

# Chapter 10

## Targeted Gene Capture by Hybridization to Illuminate Ecosystem Functioning

Céline Ribière, Réjane Beugnot, Nicolas Parisot, Cyrielle Gasc, Clémence Defois, Jérémie Denonfoux, Delphine Boucher, Eric Peyretailade, and Pierre Peyret

### Abstract

Microbial communities are extremely abundant and diverse on earth surface and play key role in the ecosystem functioning. Thus, although next-generation sequencing (NGS) technologies have greatly improved knowledge on microbial diversity, it is necessary to reduce the biological complexity to better understand the microorganism functions. To achieve this goal, we describe a promising approach, based on the solution hybrid selection (SHS) method for the selective enrichment in a target-specific biomarker from metagenomic and metatranscriptomic samples. The success of this method strongly depends on the determination of sensitive, specific, and explorative probes to assess the complete targeted gene repertoire. Indeed, in this method, RNA probes were used to capture large DNA or RNA fragments harboring biomarkers of interest that potentially allow to link structure and function of communities of interest.

**Key words** Solution hybrid selection, Metagenomics, Metatranscriptomics, Microbial diversity, RNA probes, Next-generation sequencing

---

## 1 Introduction

Microbial communities show the greatest organisms diversity on earth and are key players for the functioning of all the ecosystems. For example, 1 g of soil may contain up to  $10^9$  bacterial cells [1] and assuming 3000 genes per single bacteria genome [2] and an average of 1000 bp per gene, such cells will thus represent up to  $3 \times 10^{15}$  bp. To explore such diversity, next-generation sequencing (NGS) technologies, especially Illumina systems, produce a great amount of sequence information (e.g. HiSeq 2500 produces six billion paired-end reads corresponding to 600 Gb of data). High-throughput sequencing greatly improved the resolution for microbial diversity description [3]. However, a substantial number of runs (6000) must be realized with a global cost of \$267 million to

produce a dataset representing onefold coverage of the microbial from 1 g of soil [4].

To reduce this biological complexity, barcoding is an efficient method [5] but cannot establish the link between the microbial structure and the realized functions limiting the understanding levels [6]. Furthermore, various PCR biases could alter these descriptions [7]. Recently, promising approaches, based on the SHS (Solution Hybrid Selection) capture method for the selective enrichment in a target-specific biomarker from metagenomic [8] and metatranscriptomic [9] (*see* Chapter 14 for metatranscriptomic application) samples have been developed (Fig. 1). First results have showed that this technology allows the identification of rare populations within the studied environment but also to participate to large DNA fragments reconstruction potentially allowing to link structure and function in microbial communities. The success of this innovating gene capture approach, however strongly depends on the determination of the best probe set while taking the biological question into account [10]. Consequently, capture probe design is of critical importance and should therefore consider multiple parameters in order to assess the complete targeted gene repertoire. In addition, to being sensitive and specific, probes must also anticipate genetic variations and thus must be able to detect known and unknown sequences in environmental samples. To design such explorative probes, three algorithms, PhylGrid, KASpOD and HiSpOD have been developed. PhylGrid is a large-scale probe design software linked to the EGI (European Grid Infrastructure) [11]. It is an improvement of the PhylArray algorithm presented in Milton et al. [12] which relies on initial multiple sequence alignments to define regular and explorative oligonucleotide probes for SSU rRNA genes. KASpOD is a web service dedicated to the design of signature sequences using a k-mer-based algorithm [13, 14]. PhylGrid and KASpOD software were used to define 74,003 probes of 25 mer targeting SSU rRNA genes from 2178 genera including Bacteria and Archaea. These probes are available using PhylOPDb, an online resource for a comprehensive phylogenetic oligonucleotide probe database [15]. Finally, the HiSpOD program allows designing both gene-specific and sequence-specific probes to target any functional biomarker [16]. All these software, developed in the context of microbial ecology, are then particularly appropriate for the design of highly sensitive, specific, and explorative probes in the context of gene capture by hybridization. Indeed, these probes, used for the gene capture molecular approach described below will ensure the selective enrichment of DNA or RNA (*see* Chapter 14 for metatranscriptomic application) from targeted phylogenetic or functional biomarker genes of interest in complex environments.

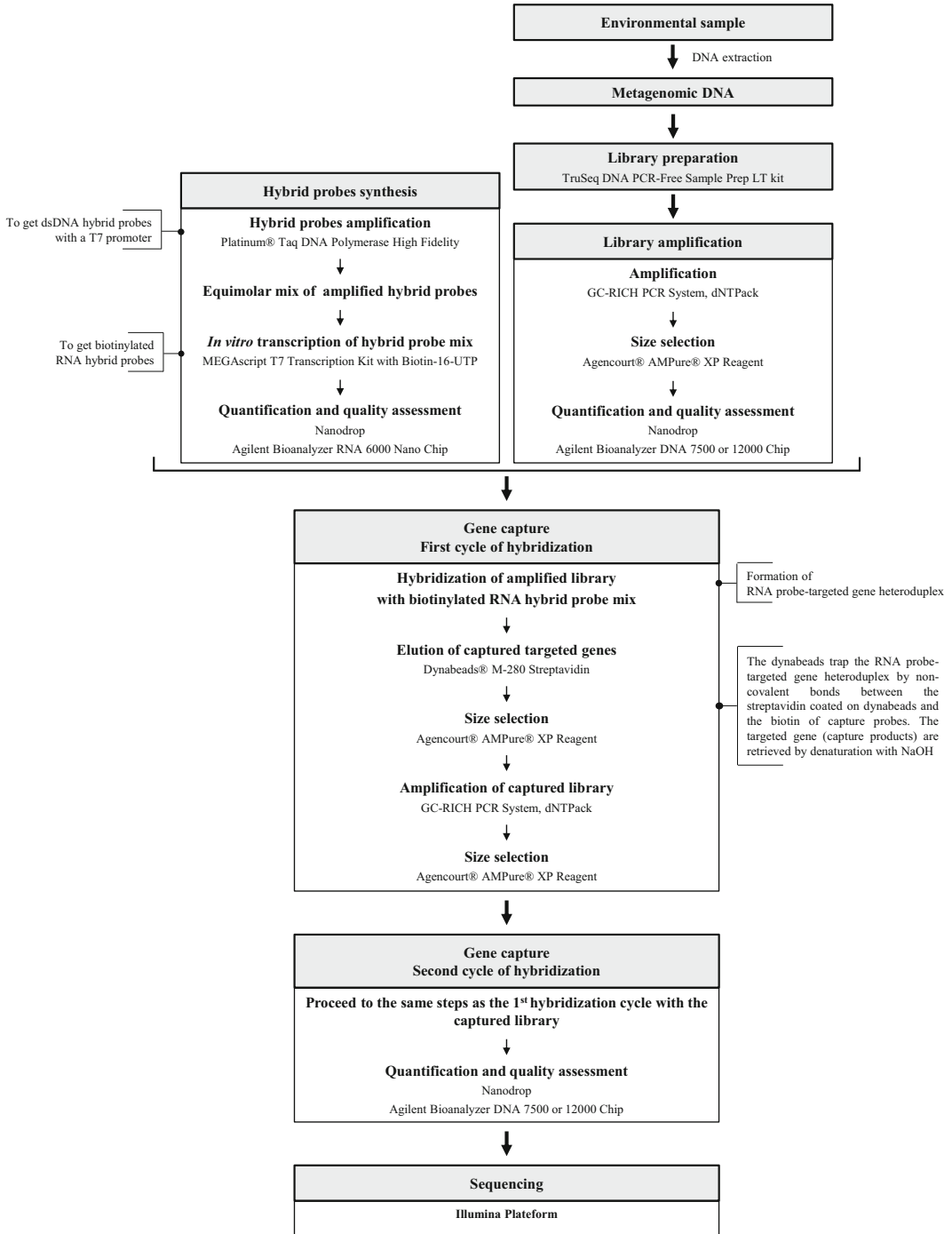


Fig. 1 Process workflow of targeted gene capture by hybridization

---

## 2 Materials

### 2.1 Reagents and Kits

1. Agencourt® AMPure® XP Reagent (Beckman Coulter, Brea, CA, USA).
2. Agilent DNA 7500 or 12000 Kit (Agilent Technologies, Santa Clara, CA, USA).
3. Agilent RNA 6000 Nano Kit (Agilent Technologies, Santa Clara, CA, USA).
4. Biotin-16-UTP (Epicentre, Madison, WI, USA).
5. Dynabeads® M-280 Streptavidin (Life Technologies, Carlsbad, CA, USA).
6. GC-RICH PCR System, dNTPack (Roche Applied Science, Basel, Switzerland).
7. Glycogen (molecular biology grade).
8. MEGAscript® T7 Transcription Kit (Life Technologies, Carlsbad, CA, USA).
9. MinElute Gel Extraction Kit (Qiagen, Hilden, Germany).
10. MinElute PCR Purification Kit (Qiagen, Hilden, Germany).
11. Platinum® Taq DNA Polymerase High Fidelity (Life Technologies, Carlsbad, CA, USA).
12. QIAquick PCR Purification Kit (Qiagen, Hilden, Germany).
13. RNeasy Plus Mini Kit (Qiagen, Hilden, Germany).
14. 10 mg/mL sheared salmon sperm DNA (Life Technologies, Carlsbad, CA, USA).
15. TruSeq DNA PCR-Free Sample Prep LT Set A or Set B (Illumina, San Diego, CA, USA).

### 2.2 Buffers and Solutions

All buffers and solutions could be prepared in laboratory under DNase/RNase-free conditions or purchased in general lab supplier. Prepare all solutions using ultrapure water (prepared by purifying deionized water to attain a sensitivity of 18 M $\Omega$  cm at 25 °C) and molecular biology grade reagents. Prepare and store all solutions and buffers at room temperature (unless indicated otherwise).

1. 100× Denhardt's solution: 2 % bovine serum albumin (BSA) (Fraction V), 2 % Ficoll 400, 2 % polyvinylpyrrolidone. Weigh 1 g BSA, 1 g Ficoll 400 and 1 g polyvinylpyrrolidone and transfer to a 50 mL graduated cylinder. Add water to a volume of 50 mL. Filter through a 0.2  $\mu$ m syringe filter to sterilize. Divide into aliquots of 2 mL, and store at -20 °C.
2. 0.5 M ethylenediaminetetraacetic acid (EDTA): pH 8.0. Weigh 46.53 g EDTA and transfer to a 250 mL graduated cylinder. Add water to a volume of 150 mL. Mix and adjust pH with

NaOH. Make up to 250 mL with water. Sterilize by autoclaving (*see Note 1*).

3. 80 % Ethanol (*see Note 2*).
4. 5 M NaCl. Weigh 73.08 g NaCl and transfer to a 250 mL graduated cylinder. Add water to a volume of 250 mL. Sterilize by autoclaving.
5. 1 M NaOH. Weigh 1 g NaOH and transfer into the plastic beaker. Add water to a volume of 25 mL. Stir vigorously and as precaution, place the beaker on ice (*see Note 2*).
6. 0.1 M NaOH. Make a dilution at  $1/10^6$  in nuclease-free water of 1 M NaOH solution (*see Note 2*).
7. 3 M sodium acetate: pH 5.2. Weigh 40.83 g sodium acetate and transfer to a 100 mL graduated cylinder. Add water to a volume of 80 mL. Adjust pH with glacial acetic acid. Make up to 100 mL with water. Filter through a 0.2  $\mu\text{m}$  syringe filter to sterilize.
8. 10 % sodium dodecyl sulfate (SDS). Weigh 10 g SDS and transfer to a 100 mL graduated cylinder. Add water to a volume of 85 mL. Heat to 68 °C and stir with a magnetic stirrer to assist dissolution. Adjust pH to 7.2 by adding a few drop of concentrated HCl (36 %). Make up to 100 mL with water. Filter through a 0.2  $\mu\text{m}$  syringe filter to sterilize.
9. 10 $\times$  Tris-Borate-EDTA (TBE): 450 mM Tris-borate, 10 mM EDTA. Weigh 108 g Tris base, 27.5 g boric acid and transfer to a 1 L graduated cylinder. Add 10 mL of 0.5 M EDTA (pH 8.0) and water to a volume of 800 mL. Mix and adjust pH to 8. Make up to 1 L with water. Sterilize by autoclaving.
10. 10 $\times$  Tris-EDTA (TE): 100 mM Tris-HCl, 10 mM EDTA. Weigh 0.24 g Tris base and 0.23 g EDTA and transfer to a 50 mL graduated cylinder. Add water to a volume of 35 mL. Adjust pH to 7.5 with HCl. Make up to 50 mL with water. Sterilize by autoclaving.
11. 1 M Tris-HCl: pH 7.5. Weigh 60.57 g Tris base and transfer to a 500 mL graduated cylinder. Add water to a volume of 350 mL. Adjust pH to 7.5 with HCl. Make up to 500 mL with water. Sterilize by autoclaving.
12. 20 $\times$  SSC: 3 M NaCl, 0.3 M trisodium citrate. Weigh 17.53 g NaCl and 8.82 g trisodium citrate and transfer to a 100 mL graduated cylinder. Add water to a volume of 80 mL. Adjust pH to 7.0 by adding HCl. Make up to 100 mL with water. Sterilize by autoclaving.
13. 20 $\times$  SSPE: 3 M NaCl, 0.2 M  $\text{NaH}_2\text{PO}_4$ , 0.02 M EDTA. Weigh 17.53 g NaCl, 2.76 g  $\text{NaH}_2\text{PO}_4$ , and 0.74 g EDTA and transfer to the cylinder. Add water to a volume of 80 mL. Adjust pH to 7.4 with NaOH. Make up to 100 mL with water. Sterilize by autoclaving.

14. Binding buffer: 1 M NaCl, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA. Transfer 10 mL of 5 M NaCl, 500  $\mu$ L of 1 M Tris-HCl (pH 7.5) and 100  $\mu$ L of 0.5 M EDTA (pH 8.0) to the cylinder. Make up to 50 mL with water. Filter through a 0.2  $\mu$ m syringe filter to sterilize (*see Note 2*).
15. 2 $\times$  Hybridization buffer: 10 $\times$  SSPE, 10 $\times$  Denhardt's solution, 10 mM EDTA, 0.2 % SDS. Transfer 10 mL of 20 $\times$  SSPE, 2 mL of 100 $\times$  Denhardt's solution, 400  $\mu$ L of 0.5 M EDTA (pH 8.0), and 400  $\mu$ L of 10 % SDS to a 20 mL graduated cylinder. Make up to 20 mL with water. Filter through a 0.2  $\mu$ m syringe filter to sterilize. Divide into aliquots of 2 mL, and store at -20 °C.
16. Wash buffer n°1: 1 $\times$  SSC, 0.1 % SDS. Transfer 2.5 mL of 20 $\times$  SSC and 500  $\mu$ L of 10 % SDS to a 50 mL graduated cylinder. Make up to 50 mL with water. Filter through a 0.2  $\mu$ m syringe filter to sterilize (*see Note 2*).
17. Wash buffer n°2: 0.1 $\times$  SSC, 0.1 % SDS. Transfer 250  $\mu$ L of 20 $\times$  SSC and 500  $\mu$ L of 10 % SDS to a 50 mL graduated cylinder. Make up to 50 mL with water. Filter through a 0.2  $\mu$ m syringe filter to sterilize (*see Note 2*).

### 2.3 Oligonucleotides (Probes and Primers)

1. Hybrid probes. Purchase hybrid probes at 100  $\mu$ M. Adaptor sequences must be added to the 5' and 3' ends of the specific capture probes. These hybrid probes consist of 5'-ATCGCA CCAGCGTGT(X)CACTGCGGCTCCTCA-3', with X indicating the specific capture probe.
2. Primers for probe amplification. Purchase oligonucleotides at 100  $\mu$ M, T7-A 5'-GGATTCTAATACGACTCACTATAGG GATCGCACCAGCGTGT-3' and B 5'-CGTGGATGAGGAG CCGCAGTG-3'.
3. Primers for library amplification. Purchase oligonucleotides at 100  $\mu$ M, TS-PCR Oligo 1 5'-AATGATACGGCGACCAC CGAGA-3' and TS-PCR Oligo 2 5'-CAAGCAGAAGACGGCA TACGAG-3'.

### 2.4 Equipments

1. Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).
2. AFA System (Covaris, Woburn, MA, USA). One of these following items: M220, S220, S2 or E210 Focused-Ultrasonicator with the corresponding AFA Tubes (*see Note 3*).
3. DynaMag<sup>TM</sup>-2 Magnet (Life Technologies, Carlsbad, CA, USA).
4. HulaMixer<sup>®</sup> Sample Mixer (Life Technologies, Carlsbad, CA, USA) (optional).
5. Nanodrop spectrophotometer (Thermo Scientific, Wilmington, DE, USA) or other systems for DNA quantification (e.g.:

Qubit® 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA) or other fluorometers).

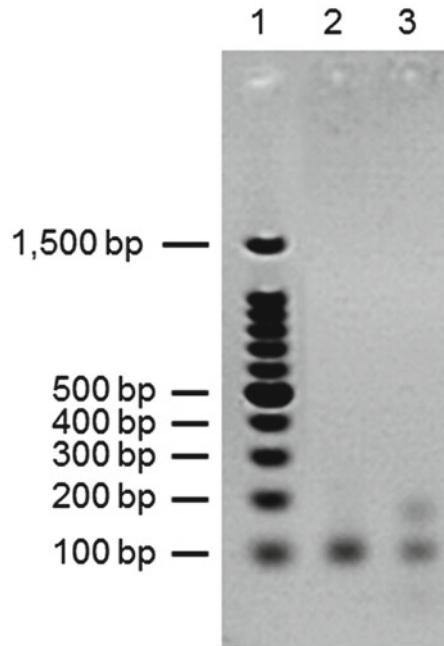
6. Speed vacuum.
7. Thermal cycler (with heated lid).

---

### 3 Methods

#### 3.1 Hybrid Probe Synthesis

1. First step of hybrid probe synthesis consists in amplification of oligonucleotide to obtain double-stranded DNA (dsDNA). Each amplification reaction should contain 5  $\mu$ L of 10 $\times$  high fidelity buffer, 1  $\mu$ L of dNTPs (10 mM), 2  $\mu$ L of MgSO<sub>4</sub> (50 mM), 1  $\mu$ L of primer T7-A (10  $\mu$ M), 1  $\mu$ L of primer B (10  $\mu$ M), 0.2  $\mu$ L of Platinum® Taq DNA polymerase high fidelity, 38.8  $\mu$ L of nuclease-free water and 1  $\mu$ L of hybrid probe diluted at 10  $\mu$ M (*see Note 4*). Include a negative control with 1  $\mu$ L of nuclease-free water instead of 1  $\mu$ L of hybrid probe. Use a thermal cycler with the following conditions: 2 min at 94 °C then 35 cycles of 30 s at 94 °C, 30 s at 58 °C and 20 s at 68 °C and a final elongation step at 68 °C for 5 min.
2. Check the probe amplification by electrophoresis on a 2 % agarose-TBE gel containing 0.5 $\times$  syber safe (or comparable nucleic acid stain). Deposit 5  $\mu$ L of amplified product (with loading buffer) (*see Note 5*). One lane is reserved for 100 bp DNA ladder. The gel is run in TBE buffer at 100 V for 45 min. The DNA is visualized on a UV transilluminator.
3. If one band at the expected size is observed, proceed to the purification of the remaining 45  $\mu$ L of amplified products using the MinElute PCR Purification Kit, following the manufacturer's instructions. If two amplification bands are observed, deposit the remaining PCR product (i.e. 45  $\mu$ L), excise with a clean razor blade or scalpel the band corresponding to the size of hybrid probes and proceed to their purification using the MinElute Gel Extraction Kit, following the manufacturer's instructions. The purified product is eluted in 15  $\mu$ L of nuclease-free water (Fig. 2).
4. Evaluate the concentration of purified amplified hybrid probes with Nanodrop spectrophotometer.
5. For RNA synthesis, mix all hybrid probes in an equimolar amount taking into account the degeneracy of each probe. Each probe combination must be present in the same molecular amount (*see Note 6*). Validate the concentration of hybrid probe mix with Nanodrop spectrophotometer. Take 150 ng of hybrid probe mix, evaporate to dryness with a speed vacuum and resuspend in 4.75  $\mu$ L of nuclease-free water. If the 150 ng

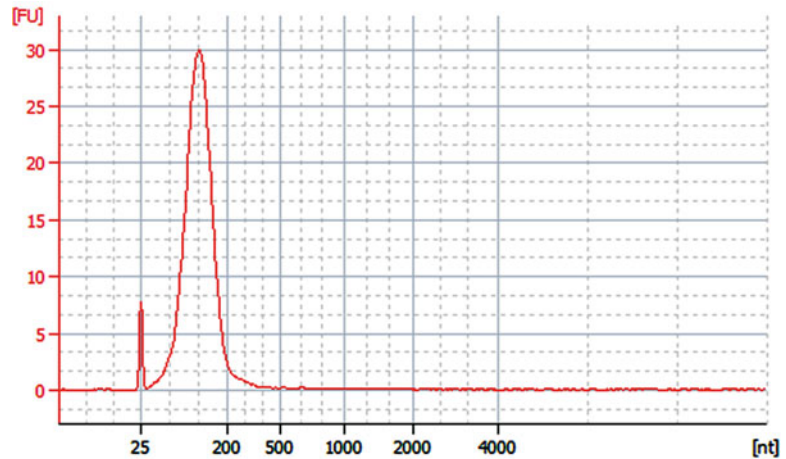


**Fig. 2** Amplified hybrid probes electrophoresis on a 2 % agarose-TBE gel. *Lane 1* shows DNA ladder (100 bp). *Lane 2* shows one amplification band at the expected size, in this case at 112 bp corresponding to a hybrid probe of 50 bp with amplification primers T7-A (41 bp) and B (21 bp). *Lane 3* shows two amplification bands, one at the expected size (same as *lane 2*) and another one (at 150 bp) corresponding to aberrant amplification of the T7 promoter

of hybrid probe mix is in a volume lower than 4.75  $\mu\text{L}$ , do not evaporate and adjust the volume to 4.75  $\mu\text{L}$  with nuclease-free water.

6. The *in vitro* transcription (IVT) is realized with the MEGAscript<sup>®</sup> T7 Transcription Kit and using Biotin-16-UTP to produce biotinylated RNA. Each IVT reaction should contain 2  $\mu\text{L}$  of 10 $\times$  reaction buffer, 2  $\mu\text{L}$  of ATP solution (75 mM), 2  $\mu\text{L}$  of CTP solution (75 mM), 2  $\mu\text{L}$  of GTP solution (75 mM), 1.5  $\mu\text{L}$  of UTP solution (75 mM), 3.75  $\mu\text{L}$  of biotin-16-UTP (10 mM), 2  $\mu\text{L}$  of T7 enzyme mix, and the 4.75  $\mu\text{L}$  of previously prepared hybrid probe mix (*see Notes 4 and 7*). Incubate at 37  $^{\circ}\text{C}$  for at least 6 h (or overnight).
7. Add 1  $\mu\text{L}$  of TURBO DNase (include in the MEGAscript<sup>®</sup> T7 Transcription Kit) to each IVT reaction and incubate at 37  $^{\circ}\text{C}$  for 30 min.
8. For RNA precipitation, transfer the IVT reaction mix in a 1.5 mL microcentrifuge tube. Add 1/10<sup>e</sup> volume of 3 M sodium acetate (pH 5.2), 3 volumes of cold 100 % ethanol, and 1  $\mu\text{L}$  of glycogen (20  $\mu\text{g}/\mu\text{L}$ ). The reaction is incubated at





**Fig. 3** Quality of biotinylated RNA probes assessed on Agilent Bioanalyzer RNA 6000 Nano chip. The electropherogram shows a resolved peak at the expected size of a hybrid probe of 50 bp with amplification primers T7-A (41 bp) and B (21 bp) (same probe as Fig. 2)

–80 °C for 30 min. Centrifuge at  $18,000\times g$  for 15 min at 4 °C. Discard the supernatant and wash two times the pellet as following: add 500  $\mu$ L of cold 70 % ethanol, centrifuge at  $18,000\times g$  for 10 min at 4 °C and discard the supernatant. Dry the pellet with a speed vacuum. Add 100  $\mu$ L of TE for pellet resuspension.

9. Proceed to the purification of biotinylated RNA probe mix with the RNeasy Plus Mini Kit following the “Appendix D: Purification of Total RNA Containing Small RNAs from Cells” instructions excepted the step D2 with the gDNA Eliminator spin column. Make two RNeasy Mini spin columns per hybrid probe mix (apply 50  $\mu$ L of biotinylated RNA probe mix on each column), elute the product in 40  $\mu$ L of nuclease-free water and pool the two obtained eluates.
10. Evaluate the concentration of purified biotinylated RNA probe mix with Nanodrop spectrophotometer. Assess their quality on an Agilent Bioanalyzer RNA 6000 Nano chip, according to the manufacturer’s instructions (Fig. 3).
11. Store at –80 °C.

### 3.2 Library Preparation (550 bp Insert)

The library is prepared for 550 bp insert using the TruSeq DNA PCR-Free Sample Prep LT kit by Illumina following the manufacturer’s instructions.

### 3.3 Library Amplification

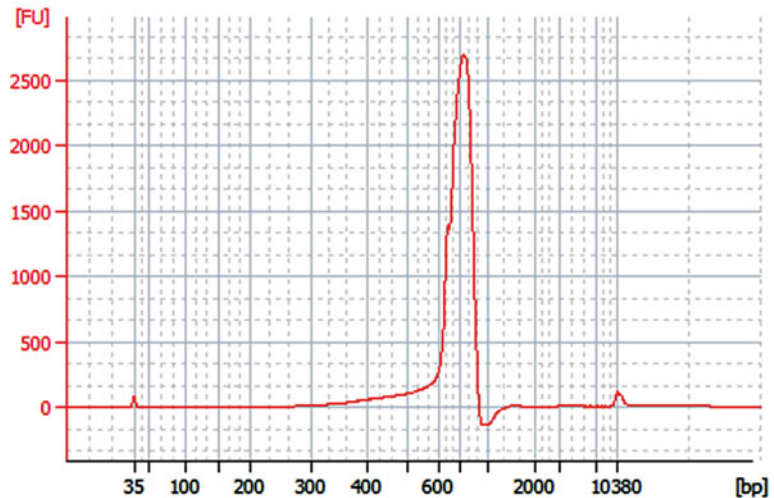
1. Add 30  $\mu$ L of nuclease-free water to the library.
2. Proceed to the library amplification with the GC-RICH PCR System, dNTPack. Realize ten 50  $\mu$ L PCR reactions per library.

Each amplification reaction should contain 10  $\mu\text{L}$  of 5 $\times$  GC-RICH PCR reaction buffer, 2  $\mu\text{L}$  of 25 mM  $\text{MgCl}_2$ , 1  $\mu\text{L}$  of PCR grade nucleotide mix, 1  $\mu\text{L}$  of 25  $\mu\text{M}$  TS-PCR Oligo 1, 1  $\mu\text{L}$  of 25  $\mu\text{M}$  TS-PCR Oligo 2, 29  $\mu\text{L}$  of PCR grade water, 1  $\mu\text{L}$  of GC-RICH enzyme mix and 5  $\mu\text{L}$  of prepared library (*see Note 4*). Use the following thermal conditions: 4 min at 94  $^\circ\text{C}$  then 20 cycles of 30 s at 94  $^\circ\text{C}$ , 1 min at 58  $^\circ\text{C}$  and 1 min 30 s at 68  $^\circ\text{C}$  and a final elongation step at 68  $^\circ\text{C}$  for 3 min.

3. Purify the amplified library using the QIAquick PCR Purification Kit following the manufacturer's instructions. Use one column of the kit for two PCR reactions pooled from a same library (i.e. five columns per library). The purified product is eluted in 50  $\mu\text{L}$  of nuclease-free water.
4. Select the DNA fragments size with the Agencourt<sup>®</sup> AMPure<sup>®</sup> XP Reagent. Check that the eluate volume is equal to 50  $\mu\text{L}$ . If necessary make up to 50  $\mu\text{L}$  with nuclease-free water. Add 50  $\mu\text{L}$  of AMPure beads, gently mix by pipetting and incubate for 5 min at room temperature (*see Note 8*). Place the tubes on the magnetic stand for at least 5 min at room temperature (until the supernatant is clear). Remove and discard the supernatant from each tube. Keep the tubes on the magnetic stand and wash two times the beads like following: add 500  $\mu\text{L}$  of 80 % ethanol to each tube without disturbing them, incubate for 30 s at room temperature, and then remove and discard all of the supernatant from each tube. Take care not to disturb the beads. Remove and discard any remaining ethanol with a 10  $\mu\text{L}$  pipette and let the beads air-dry for 5 min at room temperature. Add 50  $\mu\text{L}$  of nuclease-free water to each tube. Remove the tubes from the magnetic stand. By pipetting, resuspend the beads by repeatedly dispensing the water over the bead pellet until it is immersed in the solution. Incubate for 2 min at room temperature. Place the tubes on the magnetic stand for at least 5 min at room temperature (until the supernatant is clear). Transfer all of the supernatant from each of the five tubes into a new 1.5 mL microcentrifuge tube (*see Note 9*).
5. Evaluate the concentration of purified amplified library with NanoDrop spectrophotometer. Assess its quality on an Agilent DNA 7500 or 12000 chip, according to the manufacturer's instructions (Fig. 4).
6. Store the purified amplified libraries at  $-20\text{ }^\circ\text{C}$ .

### 3.4 Gene Capture by Hybridization

1. Transfer 2.5  $\mu\text{g}$  of sheared salmon sperm DNA and 500 ng of purified amplified library into a 0.2 mL PCR tube (*see Note 10*). Evaporate to dryness with a speed vacuum and resuspend in 7  $\mu\text{L}$  of nuclease-free water.
2. Thaw an aliquot of 2 $\times$  hybridization buffer, prewarmed it at 65  $^\circ\text{C}$  and transfer 20  $\mu\text{L}$  into a 0.2 mL PCR tube (*see Note 11*).



**Fig. 4** Quality of amplified library (prepared for 650 bp insert) assess on an Agilent Bioanalyzer High Sensitivity DNA chip. The electropherogram shows a peak focused on 770 bp for the case of library prepared for 650 bp insert with 120 bp Illumina adaptors. With a library prepared for 550 bp, the same profile will be observed but with a peak focused on 670 bp

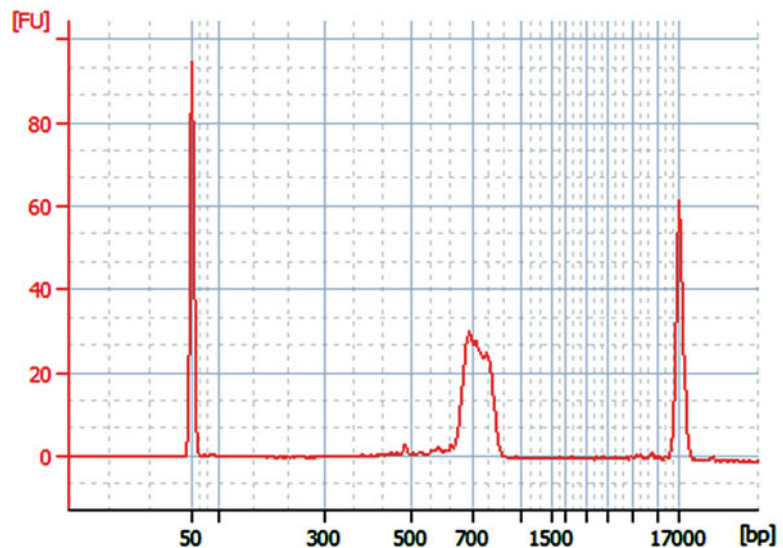
3. Thaw the biotinylated RNA probe mix on ice, transfer 500 ng into a 0.2 mL PCR tube and adjust the volume to 6  $\mu$ L with nuclease-free water (*see* **Notes 11** and **12**).
4. Incubate the salmon sperm DNA/purified amplified library (SL) mix in a thermal cycler with the following conditions: 95  $^{\circ}$ C for 5 min and 65  $^{\circ}$ C at 5 min.
5. Without removing SL mix from the thermal cycler, incubate at 65  $^{\circ}$ C the 0.2 mL PCR tube with 2 $\times$  hybridization buffer. Add quickly 13  $\mu$ L of prewarmed 2 $\times$  hybridization buffer to the tube containing the SL mix and homogenize by pipetting.
6. Always without removing SL mix from the thermal cycler, incubate at 65  $^{\circ}$ C the 0.2 mL PCR tube with the biotinylated capture probes mix. Add quickly 6  $\mu$ L of probe mix to SL mix (hybridization mix) and homogenize by pipetting. Incubate at 65  $^{\circ}$ C for the obtained hybridization mix 24 h in the thermal cycler.
7. Prior to removing the hybridization mix from the thermal cycler, prepare the Dynabeads<sup>®</sup> M-280 Streptavidin as following: transfer 50  $\mu$ L of dynabeads into a 1.5 mL microcentrifuge tube (*see* **Note 11**), place the tube on the magnetic stand until the supernatant is clear, remove and discard it. Wash three times the dynabeads as following: add 200  $\mu$ L of binding buffer, gently tap the tube to resuspend the dynabeads, place the tube on the magnetic stand until the supernatant is clear, remove and discard the supernatant. Take care not to disturb

the dynabeads. After the three washes, resuspend the dynabeads in 200  $\mu\text{L}$  of binding buffer.

8. Add the 26  $\mu\text{L}$  of hybridization mix to the washed dynabeads. Gently tap the tube to resuspend the dynabeads and incubate for 30 min at room temperature (off the magnetic stand). Regularly resuspend the dynabeads during the incubation by gently taping the tube (*see Note 13*).
9. During the incubation, pre-warm the wash buffer n°2 at 65 °C (at least 1.5 mL per captured library).
10. After the incubation, place the tube on the magnetic stand until the supernatant is clear. Remove and discard the supernatant. Take care not to disturb the dynabeads. Add 500  $\mu\text{L}$  of wash buffer n°1 and resuspend the dynabeads by gently taping the tube. Incubate for 15 min at room temperature (off the magnetic stand). Regularly resuspend the dynabeads during the incubation by gently taping the tube (*see Note 13*).
11. After the incubation, place the tube on the magnetic stand until the supernatant is clear, remove and discard it. Take care not to disturb the dynabeads. Wash three times the dynabeads as following: resuspend the dynabeads in 500  $\mu\text{L}$  of pre-warmed wash buffer n°2, incubate for 10 min at 65 °C (off the magnetic stand). Regularly resuspend the dynabeads during the incubation by gently taping the tube. Place the tube on the magnetic stand until the supernatant is clear. Remove and discard the supernatant. Take care not to disturb the dynabeads.
12. Resuspend the dynabeads in 50  $\mu\text{L}$  of 0.1 M NaOH by vortexing the tube for 5 s (*see Note 2*). Incubate for 10 min at room temperature (off the magnetic stand).
13. Place the tube on the magnetic stand until the supernatant is clear and transfer it to a 1.5 mL microcentrifuge tube containing 70  $\mu\text{L}$  of 1 M Tris-HCl (pH 7.5). Take care not to disturb the Dynabeads.
14. Purify the captured library using the QIAquick PCR Purification Kit following the manufacturer's instructions. The purified product is eluted in 50  $\mu\text{L}$  of nuclease-free water.
15. Select the DNA fragment size with the Agencourt® AMPure® XP Reagent as indicated in **step 4** of Subheading **3.3**.
16. Amplify the captured library using the GC-RICH PCR System, dNTPack. Make five 50  $\mu\text{L}$  PCR reactions per captured library and proceed in the same way as the **step 2** of Subheading **3.3** but realize 25 amplification cycles instead of 20 in the thermal conditions.
17. Purify the amplified captured library using the QIAquick PCR Purification Kit following the manufacturer's instructions. Realize one column for 2.5 amplification reactions from a same

library (i.e. two columns per library (125  $\mu$ L)). The purified product is eluted in 50  $\mu$ L of nuclease-free water.

18. Select the DNA fragment size with the Agencourt® AMPure® XP Reagent as indicated in **step 4** of Subheading **3.3**.
19. Evaluate the concentration of purified amplified captured library with Nanodrop spectrophotometer.
20. Proceed to a second cycle of hybridization by repeating the **steps 1–15** with the purified amplified capture products obtained previously (*see Notes 14 and 15*).
21. Amplify the captured library using the GC-RICH PCR System, dNTPack. Make ten PCR reactions per library and proceed in the same way as the **step 2** of Subheading **3.3** but realize 25 amplification cycles instead of 20 in the thermal conditions.
22. Purify the amplified captured library using the QIAquick PCR Purification Kit following the manufacturer's instructions. Realize one column for two amplification reactions from a same library (i.e. five columns per library). The purified product is eluted in 50  $\mu$ L of nuclease-free water.
23. Select the DNA fragment size with the Agencourt® AMPure® XP Reagent as indicated in **step 4** of Subheading **3.3**.
24. Evaluate the concentration of purified amplified captured library with Nanodrop spectrophotometer. Assess its quality on an Agilent DNA 7500 or 12000 chip, according to the manufacturer's instructions (Fig. 5).



**Fig. 5** Quality of captured library (prepared for 650 bp insert) assess on an Agilent Bioanalyzer 12000 DNA chip. The electropherogram shows a peak focused on 770 bp for the case of library prepared for 650 bp insert with 120 bp Illumina adaptor. With a library prepared for 550 bp, the same profile will be observed but with a peak focused on 670 bp

25. Store the purified amplified captured library products at  $-20^{\circ}\text{C}$ .
26. Proceed to the sequencing of captured library on an Illumina sequencer compatible with the kit use to prepare the library.

---

## 4 Notes

1. EDTA will not go into solution until the pH of the solution is adjusted to  $\sim 8.0$  by the addition of 1 M NaOH.
2. Buffers and solutions must be extemporaneously prepared.
3. For the fragmentation of gDNA, we recommend, as Illumina, to use Covaris microTUBES with a focused-ultrasonicator. Covaris offers different models of focused-ultrasonicator and the material necessary for the fragmentation depends on the focused-ultrasonicator used.
4. Include 10 % excess for multiple samples.
5. The verification of the amplified hybrid probe size by electrophoresis on a 2 % agarose-TBE gel is absolutely necessary. It is possible to get two amplification bands, one at the expected size (i.e. size of hybrid probe with amplification primers T7-A and B) and another due to aberrant amplification of the T7 promoter. Only the correct band at the expected size must be excised and purified.
6. For example, three hybrid probes (specific capture probe and adaptor sequences) are necessary for the mix with specific capture probe as following: probe A 5'-CCCAGGATWAGATACCKCCYAGTTTAYRC-3', probe B 5'-TTCAGAAGTAGATATGCTGGTAGTCTACCA-3' and probe C 5'-TGCACAATCAGATAGTYTGGYAGTGACCGC-3'. The probe A has one W base (two combinations, T or A), one K base (two combinations, T or G), two Y bases (two combinations, C or T), and one R base (two combinations, A or G), therefore there is 32 possible combinations for the probe A ( $1\text{ W} \times 1\text{ K} \times 2\text{ Y} \times 1\text{ R} = 2 \times 2 \times (2 \times 2) \times 2 = 32$ ). In the same way, the probes B and C have a degeneracy of 1 (none degenerated base) and 4 ( $2 \times 2$ ) respectively. To make the hybrid probe mix, use  $1/32^{\text{e}}$  in quantity of the probe B and  $1/8^{\text{e}}$  in quantity of the probe C compared to the probe A. Thus, for 500 ng of probe A, add 15.6 ng of probe B and 62.5 ng of probe C. In this example, the three probes have the same size, otherwise you must calculate a number of molecules (taking into account the size of each probe) to realize an equimolar mix.
7. The IVT reaction mix must be realized at room temperature. Indeed the spermidine in the  $10\times$  reaction buffer can coprecipitate the template DNA if the reaction is assembled on ice.

8. Remove the AMPure beads from 2 to 8 °C storage and let stand for at least 30 min to bring them to room temperature. Vortex the room temperature AMPure beads for at least 1 min or until they are well dispersed. Vortex the beads frequently and collect them by slowly pipetting (due to their viscosity) to make sure that they are evenly distributed.
9. Be sure not to retrieve beads with the supernatant because the beads can interfere with the following steps. It is better to retrieve less volume of clear supernatant and if necessary to adjust the volume by adding nuclease-free water to the purified product.
10. The sheared salmon sperm DNA commercial solution is concentrated at 10 mg/mL. Make a dilution at 1/10<sup>e</sup> in nuclease-free water and pipette 2.5 µL to get 2.5 µg of sheared salmon sperm DNA.
11. Make as many aliquots as captured library.
12. It could be necessary to dilute the biotinylated RNA probe mix. The 500 ng of the biotinylated RNA probe mix should be in a volume less than or equal to 6 µL.
13. Alternatively, incubate the appropriate time at room temperature with tilting and gentle rotation on a HulaMixer Sample Mixer or other similar device.
14. If you do not plan to proceed immediately to the second cycle of hybridization, the protocol can be safely stopped here. If you are stopping, store the tubes at -20 °C.
15. If the quantity of purified amplified capture products, obtained after the first cycle of hybridization, is less than 500 ng, proceed to the second cycle of hybridization with the totality of the sample.

## References

1. Schloss PD, Handelsman J (2006) Toward a census of bacteria in soil. *PLoS Comput Biol*. doi:[10.1371/journal.pcbi.0020092](https://doi.org/10.1371/journal.pcbi.0020092)
2. Vieites JM, Guazzaroni M-E, Beloqui A et al (2009) Metagenomics approaches in systems microbiology. *FEMS Microbiol Rev* 33:236–255
3. Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Next-generation sequencing technologies for environmental DNA research. *Mol Ecol* 21:1794–1805
4. Desai N, Antonopoulos D, Gilbert JA et al (2012) From genomics to metagenomics. *Curr Opin Biotechnol* 23:72–76
5. Klindworth A, Pruesse E, Schweer T et al (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res*. doi:[10.1093/nar/gks808](https://doi.org/10.1093/nar/gks808)
6. Allen EE, Banfield JF (2005) Community genomics in microbial ecology and evolution. *Nat Rev Microbiol* 3:489–498
7. Hong S, Bunge J, Leslin C et al (2009) Polymerase chain reaction primers miss half of rRNA microbial diversity. *ISME J* 3:1365–1373
8. Denonfoux J, Parisot N, Dugat-Bony E et al (2013) Gene capture coupled to high-throughput sequencing as a strategy for targeted metagenome exploration. *DNA Res* 20:185–196
9. Bragalini C, Ribière C, Parisot N et al (2014) Solution hybrid selection capture for the recovery of functional full-length eukaryotic cDNAs

- from complex environmental samples. *DNA Res* 21:685–694
10. Dugat-Bony E, Peyretailade E, Parisot N et al (2012) Detecting unknown sequences with DNA microarrays: explorative probe design strategies. *Environ Microbiol* 14:356–371
  11. Jaziri F, Peyretailade E, Missaoui M et al (2014) Large scale explorative oligonucleotide probe selection for thousands of genetic groups on a computing grid: application to phylogenetic probe design using a curated small subunit ribosomal RNA gene database. *Sci World J*. doi:[10.1155/2014/350487](https://doi.org/10.1155/2014/350487)
  12. Militon C, Rimour S, Missaoui M et al (2007) PhylArray: phylogenetic probe design algorithm for microarray. *Bioinformatics* 23:2550–2557
  13. Parisot N, Denonfoux J, Dugat-Bony E et al (2012) KASpOD—a web service for highly specific and explorative oligonucleotide design. *Bioinformatics* 28:3161–3162
  14. Parisot N, Denonfoux J, Dugat-Bony E et al (2014) Software tools for the selection of oligonucleotide probes for microarrays. In: He Z (ed) *Current technology, innovations and applications*. Academic, New York
  15. Jaziri F, Parisot N, Abid A et al (2014) PhyLOPDb: a 16S rRNA oligonucleotide probe database for prokaryotic identification. Database. doi:[10.1093/database/bau036](https://doi.org/10.1093/database/bau036)
  16. Dugat-Bony E, Missaoui M, Peyretailade E et al (2011) HiSpOD: probe design for functional DNA microarrays. *Bioinformatics* 27:641–648