Library preparation and gene capture

PROTOCOL FOR:

Capturing protein-coding genes across highly divergent species

Chenhong Li^{1,3}, Michael Hofreiter², Nicolas Straube³, Shannon Corrigan³, Gavin J P Naylor³ ¹College of Fisheries and Life Science, Shanghai Ocean University, Shanghai, China, ²Department of Biology, The University of York, Heslington, York, UK, and ³Hollings Marine Laboratory, College of Charleston, Charleston, SC

BioTechniques 54:321-326 (June 2013) doi 10.2144/000114039

Keywords: comparative biology; cross-species capture; DNA hybridization capture; next-generation sequencing; protein-coding genes; phylogenetics; target enrichment

Reagents

- " DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA)
- Agencourt AMPure XP beads (Beckman Coulter Inc, Atlanta, GA)
 Dynabeads M-270 streptavidin beads (Life Technologies, Grand Island New York)
- " Polyethylene Glycol (Life Technologies)
- " **Tween-20** (Life Technologies)
- ["] 5 M NaCl (Life Technologies)
- " **1 M Tris-Cl, pH 8.0** (Life Technologies)
- " **0.5 M EDTA, pH 8.0** (Life Technologies)
- " 10× Buffer Tango (Fermentas, Thermo Fisher Scientific, Waltham, MA)
- ²⁵ **mM dNTPs** (Fermentas, Thermo Fisher Scientific)
- " 100 mM ATP (Fermentas, Thermo Fisher Scientific)
- ["] 50% PEG-4000 (Fermentas, Thermo Fisher Scientific)
- " T4 DNA ligase (Fermentas, Thermo Fisher Scientific)
- " 10× T4 DNA ligase buffer (Fermentas, Thermo Fisher Scientific)
- " 10× ThermoPol reaction buffer (New England Biolabs, Ipswich, MA)
- " Bst polymerase, large fragment (New England Biolabs)
- **T4 polynucleotide kinase** (New England Biolabs)
- " T4 DNA polymerase (New England Biolabs)
- ² 2× KAPA HiFi Taq Ready Mix (KAPA Biosystems, Inc., Woburn, MA)
- KAPA Library Quantification Kit (KApA Biosystems, Inc.)
- * **MYBaits Target Enrichment Kit** (MYcroarray, Ann Arbor, Michigan) including:

Block#1/Human Cot-1 (1 μg/μL) HYB#1/20× SSPE HYB#2/0.5 M EDTA HYB#3/50× Denhardt's HYB#4/1% SDS SUPERase•In (20 U/μl) RNA Baits BindingBuffer (1 M NaCl, 10 mM Tris-HCl, 1mM EDTA pH75) Wash Buffer 1 (1× SSC, 0.1% SDS) Wash Buffer 2 (0.1× SSC, 0.1% SDS)

- " 50 µM Adapter mix see recipes
- " 10× Oligo hybridization buffer see recipes
- " Adapter: 500 µM IS1_adapter_P5.F (Integrated DNA Technologies, Inc. Coralville, Iowa)
- wAdapter: 500 μM IS3_adapter_P5+P7.R (Integrated DNA Technologies, Inc.)
- " Indexing primer: 10 µM Primer IS4_indPCR.P5 (Integrated DNA Technologies, Inc.)
- " Primer: 10 µM Primer IS7_short_amp.P5 (Integrated DNA Technologies, Inc.)
- " Primer: 10 µM Primer IS8_short_amp.P7 (Integrated DNA Technologies, Inc.)
- [™] Blocking oligo for adapter P5: 200 μM BO1.P5short.F (Integrated DNA Technologies, Inc.)
- [™] Blocking oligo for part 1 of adapter P7: 200 μM BO3.P7.part1.F (Integrated DNA Technologies, Inc.)

Sequences of adapters, primers, and blocking oligos

Name	Sequences
IS1_adapter.P5	a*c*a*c*TCTTTCCCTACACGACGCTCTTCCg*a*t*c*t
IS2_adapter.P7	g*t*g*a*CTGGAGTTCAGACGTGTGCTCTTCCg*a*t*c*t
IS3_adapter.P5+P7	a*g*a*t*CGGAa*g*a*g*c
IS4_indPCR.P5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACAC- GACGCTCTT
IS7_short_amp.P5	ACACTCTTTCCCTACACGAC
IS8_short_amp.P7	GTGACTGGAGTTCAGACGTGT
B01.P5short.F	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-Pho
BO3.P7.part1.F	AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC-Pho

* indicates a phosphothioate oligonucleotide (PTO) bond. Pho indicates a 3'-phosphate.

Legend ⇒*ATTENTION*

* HINT

Procedure

1. Library preparation using the with-bead method

Genomic DNA was extracted from freshly preserved muscle tissue for all species using the DNeasy Blood and Tissue Kit, according to the manufacturer's instructions. Extracted genomic DNA was subsequently sheared to ~500 bp using a Covaris M220 Focusedultrasonicator before library preparation. The protocol used for library preparation is based on the Meyer and Kircher (1) method for preparing Illumina sequencing libraries for multiplexed target capture and sequencing, with modifications to accommodate capture of divergent species. The with-bead method of Fisher et al. (2) is also adopted, meaning that the beads are kept in each of the DNA clean up steps. The samples are not eluted from the beads until after the library preparation. Thus, all library preparation steps occur on the beads. This modification to the regular Solid Phase Reversible Immobilization (SPRI) method can increase the complexity of the resulting library.

I. Shearing genomic DNA

Skip the shearing and size selection step if using samples with fragmented DNA, e.g. ancient DNA or DNA extracted from museum samples.

- Start with 0.5–3 μg genomic DNA and shear it to ~500 bp using the Covaris machine according to the manufacturer's instructions (130 μl sample in Covaris microTUBE; temperature: 18 to 22°C; treatment: 60 s, Peak Power: 50, Duty Factor: 10, Cycles/Burst: 200).
- 2. Check the size distribution of the sheared DNA on an agarose gel to confirm accuracy of the shearing process and appropriate fragment size.

II. Size selection

*If the sheared DNA has a broad size distribution, a size selection step may be applied using the SPRI bead method at different concentrations of polyethylene glycol (PEG). The following protocol is used to select DNA with fragment sizes > 250 bp. Skip the size selection step if using samples with fragmented DNA < 100 bp.

- 1. Add 50 μ L Agencourt AMPure XP beads to a 200 μ L tube. Use individual tubes to avoid cross contamination. Dry the beads for two min by placing the tubes on a magnetic plate, allowing the beads to move to the walls of the tube before removing the supernatant.
- 2. Add 50 μ L sheared sample and 37.5 μ L 20% PEG to the dried beads. Prepare 1 positive and 1 negative control: Add 50 μ L AMPure XP beads to a tube with 30 μ L nuclease free water and to another tube with 30 μ L positive DNA (1:100 diluted PCR product of any DNA fragment with a size of 100–200 bp). Pipette up and down several times to mix the samples.
- 3. Incubate at room temperature for 10 min. Collect the liquid at the bottom of the tube by brief centrifugation. Place the tube back on the magnetic plate, and let it stand for 5–10 min to separate the beads from the solution. Pipette off and discard the supernatant without removing the beads.
- 4. Leaving the tube on the magnetic rack, wash the beads by adding $186 \,\mu$ L of freshly prepared 70% ethanol. Let stand for 1 min and remove the supernatant. Keep the tube on the magnetic rack, do not disturb the beads.
- 5. Repeat step 4 one more time, for a total of two washes.
- 6. Remove residual traces of ethanol. Let the beads air-dry for 5 min at room temperature without caps.
- 7. Proceed to the next step immediately.

III. Blunt-end repair

*After shearing, overhanging 5'- and 3'-ends are repaired by T4 DNA polymerase and 5'-phosphates are attached using T4 polynucleotide kinase.

⇒ If working with ancient or fragmented (<100 bp) DNA, skip the size selection step and concentrate the samples using a speed vac. Elute the concentrated DNA in 16.32 µL nuclease free water. Replace the water in the recipe below, adding the entire concentrated sample in its place. After the blunt-end repair, skip step 1.111.3. Instead, incubate the reaction at 75°C for 20 min to deactivate the enzymes, followed by a ramp step decreasing to 12°C at the rate of -1°C/s. Immediately proceed to the ligation step.

1. Prepare a master mix for the required number of samples. Add 20 μ L of the master mix to each sample. Mix the samples well by pipetting.

Reagent	Volume (µL) per sample	Final concentration in 20 µL reaction
Buffer Tango (10×)	2	1×
dNTPs (25 mM each)	0.08	100 µM each
ATP (100 mM)	0.2	1 mM
T4 polynucleotide kinase (10 U/µL)	1	0.5 U/µL
T4 DNA polymerase (5 U/µL)	0.4	0.1 U/µL
H ₂ O	16.32	

- 2. Incubate the samples in a thermal cycler for 15 min at 25°C followed by 5 min at 12°C.
- 3. Add 15 μL 20% PEG to the sample and clean up the reaction according to the SPRI protocol. Keep the dried beads.

IV. Adapter ligation

*P5 and P7 adapters (2) are ligated to the ends of the repaired molecules using T4 DNA ligase.

⇒ If working with ancient or fragmented (<100 bp) DNA, add 9 μ L instead of 29 μ L of water in step 1.IV.1, then add 19 μ L of master mix to each sample.

1. Prepare a master mix for the required number of samples. Add 39 μ L of the master mix to each sample tube. Mix the samples well by pipetting:

Reagent	Volume (µL) per sample	Final concentration in 40 µL reaction
T4 DNA ligase buffer (10×)	4	1×
PEG-4000 (50%)	4	5%
Adapter mix (50 μ M each)	2	1.25 µM each
H ₂ O	29	

- 2. Spin down the liquid by brief centrifugation. Add 1 μ L T4 DNA ligase (5 U/ μ L) to the sample. Pipette up and down to mix the sample and collect the liquid at the bottom of the tube by brief centrifugation, then incubate for 30 min at 22°C in a thermal cycler.
- 3. Clean up the reaction using the SPRI method (adding $30 \,\mu L$ 20% PEG for a regular library, or $80 \,\mu L$ 20% PEG for library with short inserts such as ancient or fragmented DNA). Keep the dried beads. Proceed immediately to the next step.

V. Fill-in

*Adapters are non-phosphorylated and thus ligate to only one of the template strands. Resulting single-strand nicks are filled in using Bst polymerase to allow amplification of the insert.

1. Prepare a master mix for the required number of samples.

Reagent	Volume (µL) per sample	Final concentration in 40 µL reaction
ThermoPol reaction buffer (10×)	4	1×
dNTPs (25 mM each)	0.4	250 µM each
Bst polymerase, large fragment (8 U/ μ L)	1.5	0.3 U/µL
H ₂ O	34.1	

- 2. Add 40 μ L of master mix to the samples. Pipette up and down to mix the sample and collect the liquid at the bottom of the tube by brief centrifugation. Incubate the samples for 20 min at 37°C.
- 3. Clean up the samples using the SPRI method (adding 30 μ L 20% PEG for regular libraries or 80 μ L 20% PEG for libraries with short inserts such as ancient or fragmented DNA). Elute the samples with 20 μ L of nuclease free water. Transfer the supernatant to a new tube labeled accordingly. The libraries can be kept frozen at -20°C for a short period, but we recommend to further process the libraries as soon as possible to avoid degradation of DNA.

VI. Pre-hybridization PCR

*Libraries are amplified using primers IS7_short_amp.P5 and IS8_short_amp.P7 prior to gene capture and indexing

1. Prepare a master mix for the required number of samples.

Reagent	Volume (µL) per sample	Final concentration in 50 µL reaction
KAPA HiFi Taq Ready Mix (2×)	25	1×
Primer IS7_short_ amp.P5 (10 μM)	1.5	0.3 µM
Primer IS8_short_ amp.P7 (10 μM)	1.5	0.3 µM
H ₂ O	16	

- 2. Add 44 μ L of master mix to empty tubes, and then add 6 μ L of the cleaned libraries from step 1.V.3. Mix well and amplify the samples using the following thermal profile: 98°C for 45 s, 12 to 18 cycles of 98°C for 15 s, 65°C for 30 s, and 72°C for 45 s, followed by 72°C for 1 min, and hold at 4°C for 10 min. The number of cycles can be adjusted according to the amount of starting material used to construct the library.
- 3. Clean up the PCR product using the SPRI bead method (adding 30 μ L 20% PEG). Elute the DNA using 20 μ L of nuclease free water and transfer it to a new tube labeled accordingly. Measure the concentration of the amplified library using a Qubit 2.0 Fluorometer or similar. The concentration should be around 6–10 ng/ μ L or higher if more starting material was used.

2. Gene capture

We applied two sets of conditions for hybridization and washing: the standard hybridization conditions [hybridization at 65°C for 36 h; 1st wash at room temperature $(18-22^{\circ}C)$; 2nd wash at 65°C] and the relaxed hybridization conditions [hybridization under a touchdown scheme (Figure 1): 65°C for 11 h, followed by 60°C for 11 h, 55°C for 11 h, and 50°C for 11 h; 1st wash at room temperature $(18-22^{\circ}C)$; 2nd Wash at 45°C]. The protocol used is based on the user manual of the MYBaits target enrichment kit.

I. Hybridization

*Amplified target DNA libraries are hybridized to biotinylated RNA baits.

- 1. Set the following program on a thermal cycler: 95°C for 5 min, 65°C for 3 min, 65°C for 2 min, 65°C for 11 h, 60°C for 11 h, 55°C for 11 h, 50°C for 11 h, and hold at 50°C.
- 2. Prepare a Library Master Mix for the required number of samples:

*Block #1/Human Cot-1 binds and blocks repetitive elements in the target DNA to prevent non-specific binding to the baits. The blocking oligos are used to blanket the adapters and prevent adapter to adapter hybridization.

Reagent	Volume (µL) per sample
Block#1/Human Cot-1 (1 µg/µL)	2.5
B01.P5short.F (200 µM)	0.08
B03.P7.part1.F (200 µM)	0.08

Add 2.5 μ L of Library Master Mix from step 2.I.2 to empty 200 μ L tubes, and then add 6.5 μ L of target DNA library from step 1.VI.3 to each tube and label accordingly. Mix by pipetting. Collect the liquid at the bottom of the tube by brief centrifugation. Set aside until step 2.I.5.

3. Prepare a Hybridization Master Mix for the required number of samples:

*This master mix contains various buffers to aid successful hybridization of baits to the target DNA library.

Reagent	Volume (µL) per sample
HYB#1/20× SSPE	10
HYB#2/0.5 M EDTA	0.4
HYB#3/50× Denhardt's	4
HYB#4/1% SDS	4

Mix the reagents by pipetting, and collect the liquid at the bottom of the tube by brief centrifugation. Label accordingly and set aside until 2.I.5.

Prepare a Bait Master Mix for the required number of samples:
 **This master mix contains the RNA baits and an RNase inhibitor*

Reagent	Volume (µL) per sample
SUPERase•In (20U/µI)	1
RNA Baits	0.5
H ₂ O	4.5

Mix the reagents by pipetting. Label accordingly and set aside until 2.I.5.

- 5. Transfer the tube(s) containing the Library Master Mix from step 2.I.2. to the thermocycler and start the program set in step 2.I.1. This will denature the DNA library for 5 min at 95°C.
- 6. Once the thermocycler program reaches step 2 (65°C, see step 2.I.1), transfer the tube containing the Hybridization Master Mix from step 2.I.3 to the thermocycler. This will pre-warm the Hybridization Master Mix for 3 min at 65°C. Leave the Library Master Mix in the thermocycler at this time.
- 7. Once the thermocycler program reaches step 3 (65°C, see step 2.I.1), transfer the tube containing the Baits Master Mix from step 2.I.4 to the thermocycler. This will pre-warm the Bait Master Mix for 2 min at 65°C. Leave the Library Master Mix and the Hybridization Master Mix in the thermocycler at this time.
- While keeping tubes at 65°C, transfer 13 μl of Hybridization Master Mix and 7 μl of Library Master Mix to the tube(s) containing the Baits Master Mix and mix via pipetting.
- 9. Close the lid and incubate the hybridization solution in the thermocycler for 48 h under the touchdown scheme described in step 2.1.1. This scheme represents the optimized conditions for capturing across divergent species. Depending on the application, hybridization time may require some optimization.



Figure 1. Schematic representation of the 'touchdown' hybridization procedure described in steps 2.1.5 to 2.1.9.

II. Binding hybridized libraries to beads and wash

*The resulting biotinylated target-bait complex is bound to streptavidin coated beads. A low stringency washing step is used to remove unbound material. This is followed by higher stringency washing to remove bound non-target DNA. We have optimized the temperatures and number of washes for the higher stringency wash, which is important to capture divergent sequences.

- 1. Add $n \times 10 \,\mu\text{L}$ (*n* is the number of samples) of Dynabeads M-270 streptavidin beads to a 200 μ L or 1.5 mL tube according to the volume of the mixture.
- 2. Pellet the beads for two min using a magnetic particle stand and discard the supernatant.
- 3. Add 200 μL MYBaits target enrichment kit Binding Buffer (1M NaCl, 10mM Tris-HCl, 1mM EDTA pH 7.5) at room temperature to beads to wash. Vortex tube for 5–10 s, place on magnetic particle stand for 2 min to pellet the beads and remove and discard supernatant.
- 4. Repeat step 3 twice for a total of three washes.
- 5. Resuspend the beads in $n \times 20 \,\mu\text{L}$ Binding Buffer, add 1 μL 10% Tween.
- 6. For each sample, add 180 μ L Binding Buffer to an empty 200 μ L tube, and then add 20 μ L resuspended beads to it.
- 7. Transfer the hybridization solution(s) from step 2.I.9 to the tube(s) containing the Binding Buffer and beads, and incubate for 30 min at room temperature on a rotator. Collect the liquid at the bottom of the tube by brief centrifugation. Pellet the beads with the magnetic plate for two min and remove supernatant.
- 8. Add 186 μ L MYBaits target enrichment kit Wash Buffer 1 (1× SSC, 0.1% SDS) to the beads and pipette up and down to resuspend. Incubate for 10 min at room temperature. Collect the liquid at the bottom of the tube by brief centrifugation. Pellet beads with magnetic particle stand for two min and remove supernatant. Repeat step 2.II.8 one more time for a total of two low stringency washes. In the meantime, preheat MYBaits target enrichment kit Wash Buffer 2 (0.1× SSC, 0.1% SDS) to 45°C.
- 9. Add 186 ml 45°C Wash Buffer 2 (0.1× SSC, 0.1% SDS) to the beads and pipette up and down to mix the sample. Incubate for 10 min at 45°C. Pellet beads with the magnetic plate for two min and remove supernatant.
- 10. Repeat step 2.II.9 2 times for a total of 3 higher stringency washes at 45°C. After the last wash make sure all additional buffer is removed.
- 11. Add 50 μL nuclease free water to the beads and label accordingly.

III. Post-hybridization indexing PCR (off-beads amplification)

*The captured library is recovered and amplified using off-beads amplification (1), i.e. the captured target is amplified off the targetbait-bead complex during the indexing PCR. This avoids the need for chemical denaturation and maximizes retention of captured products.

→ If performing double capture, amplify the captured library off the beads as in 2.111.1 but replace the indexing primers with primers IS7_ short_amp.P5 and IS8_short_amp.P7 (as in step 1.VI.1). Clean up the amplified libraries using the SPRI method. Use the product as template to repeat steps 2.1.1 through 2.11.11. After the second round of capture, perform the post hybridization indexing PCR as described in 2.111.1 and continue the protocol to completion.

1. Prepare a master mix for the required number of samples.

Reagent	Volume (µL) per sample	Final concentration in 50 µL reaction
KAPA HiFi Taq Ready Mix (2×)	25	1×
Primer IS4_indPCR. P5 (10 µM)	0.5	0.2 µM

- 2. Add 25.5 μ L of master mix, 23 μ L well-mixed sample (including streptavidin beads) from step 2.11.11, and 0.5 μ L indexing primer (10 μ M) to each tube. Mix well and amplify the samples using the following thermal profile: 98°C for 45 s, 12 to 18 cycles of 98°C for 15 s, 65°C for 30 s and 72°C for 45 s, followed by 72°C for 1 min, and hold at 4°C for 10 min. The number of PCR cycles can be adjusted according to the starting material used to construct the library.
- 3. Load 3 μ L of PCR product to a mini agarose gel to check the size of the captured library. The band should be barely visible.
- 4. Clean up the PCR product using the SPRI bead method. Elute the DNA using 20 μL of nuclease free water, transfer it to a new tube and label accordingly.

IV. Pooling multiple samples for sequencing

- 1. Determine the DNA concentration of indexed captured libraries using Qubit 2.0 Fluorometer or similar. The concentration of the samples should be around $0.1-0.9 \text{ ng/}\mu\text{L}$, or higher if more starting material was used.
- 2. Pool all samples in equimolar ratios for sequencing.
- 3. Clean up the pooled sample with SPRI beads adding 3/4 of the sample volume of 20% PEG. Elute the sample with 20 μ L nuclease free water.
- 4. Quantify the pooled library using qPCR. We utilized the KAPA Library Quantification Kit. The pooled library should be 20 μ L final volume at a concentration of > 10 nM (2nM-50nM), which is 2.6 ng/ μ L (0.5 ng/ μ L-13 ng/ μ L) for DNA ~400 bp.

Recipes

Oligo hybridization buffer (10×):

Reagent	Volume	Final concentration in 10 ml
NaCI (5 M)	1 ml	500 mM
Tris-Cl, pH 8.0 (1 M)	100 µL	10 mM
EDTA, pH 8.0 (0.5 M)	20 µL	1 mM
H ₂ O	8.88 ml	

Adapter Mix (50 µM):

Hybridization mix for adapter P5 (100 µM):

Reagent	Volume (µL)	Final concentration in 100 µL reaction
IS1_adapter_P5.F (500 µM)	20	100 µM
IS3_adapter_P5+P7.R (500 µM)	20	100 µM
Oligo hybridization buffer (10×)	10	1×
H _* O	50	

Hybridization mix for adapter P7 (100 µM):

Reagent	Volume (µL)	Final concentration in 100 µL reaction
IS2_adapter_P7.F (500 µM)	20	100 µM
IS3_adapter_P5+P7.R (500 µM)	20	100 µM
Oligo hybridization buffer (10×)	10	1×
H ₂ O	50	

Mix and incubate both reactions in a thermal cycler for 10 s at 95°C, followed by a ramp from 95°C to 12°C at a rate of -0.1°C/s. Combine both reactions to obtain a ready-to-use adapter mix (50 μ M each adapter).

Equipment

- " Covaris M220 Focused-ultrasonicator (Covaris, Inc., Woburn, MA)
- " **Qubit 2.0 Fluorometer** (Life Technologies)
- " Eppendorf Mastercycler pro S (Hauppauge, New York)
- " CFX Connect Real-Time PCR system (Bio-Rad, Hercules, CA)
- " MiSeq sequencer (Illumina, Inc, San Diego, CA)

References

- 1. Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010(6): pdb prot5448.
- 2. Fisher S, Barry A, Abreu J, Minie B, Nolan J, Delorey TM, Young G, Fennell TJ, Allen A, Ambrogio L et al. 2011. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. Genome Biol 12(1): R1.