWTGAACWGTWTAYCCYCC**
mlCOIintF_Tag5	ATCGAC**GGWACWGGWTGAACWGTWTAYCCYCC**
mlCOIintF_Tag6	ATGTCG**GGWACWGGWTGAACWGTWTAYCCYCC**
mlCOIintF_Tag7	ATAGCA**GGWACWGGWTGAACWGTWTAYCCYCC**
dgHCO2198_Tag1	AGACGC**TAAACTTCAGGGTGACCAAARAAYCA**
dgHCO2198_Tag2	AGTGTA**TAAACTTCAGGGTGACCAAARAAYCA**
dgHCO2198_Tag3	ACTAGC**TAAACTTCAGGGTGACCAAARAAYCA**
dgHCO2198_Tag4	ACAGTC**TAAACTTCAGGGTGACCAAARAAYCA**
dgHCO2198_Tag5	ATCGAC**TAAACTTCAGGGTGACCAAARAAYCA**
dgHCO2198_Tag6	ATGTCG **TAAACTTCAGGGTGACCAAARAAYCA**
dgHCO2198_Tag7	ATAGCA**TAAACTTCAGGGTGACCAAARAAYCA**

2.3 Bead Purification

1. Magnetic beads for DNA purification (also called SPRI beads for solid-phase reversible immobilization).
2. Magnetic 96-well plate.
3. Freshly prepared 70% ethanol.
4. Nuclease-free water or resuspension buffer provided in TruSeq DNA PCR-free LT Library Prep Kit (Illumina).

2.4 Normalization and Pooling

1. Qubit fluorometric quantitation apparatus.
2. Qubit dsDNA HS Assay Kit.
3. Nuclease-free water.

2.5 End Repair

1. End Repair Mix 2 from TruSeq DNA PCR-free LT Library Prep Kit (Illumina).
2. Resuspension Buffer from TruSeq DNA PCR-free LT Library Prep Kit (Illumina).
3. Thermocycler with heated lid.

2.6 A-Tailing

1. A-Tailing Mix from TruSeq DNA PCR-free LT Library Prep Kit (Illumina).
2. Resuspension Buffer from TruSeq DNA PCR-free LT Library Prep Kit (Illumina).
3. Thermocycler with heated lid.

2.7 Ligation of Y-Adapters

1. DNA Adapter Index tubes from TruSeq DNA PCR-free LT Library Prep Kit (Illumina).
2. Ligation Mix from TruSeq DNA PCR-free LT Library Prep Kit (Illumina).
3. Resuspension Buffer from TruSeq DNA PCR-free LT Library Prep Kit (Illumina).
4. Stop Ligation Buffer from TruSeq DNA PCR-free LT Library Prep Kit (Illumina).
5. Thermocycler with heated lid.

2.8 Final Normalization and Pooling Step

1. TapeStation.
2. Qubit fluorometric quantitation apparatus.
3. Qubit dsDNA HS Assay Kit.
4. Tris–Cl 10 nM buffer, pH 8.5 with 0.1 % Tween 20.

2.9 Library Validation Using Quantitative PCR

1. ViiA 7 real-time PCR system (Applied Biosystems).
2. KAPA library quantification kit for Illumina platform (ROX low).
3. Tris–Cl 10 nM buffer, pH 8.5 with 0.1 % Tween 20.
4. Nuclease-free water.

3 Methods

3.1 DNA Extraction

For total DNA extraction from sediment samples or samples with large biomass, use MoBio's PowerSoil DNA isolation kit according to the manufacturer's instructions. For other sample types, any DNA extraction kit can be used, with an additional DNA purification step to remove inhibitors using MoBio's PowerClean Pro DNA Clean-Up Kit according to the manufacturer's instructions.

3.2 PCR Amplification Using Tailed Locus-Specific PCR

Here, we use locus-specific primers with a 6-nucleotide tag. To minimize PCR errors, use a proofreading polymerase. Also, we recommend running three PCR replicates to minimize biases:

1. For a 20 µl reaction volume, use 2 µl Clontech Advantage2 PCR buffer (10×), 1.4 µl dntp mix (10 mM), 1 µl of each primer (10 µM), 1 µl DNA template (~10 ng), 0.4 µl Clontech Advantage2 Polymerase Mix (50×), and 13.2 µl of nuclease-free water.
2. Run three PCR replicates (e.g., three independent PCR for the same sample) using the following cycling conditions: 5 min at 95 °C (1×); 1 min at 95 °C, 45 s at 48 °C, and 30 s at 72 °C (35×); and 10 min at 72 °C (1×); hold at 4 °C.
3. Run a 1.5% agarose gel to check the size of the amplicons using a DNA ladder.
4. If any additional bands appear that are not the size of the desired product, increase the annealing temperature of the PCR or perform additional purification steps (*see* **Note 1**).
5. Pool PCR replicates.

3.3 Bead Purification

Purify amplicons using magnetic beads and a magnetic stand. Size selection can be achieved using different ratios of beads to sample (*see* **Note 1**). A ratio bead–sample of 1.6:1 will efficiently purify the amplicons away from primers and primer dimers [11]:

1. Vortex the beads before use. Add 80 µl beads to 50 µl of PCR product to obtain a ratio of 1.6. Pipette up and down ten times. Incubate a room temperature without shaking for 5 min.
2. Place the plate on the magnetic stand until the supernatant has cleared, at least 3 min.
3. Remove the supernatant with a multichannel pipette if you are using a 96-well plate, making sure not to disturb the beads.
4. With the samples on the magnetic stand, wash the beads by adding 200 µl of freshly prepared 70% ethanol, and incubate for 30 s. Carefully remove the supernatant, without disturbing the beads.
5. Repeat washing **step 4**.
6. Remove all residual ethanol using a pipette and air dry with the samples on the magnetic stand for 3 min.
7. Remove the plate from the stand and add 40 µl of nuclease-free water for elution, gently pipetting up and down ten times to resuspend the beads. Incubate the plate at room temperature for 5 min.
8. Place the plate back on the magnetic stand at least 5 min or until the supernatant has cleared. *See* also **Note 2** about bead carryover.
9. Carefully transfer the supernatant to a new plate.

3.4 Normalization and Pooling

At this step, equimolar amounts of purified amplicons are pooled, so that each pool contains amplicons generated using different tailed PCR primers (*see* Fig. 2):

1. Check the concentration of the purified amplicons using Qubit.
2. Pool equimolar amounts of each purified amplicon for a final amount of 1 μg in a final volume of 60 μl (add water if necessary).

3.5 End Repair

1. Thaw End Repair Mix 2 and resuspension buffer.
2. Add 40 μl of End Repair Mix 2 to each tube containing 60 μl of pooled PCR amplicon (1 μg) and mix gently by pipetting up and down ten times.
3. Incubate the samples for 30 min at 30 °C in a thermocycler with pre-heated lid.

3.6 Bead Purification

Perform a magnetic bead cleanup (refer to Subheading 3.3) using a ratio of 1:1.6. Add 160 μl of beads to each tube containing 100 μl of end repair mix (*see* **Note 3**). At the end of the cleanup, elute the sample in 20 μl of resuspension buffer and transfer 17.5 μl of the clear supernatant to a new tube. If you don't proceed to the next step immediately, samples can be stored in the −20 °C freezer for up to 7 days.

3.7 A-Tailing

1. Thaw A-Tailing Mix and resuspension buffer.
2. Add 12.5 μl of A-Tailing Mix to each tube containing 17.5 μl of sample and mix gently by pipetting up and down.
3. Incubate the samples for 30 min at 37 °C, followed by 5 min at 70 °C and 5 min at 4 °C in a thermocycler with pre-heated lid.
4. Proceed immediately to the ligation.

3.8 Ligation of Y-Adapters (See Fig. 2)

1. Thaw DNA Adapter Index tubes and stop ligation buffer. Leave Ligation Mix in the freezer until immediately before use and keep on ice once removed from the freezer.
2. Add reagents in the following order to each tube containing 30 μl of end-repaired and A-tailed PCR amplicons – (1) 2.5 μl of resuspension buffer; (2) 2.5 μl of ligation mix; (3) and 2.5 μl of appropriate DNA Adapter Index – and mix gently by pipetting up and down.
3. Incubate the samples for 10 min at 30 °C in a thermocycler with pre-heated lid and then hold at 4 °C.
4. Add 5 μl of stop ligation buffer to each sample and mix gently by pipetting up and down.

3.9 Bead Purifications (Twice)

Perform a magnetic bead cleanup (refer to Subheading 3.3) using a ratio of 1:1 following Illumina TruSeq protocol. Add 42.5 μl of beads to each sample containing 42.5 μl of Adapter mix obtained

after Subheading 3.8. At the end of the cleanup, elute the sample in 52.5 µl resuspension buffer and transfer 50 µl of the clear supernatant to a new tube. Perform an additional magnetic bead cleanup (refer to Subheading 3.3) using a ratio of 1:1. At the end of the cleanup, elute the sample in 22.5 µl resuspension buffer and transfer 20 µl of the clear supernatant to a new tube.

3.10 Normalization and Pooling

1. Check fragment sizes of each sample using a TapeStation. One main peak of the right size should be seen in addition to the lower and upper markers.
2. Check sample concentration using Qubit.
3. Dilute each indexed library to 10 ng/µl.
4. Pool 5 µl of each indexed library.

3.11 Library Validation Using Quantitative PCR

1. If the KAPA kit is used for the first time, add the 10× Primer Premix (1 ml) to the bottle of 2× KAPA SYBR FAST qPCR Master Mix and mix well with vortexer.
2. Prepare 1:5000, 1:10,000, and 1:20,000 dilutions of the library using Tris–Cl 10 nM buffer, pH 8.5 with 0.1% Tween 20. It is best to prepare three independent dilutions for each individual dilution (i.e., three independent 1/5000 dilutions). Here is an example on how to prepare dilutions:

Dilution	Library input	10 mM Tris–HCl (µl)
1:50	2 µl of undiluted library	98
1:5000	2 µl of 1/50 dilution	198
1:10,000	10 µl of 1/5000 dilution	10
1:20,000	10 µl of 1/5000 dilution	30

3. Determine the total number of reactions that will be performed for the appropriate number of replicates of each of the following reactions:
   - Three replicates of each of the six DNA standards.
   - Three replicates of each dilution to be assayed.
   - Three replicates of the no template control.
4. Prepare the Master Mix for the total number of reactions. For each 20 µl reaction, add:
   - 12 µl of 2× KAPA SYBR FAST Master Mix.
   - 4 µl of PCR-grade water.
5. Mix and briefly centrifuge the reagent Master Mix.
6. Dispense 16 µl of Master Mix to each well.
7. Add 4 µl of PCR-grade water to each well of the no template control.

8. Add 4 μl of each DNA standard to the appropriate well. Always start with low concentration standard.
9. Add 4 μl of each sample dilution to the appropriate well.
10. Seal the plate, centrifuge, and transfer to the qPCR instrument.
11. Setup experiment on the Applied Biosystems ViiA 7 real-time PCR system according to the manufacturer's instructions. In the menu, select the following options:
    - Fast 96-well block.
    - Standard curve.
    - SYBR mode.
    - Fast.
12. Set the reaction volume to 20 μl.
13. Perform qPCR using the following cycling conditions: 5 min at 95 °C (1×), 30 s at 95 °C, and 45 s at 60 °C (35×).
14. Calculate sample concentration in nM using the KAPA library Quantification Data Analysis Template for Illumina.
15. Dilute the library to 4 nM.

4 Notes

1. Size selection can be carried out by gel purification. An alternative better suited to high-throughput sequencing and low DNA concentrations is to use magnetic beads, as these will give better DNA recovery. Depending on the ratio of beads to sample, different size fragments can be purified [11]. In addition, selection of fragments to the left and right side of the desired fragment range can be carried out. Left side selection is done by binding the larger fragments to the right of the desired range to the beads and eluting the smaller fragments. For right size selection, the larger fragments to the right of the desired range are bound to the beads, and the supernatant containing the smaller fragments is removed to a fresh tube. For more details on this procedure *see* [12].
2. If an unwanted product is seen at 1000 bp, it can be due to bead carryover. To ensure that all magnetic beads are removed from the sample, an additional purification step can be carried out by placing the samples on the magnetic stand for 15 min and transferring the supernatant to a new tube.
3. If bead cleaning is performed in a magnetic plate, beads of the top layer might not efficiently bind to the magnet because of the large volume. To ensure maximum recovery, first pipette off approximately 100 μl of the lower clear layer and wait for an additional 5 min before pipetting the rest of the liquid.

Acknowledgments

Financial support to M.L. was provided by the Sant Chair and the Smithsonian Tennenbaum Marine Observatories Network, for which this is Contribution No. 5. This work was also supported by Swedish research council grant C0344601 to S.J.B.

References

1. Fonseca VG, Carvalho GR, Sung W, Johnson HF, Power DM, Neill SP, Packer M, Blaxter ML, Lambshead PJD, Thomas WK, Creer S (2010) Second-generation environmental sequencing unmasks marine metazoan biodiversity. Nat Commun 1:Artn 98. doi:10.1038/Ncomms1095
2. Ji YQ, Ashton L, Pedley SM, Edwards DP, Tang Y, Nakamura A, Kitching R, Dolman PM, Woodcock P, Edwards FA, Larsen TH, Hsu WW, Benedick S, Hamer KC, Wilcove DS, Bruce C, Wang XY, Levi T, Lott M, Emerson BC, Yu DW (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. Ecol Lett 16(10):1245–1257. doi:10.1111/Ele.12162
3. Leray M, Meyer CP, Mills SC (2015) Metabarcoding dietary analysis of coral dwelling predatory fish demonstrates the minor contribution of coral mutualists to their highly partitioned, generalist diet. Peerj 3:Artn e1047. doi:10.7717/peerj.1047
4. Aylagas E, Borja A, Rodriguez-Ezpeleta N (2014) Environmental status assessment using DNA metabarcoding: towards a genetics based marine biotic index (gAMBI). PLoS One 9(3), e90529. doi:10.1371/journal.pone.0090529
5. Leray M, Knowlton N (2015) DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proc Natl Acad Sci U S A 112(7):2076–2081. doi:10.1073/Pnas.1424997112
6. Visco JA, Apotheloz-Perret-Gentil L, Cordonier A, Esling P, Pillet L, Pawlowski J (2015) Environmental monitoring: inferring the diatom index from next-generation sequencing data. Environ Sci Technol 49(13):7597–7605. doi:10.1021/es506158m
7. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6(8):1621–1624. doi:10.1038/ismej.2012.8
8. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a Dual-Index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microb 79(17):5112–5120. doi:10.1128/AEM.01043-13
9. Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (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. Frontiers in zoology 10:Unsp 34. doi:10.1186/1742-9994-10-34
10. Lobo J, Costa PM, Teixeira MAL, Ferreira MSG, Costa MH, Costa FO (2013) Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. BMC Ecol 13:34. doi:10.1186/1472-6785-13-34
11. CoreGenomics. How do SPRI beads work? http://core-genomics.blogspot.co.uk/2012/04/how-do-spri-beads-work.html
12. BeckmanCoulter. http://www.beckmancoulter.com/wsrportal/bibliography?docname=SPRIselect.pdf