Supporting Information Appendix

DNA Capture Reveals Trans Oceanic Gene Flow in Endangered River Sharks

Chenhong Li, Shannon Corrigan, Lei Yang, Nicolas Straube, Mark Harris, Michael Hofreiter, William White, Gavin J P Naylor.

This document contains the following supporting information:

Supporting Figures:

Fig. S1. Statistical parsimony network based on mitochondrial genome sequences. Large dots indicate observed sequences with colors denoting nominal species. Small red dots represent inferred sequences. Numbers along branches show the number of substitutions between nodes on the network.

Fig. S2. Frequency distribution of pairwise p-distance among unique mitochondrial genome sequences.

Fig. S3. Phylogenetic relationships among the major *Glyphis* lineages based on partitioned Maximum Likelihood analysis of 100 nuclear exons.

Fig. S4. Species tree based on SNAPP analysis of 1,041 nuclear SNP loci.

Fig. S5. Agarose gels showing DNA ladder (GeneRuler 50 bp Thermo Fisher Scientific, Massachusetts, USA) cleaned up with different amounts of PEG (lane descriptions are listed in the below table). Left – 2% agarose gel; right – 4% agarose gel.

Supporting Tables:

Table S1. Specimens of *Glyphis* plus outgroups sampled in this study.

Table S2. Sample ID and mitochondrial genome sequencing statistics for *Glyphis* and the outgroup species.

Table S3. Number of SNPs retrieved for each taxon.

Table S4. Primers used to assess the quality of extracted DNA and long-range PCR primers for bait preparation.

Detailed Protocol for Preparing Homemade Baits

Detailed Protocol for Target Capture

DNA Clean-Up Using Solid Phase Reversible Immobilization (SPRI)

References

Supporting Figures

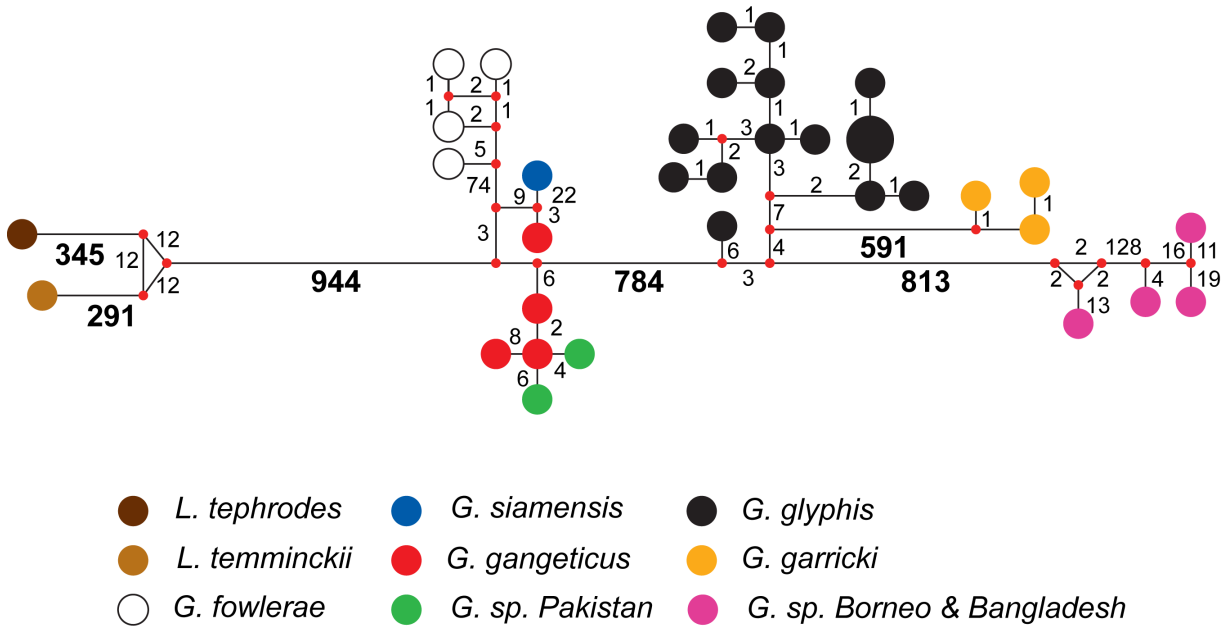

Fig. S1. Statistical parsimony network based on mitochondrial genome sequences. Large dots indicate observed sequences with colors denoting nominal species. Small red dots represent inferred sequences. Numbers along branches show the number of substitutions between nodes on the network.

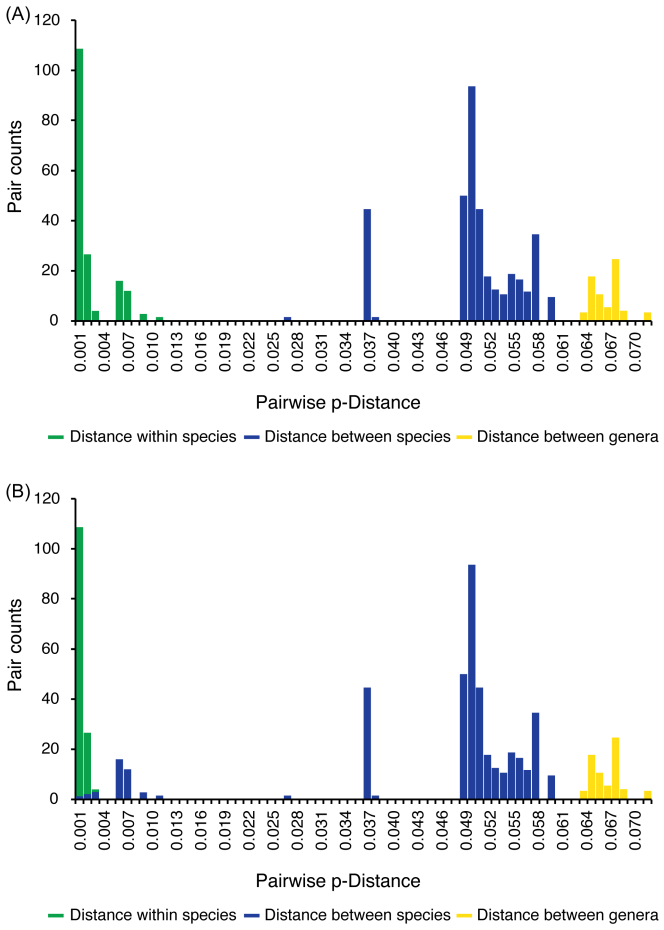

Fig. S2. Frequency distribution of pairwise p-distance among unique mitochondrial genome sequences. A) *G. gangeticus*, *G. siamensis*, *G. sp*. Pakistan and *G. fowlerae* synonymized. B) All nominal species are considered valid.

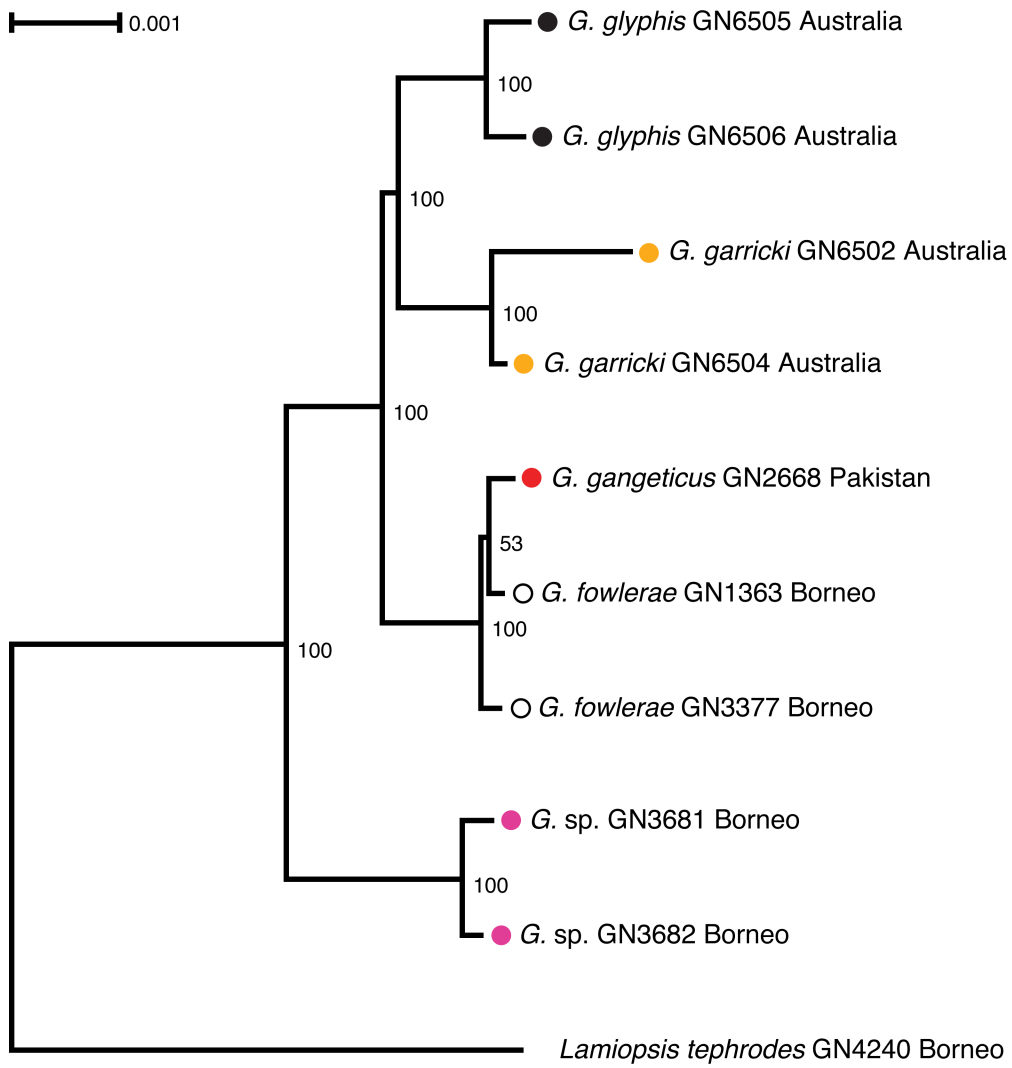

Fig. S3. Phylogenetic relationships among the major *Glyphis* lineages based on partitioned Maximum Likelihood analysis of 100 nuclear exons. Bootstrap support values are indicated on the branches. Colors correspond to those in Figs. 1 and 2.

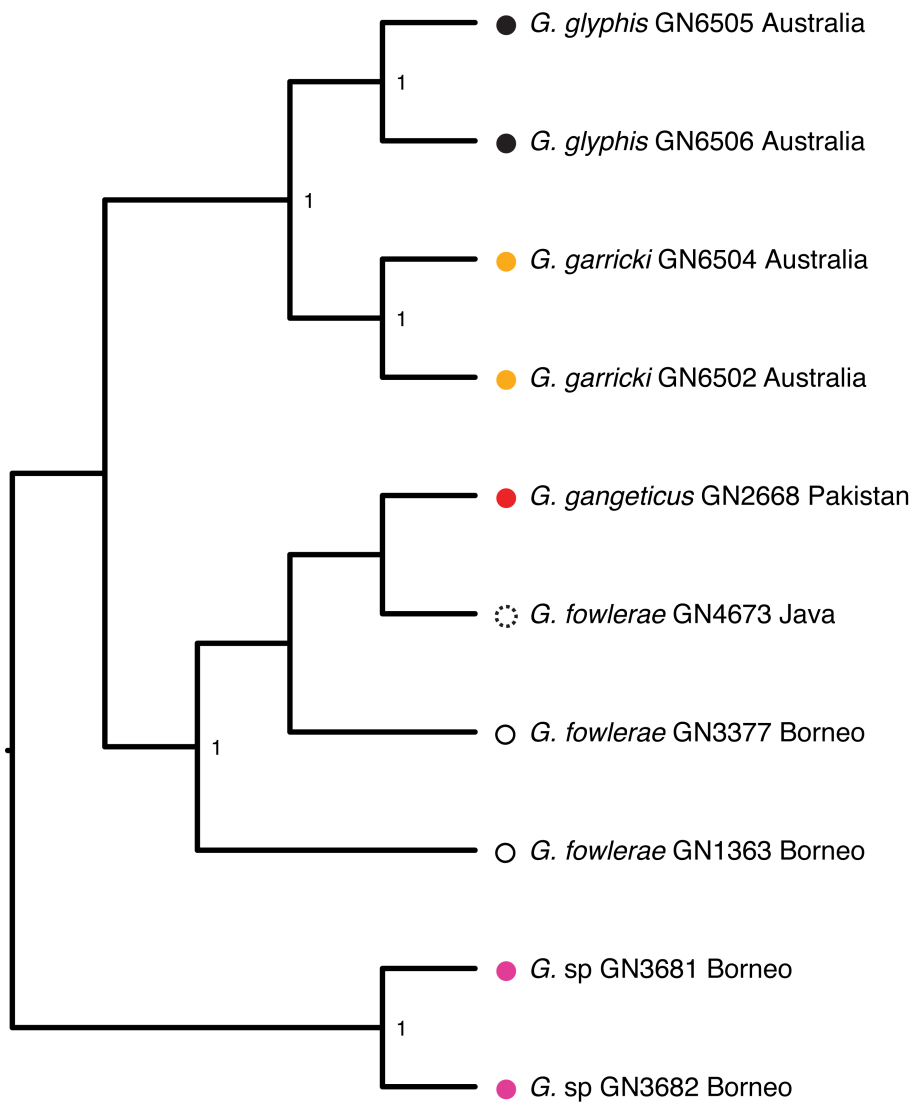

Fig. S4. Species tree based on SNAPP analysis of 1,041 nuclear SNP loci. Bayesian posterior probabilities larger than 0.9 are indicated on the branches. Colors correspond to those in Figs. 1 and 2.

Supporting Tables

Table S1. Specimens of *Glyphis* plus outgroups sampled in this study.

Sample ID	Species	Collection sites	Type specimen ID
GN6921*	*G. glyphis*	Unknown	ZMB 5265
GN6505	*G. glyphis*	Alligator River, Northern territory, Australia	
GN6506	*G. glyphis*	Alligator River, Northern territory, Australia	
GN15749	*G. glyphis*	Katatai, Papua New Guinea	
GN6923*	*G. gangeticus*	Ganges River, Hooghly, India	ZMB 4474
GN1188	*G. gangeticus*	India	
GN2668	*G. gangeticus*	Pakistan	
GN2669	*G. gangeticus*	Pakistan	
GN12170*	*G. siamensis*	Irrawaddy River, Rangoon, Myanmar	NMW 61397
GN10999*	*G. garricki*	Alligator River, Northern territory, Australia	CSIRO H 5262-01
GN6502	*G. garricki*	Alligator River, Northern territory, Australia	
GN6503	*G. garricki*	Alligator River, Northern territory, Australia	
GN6504	*G. garricki*	Alligator River, Northern territory, Australia	
GN3376*	*G. fowlerae*	Kinabatangan River, Sabah, Malaysian Borneo	IPMB 38.14.02
GN3377	*G. fowlerae*	Kinabatangan River, Sabah, Malaysian Borneo	
GN1363	*G. fowlerae*	Kinabatangan River, Sabah, Malaysian Borneo	
GN4673	*G. fowlerae*	Cirebon, Java, Indonesia	
GN3830	*G. sp.*	Bangladesh	
GN2673	*G. sp.*	Bangladesh	
GN3681	*G. sp.*	Mukah, Sarawak, Malaysian Borneo	
GN3682	*G. sp.*	Mukah, Sarawak, Malaysian Borneo	
GN2667	*G. sp.*	Pakistan	
GN2671	*G. sp.*	Pakistan	
GN1670	*Lamiopsis temminckii*	India	
GN4240	*Lamiopsis tephrodes*	South Kalimantan, Indonesian Borneo	
GN4239	*Lamiopsis tephrodes*	South Kalimantan, Indonesian Borneo	

*Type specimens are indicated by an asterisk.

Table S2. Sample ID and mitochondrial genome sequencing statistics for *Glyphis* and the outgroup species.

Taxon	Number of	Reads on target	Number of	Mitogenome Coverage	Average	Read length
G. glyphis GN6921*	10914976	6467571 (59%)	259051	100%	1299	100±39
G. glyphis GN6505	5517644	4239803 (76%)	419896	100%	2475	302±59
G. glyphis GN6506	5798230	4565517 (78%)	3197140	100%	18644	266±91
G. glyphis GN15749	1153581	625887 (54%)	544141	100%	8872	269±38
G. gangeticus GN6923*	10222930	9377899 (91%)	32105	100%	167	113±47
G. gangeticus GN1188	2553284	1736816 (68%)	89677	100%	528	215±65
G. gangeticus GN2668	2759642	630227 (22%)	194623	100%	1143	198±42
G. gangeticus GN2669	5985818	5465953 (91%)	182142	100%	1078	251±63
G. siamensis GN12170*	4068382	616806 (15%)	2658	99%	16	123±47
G. garricki GN10999+	-	-	-	-	-	-
G. garricki GN6502	4876242	3911716 (79%)	2406	100%	15	107±47
G. garricki GN6503	4516256	3368406 (74%)	2540358	100%	14884	290±89
G. garricki GN6504	4843784	3596334 (74%)	2671798	100%	15601	276±92
G. fowlerae GN3376	4892498	4786864 (97%)	3500488	100%	20713	283±94
G. fowlerae GN3377	2066202	53610 (3%)	4723	96%	33	191±46
G. fowlerae GN1363	6374022	57776 (1%)	35679	97%	201	163±55
G. fowlerae GN4673	5986462	672443 (11%)	152297	100%	859	190±42
G. sp. GN3830	2904258	584183 (20%)	17101	100%	95	201±62
G. sp. GN2673	5452246	2830794 (51%)	111437	100%	643	237±59
G. sp. GN3681	5801596	3137846 (54%)	687942	100%	3993	219±44
G. sp. GN3682	4751292	2617663 (55%)	473724	100%	2751	220±44
G. sp. GN2667	2606798	48709 (1%)	6512	97%	37	176±52
G. sp. GN2671	2575756	314846 (12%)	18379	100%	108	232±53
L. temminckii GN1670	5361586	5059749 (94%)	3624093	100%	21351	278±93
L. tephrodes GN4239+	-	-	-	-	-	-

* Museum samples. + Samples sequenced using Sanger methods

Table S3. Number of SNPs retrieved for each taxon.

Taxon	number of SNPs retrieved	percentage of total SNPs*
G. gangeticus GN2668	1028	99
G. garricki GN6502	1020	98
G. garricki GN6504	1018	98
G. sp GN3682	1004	96
G. sp GN3681	1001	96
G. fowlerae GN3377	997	96
G. fowlerae GN1363	973	93
G. glyphis GN6506	960	92
G. glyphis GN6505	938	90
G. fowlerae GN4673	900	86

*The total number of called SNP loci is 1041.

Table S4. Primers used to assess the quality of extracted DNA and long-range PCR primers for bait preparation.

Primer	Sequence 5' – 3'	Amplicon size	Used for
Cmtshort_F583	CGCTAGCTTNAAACCCAAAGGAC	131bp	Checking DNA
Cmtshort_R713	GAGCTGRCGACGGCGGTAT		
ND2_ ILEM	AAGGAGCAGTTTGATAGAGT	1048bp	Checking DNA
ND2_ ASNM	AACGCTTAGCTGTTAATTAA		
Gly_16244F_blocked*	/5SpC3/AAGGTAAACTTGAGCTATCCTCGACAC	16672bp	Long range PCR
Gly_16215R_blocked	/5SpC3/TAGCGATTTGCTTCATACCGAACT		
Gly_7543F_blocked	/5SpC3/AGATTTAACTCCAGGCCAATTTCG	16659bp	Long range PCR
Gly_7485R_blocked	/5SpC3/CTTCAGTATCATTGATGACCCATAGC		

*The 5′ end of the blocked primers were modified with a C3 spacer (C3-blocked) to improve coverage uniformity of the reads (1).

Detailed Protocol for Preparing Homemade Baits

Introduction

Mitochondrial genomes can be amplified using long-range PCR. The PCR product can be sheared to make a library, and subsequently sequenced using Illumina sequencing technology. The long-range PCR product can also be used to make homemade biotinylated baits that can be used to target the mitochondrial genome for sequencing using DNA hybridisation techniques.

Library prep procedure

I. Long-range PCR

Mitochondrial genomes can be amplified using one or more pairs of primers. Using more primer pairs that amplify different overlapping parts of the mitochondrial genome may produce cleaner PCR products. Use 5' blocked primers, such as /5SpC3/ to increase the homogeneity of capture (1).

1. Prepare a master mix for the number of samples needed as follows, adding 0.4 μL of total DNA as template:

Reagent	Volume (μL) per sample	Final concentration
Takara LA Hot Start Taq*	0.2	0.05 U/μL
Buffer (10×)*	2	1×
$MgCl_2$ (25 mM) *	2	2.5 mM
dNTPs (25 mM) *	0.32	0.4 mM each
Forward primer (10 μM)	1	0.5 μM
Reverse primer (10 μM)	1	0.5 μM
H_2O	13.08	

* TaKaRa Bio Inc., Shiga, Japan.

2. Load the tubes onto a thermo cycler and run the following program:

    1. 94 °C for 1 min

    2. 98 °C for 10 sec

    3. 68 °C for 15 min

    4. go to step 2. Repeat for 34 cycles

    5. 72 °C for 30 min

    6. 4 °C for 10 min

3. Check 2 µL of the PCR product on a mini agarose gel. For the products with clean bands in the expected size range, perform subsequent comparable PCRs in order to obtain enough DNA for shearing. Note: products > 10 kb may sometimes remain "stuck" in the well of the agarose gel.

II. Shearing the genomic DNA

1. Please read "*DNA Clean-Up Using Solid Phase Reversible Immobilization (SPRI)*" for instructions on how to use the SPRI system. "Dry" the SPRI beads using a magnetic plate. Add the PCR product and 10 µL 20% Polyethylene Glycol (PEG; Life Technologies, New York, USA) to each tube with dried beads. Clean the PCR products using SPRI methods. Elute the samples with 20 µL of nuclease free water.

2. Measure the concentration of the elution. If there are multiple amplified reactions for one sample, mix them in equimolar ratios, i.e. shorter fragments should be at lower concentrations than larger fragments.

3. Shear 3 to 6 µg PCR product in 130 µL volume to ~ 200 bp range using a Covaris M220 Focused-ultrasonicator™ (Covaris, Inc. Massachusetts USA).

III. Size selection

A size selection step may be applied using the SPRI bead method at different concentrations of PEG. The following protocol is used to select DNA with fragment sizes in the range 100 – 200 bp.

1. Add 100 µL Agencourt® AMPure® XP beads (Beckman Coulter Inc, Georgia, USA) to a 200 µL tube. "Dry" the beads using a magnetic plate as per the SPRI method.
2. Add 50 µL sheared samples and 37.5 µL 20% PEG buffer to the dried beads, vortex the tube. Incubate at room temperature for 10 min on the bench. Collect the liquid at the bottom of the tube by brief centrifugation.
3. Place the tube on a magnetic plate, and let it stand for 10 min to separate the beads from the solution. Transfer the supernatant to a new tube with dried beads prepared as in step III.1.
4. Add 12.5 µL 20% PEG buffer to the sample tube. Prepare one positive and one negative control by adding 50 µL AMPure® XP beads to a tube with 30 µL nuclease free water and to another with 30 µL positive DNA (1:100 diluted PCR product of any gene with a size 100 bp - 200 bp). Vortex all tubes. Let all tubes sit for 10 min, collect the liquid at the bottom of the tube by briefly centrifugation.
5. Place the tube on a magnetic plate, and let it stand for 10 min to separate the beads from the solution. Pipette off and discard the supernatant without removing the beads.
6. Leave the tube on the magnetic rack, and wash the beads by adding 186 µL of freshly prepared 70% ethanol. Let stand for 1 min and remove the supernatant. Keep the tube on the magnetic rack. It is important that the beads are not disturbed during this phase.
7. Repeat Step 6 one more time for a total of two washes.

8. Remove the supernatant and let the beads air-dry for 5 min at room temperature without caps to allow residual traces of ethanol to evaporate.
9. Proceed to the next step immediately.

IV. 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.

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

Reagent	Volume (µL) per sample	Final concentration
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) α	1	0.5 U/µL
T4 DNA polymerase (5 U/µL) α	0.4	0.1 U/µL
H_2O	16.32	

*Fermentas, Thermo Fisher Scientific, Massachusetts, USA
αNew England Biolabs, Massachusetts, USA

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

V. Adapter ligation

M13 adapters are ligated to the ends of the repaired molecules using T4 DNA ligase. ***Be sure to use the correct M13 adapters, not the Illumina adapters for the target library.***

1. Prepare a master mix for the required number of samples. Add 39 µL of the master mix to each sample tube. Mix the samples well by vortexing.

Reagent	Volume (µL) per sample	Final concentration
T4 DNA ligase buffer (10×)*	4	1×
PEG-4000 (50%)*	4	5%
Adapter mix (50 µM each)	2	1.25 µM each
H_2O	29	

*Fermentas, Thermo Fisher Scientific, Massachusetts, USA

2. Spin down the liquid by brief centrifugation. Add 1 µL T4 DNA ligase (5 U/µL) to each sample tube. 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. Cleanup the reaction using the SPRI method (adding 40 µL 20% PEG). Keep the dried beads. Proceed immediately to the next step.

VI. 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
ThermoPol reaction buffer (10×)*	4	1×
dNTPs (25 mM each)	0.4	250 µM each
Bst polymerase, large fragment (8 U/µL)*	1.5	0.3 U/µL
H_2O	34.1	

* New England Biolabs, Massachusetts, USA

2. Add 40 µL of master mix to the samples. Briefly vortex to mix the sample. Collect the liquid at the bottom of the tube by brief centrifugation. Incubate the samples for 20 min at 37 °C.
3. Cleanup the samples using the SPRI method (adding 40 µL 20% PEG). Elute the samples with 20 µL of nuclease free water. Transfer the supernatant to a new tube. The libraries

can be kept frozen at -20 °C for a short period.

VII. PCR to make the biotinylated baits.

1. Prepare a master mix as follows. The PCR can be performed in a larger volume if more baits are required.

Reagent	Volume (µL) per sample	Final concentration
Gold Buffer (10×)*	5	1×
$MgCl_2$ (25 mM)*	5	2.5 mM
dNTP/UTP (25 mM ATP, UTP, CTP and GTP) α	0.5	0.25 mM
Bio*M13F (-40) (10 µM)	2	0.4 µM
Bio*M13R (10 µM)	2	0.4 µM
AmpliTaq Gold 360*	0.25	1.25 U/reaction
H_2O	29.25	

*Life Technologies, New York, USA
α Fermentas, Thermo Fisher Scientific, Massachusetts, USA

2. Add 44 µL of master mix to a tube with 6 µL of the bait library. Mix well and amplify the samples using the following thermal profile: 95 °C for 9 min, ~ 32 cycles of 95 °C for 15 sec, 55 °C for 30 sec and 72 °C for 45 sec, followed by 72 °C for 7 min, and hold at 4 °C for 10 min. Check the PCR product on an agarose gel. There should be no amplification in the negative sample. There should be a size shift for the positive sample compared to the positive DNA insert.
3. Cleanup the PCR product using the SPRI method (adding 100 µL 20% PEG). Elute the DNA using 200 µL of nuclease free water and transfer it to a new tube. Measure the concentration of the baits using a Qubit® 2.0 Fluorometer or similar.

Recipes and adapter/primer sequences

M13 adapter mix

1. Assemble the following hybridization reactions in separate tubes:

Hybridization mix for adapter 1 (100 µM):

Reagent	Volume (µL)	Final concentration
Adapter 1 (500 µM)	20	100 µM
Id697 (500 µM)	20	100 µM
Oligo hybridization buffer (10×)	10	1×
H_2O	50	

Hybridization mix for adapter 2 (100 µM):

Reagent	Volume (µL)	Final concentration
Adapter 2 (500 µM)	20	100 µM
Id697 (500 µM)	20	100 µM
Oligo hybridization buffer (10×)	10	1×
H_2O	50	

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

Oligo hybridization buffer (10×)

Reagent	Volume (µL)	Final concentration
NaCl (5 M)	1000	500 mM
Tris-Cl, pH 8.0 (1 M)	100	10 mM
EDTA, pH 8.0 (0.5 M)	20	1 mM
H_2O	8880	

Adapters and primers

Name	Sequence
M13 Adapter 1	G*T*TTTCCCAGTCACGACTTCATA*C*G
M13 Adapter 2	C*A*GGAAACAGCTATGACTTCATA*C*G
Id697	C*G*TATG*A*A
M13_Forward (-40)	GTTTTCCCAGTCACGAC
M13_Reverse	CAGGAAACAGCTATGAC
bio*M13_Forward (-40)	/52-Bio/GT TTT CCC AGT CAC GAC
bio*M13_Reverse	/52-Bio/CA GGA AAC AGC TAT GAC

* indicates a PTO bond

Detailed Protocol for Target Capture

Introduction

This protocol is modified from the protocol described in Maricic *et al*. (2).

Gene capture procedure

I. Prepare baits and beads

1. Mix $n \times 10$ μL (10 ng) baits with $n \times 10$ μL of 2× BWT buffer and $n \times 0.125$ μL of each blocking oligo (BO1F, BO2F). n is the number of samples.
2. Denature the baits mixture for 5 minutes at 95 °C, chill on ice.
3. Add $n \times 10$ μL of Dynabeads® M-270 streptavidin beads (Life Technologies, New York, USA) to a 200 μL or 1.5 mL tube according to the volume of the mixture. Dry the beads with a magnet.
4. Wash the beads three times with 200 μL BWT buffer. Dry the beads.
5. Wash the beads three times with 200 μL TET buffer. Dry the beads.
6. Mix the baits with beads.
7. Rotate the baits and beads mixture for 20 minutes at room temperature.
8. Aliquot 20 μL to a tube for each sample
9. Dry the beads and wash them three times with 100 μL BWT buffer, preheated to 60 °C.
10. Wash the beads three times with 100 μL TET buffer. Dry the beads.

11. Add 25 μL 2× Hb buffer (Agilent Technologies, California, USA) to each tube. Keep in refrigerator until step III.

II. Library Master Mix

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

Reagent	Volume (μL) per sample	Final concentration
BO1.P5short.F (200 μM)	0.125	1 μM
BO3.P7.part1.F (200 μM)	0.125	1 μM
Blocking reagent (10×)	5	2×
H_2O	14.5	

2. Add 19 μL of library master mix to empty tubes, and then add 6 μL of the target library to each sample tube. Mix the sample by vortexing. Collect the liquid at the bottom of the tube by brief centrifugation.

III. Hybridization

1. Set the following program on a thermal cycler: 95 °C for 5 min, 65 °C for 3 min, 65 °C for 10 min, and hold 65 °C. Transfer the tube containing the Library Master Mix to the thermocycler and start the program. This will denature the DNA library for 5 minutes at 95 °C.
2. Once the thermocycler program reaches step 2 (temperature = 65 °C), transfer the tube containing the baits and beads mix to the thermocycler. Leave the Library Master Mix in the thermocycler. This will pre-warm the baits and beads mix for 3 minutes at 65 °C.
3. While keeping tubes at 65 °C, transfer 25 μL of library master mix to the tubes with the baits and bead mixture.

4. Heat a hybridization oven to 65 °C. Move the tube to the oven, rotate and hybridize at 65 °C for 36 hours.

IV. Washing

1. Collect the liquid at the bottom of the tube by brief centrifugation. Dry the beads.
2. Wash three times with 100 µL BWT buffer.
3. Add 100 µL preheated (60 °C) HW buffer, incubate for 2 min at 60 °C then dry the beads.
4. Wash one more time with 100 µL BWT buffer.
5. Wash one time with 100 µL TET buffer. Dry the beads.
6. Add 50 µL nuclease free water to beads.

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

The captured library is recovered and amplified using 'off-beads amplification' (3), i.e. the captured target is amplified off the target-bait-bead complex during the indexing PCR. This avoids the need for chemical denaturation and maximizes retention of captured products.

1. Prepare a master mix for the number of samples as follows:

Reagent	Volume (µL) per sample	Final concentration
Gold Buffer (10×)*	5	1×
$MgCl_2$ (25 mM)*	4	2 mM
dNTPs (25 mM each)	0.5	0.25 mM
Primer IS4 (10 mM)	0.5	0.4 mM
AmpliTaq Gold 360*	0.25	1.25 U/reaction
H_2O	16.25	

*Life Technologies, New York, USA

2. Add 26.5 µL of master mix, 23µL well-mixed sample from step IV-6, and 0.5 µL

indexing primers (10 μM) to each tube. Mix well and amplify the samples using the following thermal profile: 95 °C for 9 min, ~ 12 cycles of 95 °C for 15 sec, 60 °C for 30 sec and 72 °C for 45 sec, followed by 72 °C for 7 min, and hold at 4 °C for 10 min. The number of PCR cycles can be adjusted according to the amount of starting material used to construct the library.

3. Load 3 μL of PCR product on an agarose gel to check the size of the captured library. The band should be barely visible.
4. Add 1 μL (1U/μL) of Uracil-DNA Glycosylase (UDG) to each sample tube to digest the baits. Incubate for 10 min at 37 °C, followed by 10 min at 94 °C, then hold at 4 °C for 2 min.
5. Cleanup the PCR product using the SPRI method. Elute the DNA using 20 μL of nuclease free water and transfer it to a new tube.

VI. Pooling multiple samples for sequencing

1. Determine the DNA concentration of indexed captured libraries using a Qubit® 2.0 Fluorometer or similar.
2. Pool all samples in equimolar ratios. The pooled library should be 20 μL at a concentration of about 2nM to 50nM, which is between 0.5 ng/μL to 13 ng/μL for DNA ~ 500 bp.

Recipes and adapter/primer sequences

Blocking Oligos

Name	Sequences
BO1.P5short.F	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-Pho
BO3.P7.part1.F	AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC-Pho
BO1F	GTTTTCCCAGTCACGACTTCATACG-Pho
BO2F	CAGGAAACAGCTATGACTTCATACG-Pho

Pho indicates a 3'-phosphate

BWT buffer 2×

Reagent	Volume (µL)	Final concentration
NaCl (5 M)	4000	2 M
Tris-Cl (1 M, pH 8)	200	20 mM
EDTA (0.5 M, pH 8)	40	2 mM
Tween20	10	0.1%
H_2O	5750	

TE buffer

Reagent	Volume (µL)	Final concentration
Tris-Cl (1 M, pH 8)	100	10 mM
EDTA (0.5 M, pH 8)	20	1 mM
H_2O	9880	

HW buffer

keep frozen at -20°C

Reagent	Volume (µL)	Final concentration
10× Amplitaq Gold Buffer*	200	1×
$MgCl_2$ (25mM)	200	2.5 mM
H_2O	1600	

*Life Technologies, New York, USA

TET buffer

10 ml TE buffer, 0.5 µl Tween20

Blocking reagent (10×), Hb buffer (2×)

Part of Oligo aCGH Hybridization Kit (Agilent Technologies, California, USA)

DNA Clean-Up Using Solid Phase Reversible Immobilization (SPRI)

Introduction

DNA precipitates and binds to the surface of carboxyl coated magnetic particles when there are high concentrations of polyethylene glycol (PEG) and salt. This process can be used to purify DNA (4). The Agencourt® AMPure® XP purification system (Beckman Coulter Inc, Georgia, USA) provides paramagnetic beads in optimized buffer to selectively bind DNA. The size of retained DNA is determined by the ratio of the AMPure® XP beads solution to the sample. For example, when a 1.8-volume of beads is added to 1-volume of sample, DNA larger than 100 bp is retained.

Instead of adjusting the volume of beads, 20% PEG solution (see attached recipe) can be added to "dried" (using a magnetic plate) beads and the sample to optimize the size selection of retained DNA. The following experiments show the effect of adding different concentrations of PEG to a DNA size standard sample. It is recommended that optimal conditions should be tested empirically for different batches of PEG and beads.

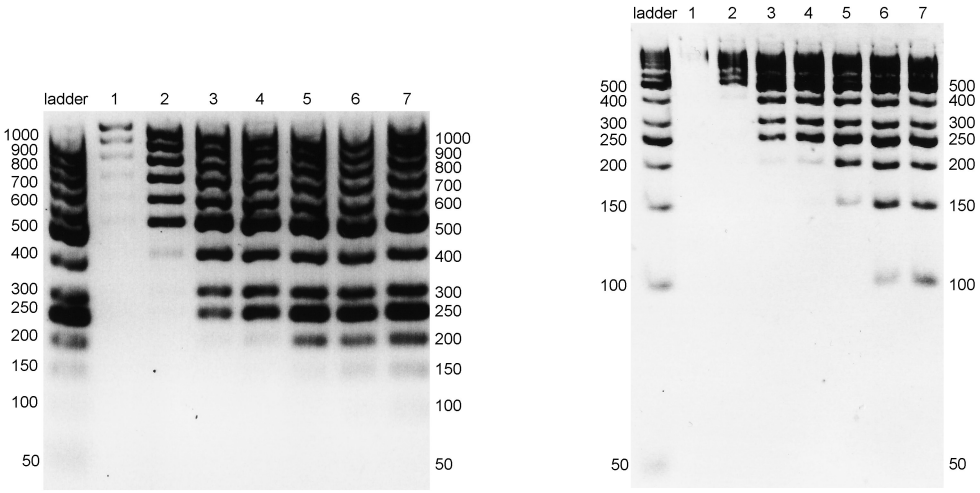

Fig. S5. Agarose gels showing DNA ladder (GeneRuler 50 bp Thermo Fisher Scientific, Massachusetts, USA) cleaned up with different amounts of PEG (lane descriptions are listed in the below table). Left – 2% agarose gel; right – 4% agarose gel.

Description for each lane in the above Figure showing the amount with of PEG added.

Lane	1	2	3	4	5	6	7
Sample (µL)	40	40	40	40	40	40	40
20% PEG (µL)	15	20	25	30	40	80	120
Conc. of PEG (%)	5.45	6.6	7.69	8.6	10	13.3	15

DNA clean-up procedure

I. Resuspend the stock solution of SPRI beads (Agencourt® AMPure® XP purification system; Beckman Coulter Inc, Georgia, USA). To make subsequent pipetting easier, add Tween 20 to the aliquot to a final concentration of 0.05% (i.e. add 1 µL of Tween 20 to 2 mL of bead suspension).

1. Add 50 µL resuspended SPRI beads to an empty tube. Dry the beads by placing the tube on a magnetic plate, allowing the beads to move to the walls of the tube before removing the supernatant. Then add the sample and 20% PEG buffer to the dried beads. The amount of

20% PEG can be adjusted to select preferred DNA size, according to Fig. S4.

2. Seal the tubes and vortex for several seconds until the beads are properly suspended.
3. Let the tube stand for 5-10 min at room temperature (more time is needed when high concentrations of PEG are used). Collect the liquid at the bottom of the tube by brief centrifugation.
4. 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.
5. Leave the tube on the magnetic rack, and wash the beads by adding 186 μL 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!
6. Repeat Step 6 one more time, for a total of two washes.
7. Pipette off the ethanol and remove residual traces by allowing the beads to air-dry for 5 min at room temperature without caps.
8. Add 20 μL of nuclease free water and seal the tube. Remove the tube from the magnetic rack, and resuspend the beads by repeated vortexing. Let it stand for 1 min, and then collect the liquid at the bottom by brief centrifugation. Occasionally the beads may appear clumpy after vortexing, this does not have a negative effect on DNA recovery.
9. Place the tube back on the magnetic rack, let stand for 1 min, and transfer the supernatant to a new tube. Carryover of small amounts of beads will not inhibit subsequent reactions.

Recipes

20% PEG solution

Reagent	Amount	Final concentration
PEG8000*	20 g	20%
NaCl	14.6 g	2.5 M
H_2O to 100 ml		

* Thermo Fisher Scientific, Massachusetts, USA

Note: After autoclaving the liquid may separate into two layers. It will become homogeneous once cooled.

References

1. Harismendy O & Frazer K (2009) Method for improving sequence coverage uniformity of targeted genomic intervals amplified by LR-PCR using Illumina GA sequencing-by-synthesis technology. *Biotechniques* 46:229-231.
2. Maricic T, Whitten M, & Pääbo S (2010) Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* 5(e14004).
3. Fisher S, et al. (2011) A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* 12(1):R1.
4. Hawkins TL, O'Connor-Morin T, Roy A, & Santillan C (1994) DNA purification and isolation using a solid-phase. *Nucleic Acids Res* 22(21):4543-4544.