Cronn, Richard, Brian J. Knaus, Aaron Liston, Peter J. Maughan, Matthew Parks, John V. Syring, and Joshua Udall. 2012. Targeted enrichment strategies for next-generation plant biology. *American Journal of Botany* 99(2): 383-396.

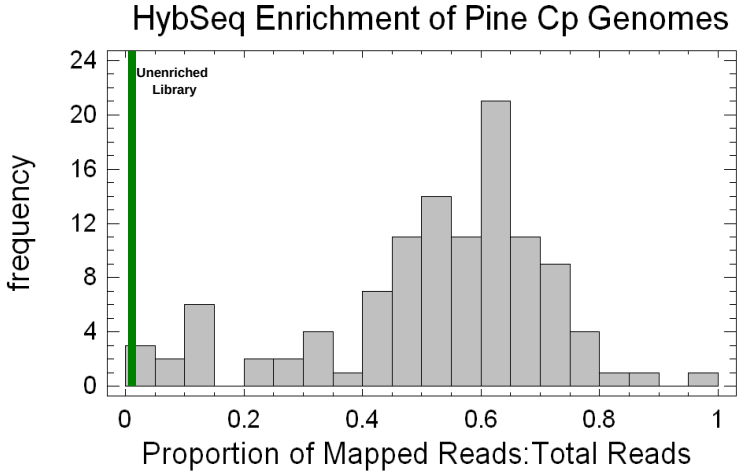
Appendix S1. Next-generation sequencing of hybridization-enriched libraries from plant organelle genomes.

This process starts with high-quality total genomic Illumina libraries, and uses biotinylated probes to enrich specific organelle targets of the library. Our example was designed for chloroplast genomes, but the process could be modified for mitochondrial genomes and nuclear targets. This process assumes that the user: (1) has access to organelle genome probes of interest (either single-stranded oligonucleotides or PCR products spanning the target region); (2) has made the choice of whether to use multiplex adapters (from Illumina or custom design); and (3) knows that the adapters are compatible with the Illumina single-end or paired-end sequencing process.

Once genome probes are prepared, hybridization adds 2-3 days to the usual Illumina library preparation time. Samples can be processed in individual tubes or multi-well plates. Most importantly, hybridizations can also be done with multiple (multiplex) samples if the libraries are prepared with custom barcode or Illumina index sequencing adapters. We have done as many as 96 hybridizations in an experiment, with barcoded libraries pooled to four per multiplex (96 samples; 24 hybridization reactions). The results were indistinguishable from non-multiplexed hybridizations.

Recovery Trends from 111 Hyb/Seq Enrichments of Conifer Plastomes

CP representation from unenriched conifer genomic libraries	0.5% - 3%
Mean CP representation following Hyb/Seq (*95% CI*)	52.6% (*48.9-56.2*)



Results depend: (1) quality of your input DNA; (2) quality of your Illumina library, (3) molar equivalence across pooled probes, and (4) probe labeling efficiency. Patience and care with these steps will reward the user with higher quality data.

REAGENTS

Regents shown below are specific to hybridization-based capture and enrichment of organelle genome sequences. Approximate numbers of samples per minimum order are indicated for enzymes. Vendor names are for descriptive purposes only, and they do not imply endorsement by the U.S. Government.

1. Chloroplast genome probe assembly

1A. Double-stranded Chloroplast genome probes

- PCR product to be used as probes. These need to be accurately quantified and pooled in an equimolar (not equal mass) pool.
- DNA blunting kit (New England Biolabs [NEB] #E1201; alternatively, you can order T4 DNA polymerase and T4 polynucleotide kinase independently)
- T4 DNA ligase and buffer (NEB #M0202S)

1B. Single-stranded Chloroplast genome probes

- Oligonucleotides to be used as probes. These can be PCR primers (20 – 25 bp) or microarray oligos (60 – 100 bp), as long as they provide high coverage of the target region.
- T4 polynucleotide kinase and buffer (NEB #M0201)
- T4 RNA ligase (ssRNA ligase; NEB #M0204S)
- Ethanol, 100%
- Sodium Acetate, 3M stock. (Ambion #AM9740 or equivalent)
- MgCl2, 1M stock. (Ambion #AM9530G or equivalent)
- co-precipitant (linear acrylamide or glycogen)

2. Probe amplification and labelling

- KOH, 4 M stock solution
- Tris-HCl, 0.4 M
- 5'-biotinylated $(N)_6$, 1 μmol . This random hexamer should be diluted to 1 nmol/μl using 10 mM Tris-HCl buffer
- ϕ29 DNA polymerase and buffer (NEB #0269S or equivalent)
- KCl, 0.5 M
- dNTPs, 10 mM
- biotin-14-dCTP, 0.4 mM (Invitrogen #19518-018)

3. Hybridization buffers and related

- SSC, 20X stock buffer (Ambion # AM9770 or equivalent)
- Disodium EDTA, 500 mM, pH 8.0 (Ambion #AM9260G or equivalent)
- SDS, 10% stock (Ambion #AM9822 or equivalent)

- Denhardt's solution, 50X stock (Invitrogen #750018 or equivalent)
- Lambda DNA, 500 µg/ml stock (NEB #N3011S or equivalent)

4. Magnetic bead preparation

- Streptavidin MagnaSphere paramagnetic beads (Promega #Z5481)
- 2X casein blocking buffer (Sigma #C7594)
- STE (1M Sodium Chloride/10 mM Tris-HCl pH 7.5/1 mM EDTA pH 8), 1X

5. Hybrid capture solutions

- 20X SSC (3 M sodium chloride and 300 mM trisodium citrate, adjusted to pH 7.0 with HCl) (Ambion #AM9770 or equivalent)
- 10% SDS (sodium dodecyl sulfate) (Ambion #AM9822 or equivalent)

6. Library enrichment

- Phusion Flash DNA polymerase 2X mix (Finnzymes Oy).
- Illumina PCR Primer 1.0. See Adapter/Primer page.
- Illumina PCR Primer 2.0. See Adapter /Primer page.

7. PCR clean-up kit

- QIAgen PCR clean-up kit, Agencourt AMPure magnetic beads, or equivalent.
- DNA storage buffer (i.e., EB or 10 mM Tris-HCl, pH 8.0).

8. Required equipment

- PCR machine or heating block (37°C – 80°C range); shaker/rotator
- Magnetic separator. Required for hybrid capture ; can also be used for paramagnetic bead DNA clean-up. Many sources, including Invitrogen, New England Biolabs, Ambion and Agencourt.

Overview

1. Probe assembly, labeling and amplification – *Probe synthesis involves the ligation of a probe set into a large, randomly concatenated molecule, which is then used as a template for probe preparation. Protocols differ depending on whether the starting probe set consists of double-stranded DNA (i.e., PCR products; step 1A) or single-stranded DNA (i.e., oligonucleotides; step 1B).*

With PCR products, the products are pooled in an equimolar mix, the 5' ends are phosphorylated, the 3' ends blunt-ended, and then the fragments are ligated to form a double-stranded concatemer. With single stranded probes, the oligos are pooled in an equimolar mix, the 5' ends are phosphorylated, and then the oligos are ligated with T4 RNA ligase to form a single-stranded concatemer.

Concatemers are amplified and biotinylated using a DNA polymerase capable of strand displacement (ϕ29 polymerase). The resulting probes are large (tens of kb), double-stranded, and are all biotinylated internally and on the 5' end.

Notes on probes:

- *We do not know the optimal size for oligonucleotide probes, but we've had success with 20- to 60-mers. Our results show that a single 25-mer probe can pull down ~1000 bp of target. Higher probe density will give fewer coverage gaps.*
- *Remember that single-stranded oligos CAN'T be cleaned with QIAgen columns. EtOH precipitation works well, but remember to add $MgCl_2$ to 10 mM final concentration.*

2. Hybridization – *Concatemer probes are mixed with Illumina libraries for selective hybridization to the chloroplast genome.*

3. Magnetic bead preparation – *Streptavidin coated paramagnetic beads are prepared for oligo capture using a blocking agent.*

4. Hybrid capture – *Concatemer probes are captured with streptavidin beads, and then washed under stringent conditions so that non-specifically bound materials are minimized.*

5. Library enrichment – *A final PCR enrichment step is used to bring the DNA concentrations up to what is required for Illumina sequencing. Cycles should be kept to the absolute minimum (10 -12, or even fewer) to prevent biasing of library and accumulation of identical reads.*

6. Analysis – *Too much here to cover, and you probably know how to handle much of this already. Some issues we've considered are identified, and they may be helpful for your experiments.*

7. Cost – *The minimum order of reagents listed here will produce sufficient probe for over $\geq$1000 samples. If the same probe pool (PCR or oligonucleotides), is used to completion, then the cost of performing hybridization/capture on an Illumina library can be as low as ~ $1 per sample.*

STEP 1A

Probe synthesis using PCR products*. Use this protocol for double-stranded PCR products.*

1A.1. End-repair treatment to phosphorylate ends

- Mix ~100 ng of pooled PCR products (equimolar representation of all), 39 µl dH_2O, 5 µl of 10X blunting buffer, 21 µl blunting enzyme mix (NEB #E1201S), 5 µl of 1 mM dNTPs, and water up to 50 µl.
- Incubate for 1 hour at 37°C.
- Heat inactivate the enzyme at 65°C for 15 minutes.
- Clean using QIAgen, AMPure or similar. Elute in 20 µl EB or 10 mM Tris-HCl.

1A.2. Probe concatenation (Ligation of double-stranded PCR products oligos)

- Mix: 14 µl of phosphorylated ds PCR products, 3 µl of 10X DNA ligase buffer, and 1 µl T4 DNA ligase and 12 µl dH2O.
- Incubate for 15 minutes at room temperature.
- Heat inactivate enzyme at 65°C for 15 minutes.
- Clean using QIAgen, AMPure or similar. Elute in 20 µl EB or 10 mM Tris-HCl.
    - *Be sure to archive this product because it can be used for many hybridization probe labeling.*

1A.3. Denaturation of ds-DNAs

- Mix the following in a 0.2 ml tube:
    - 1 µl of concatenatedd DNA probes
    - 1 µl 0.4M KOH. This needs to be made fresh from 4M KOH stock
- Let stand for 3 minutes at room temperature
    - Add 2 µl of 0.4M Tris-HCl to neutralize, then place on ice.

1A.4. Multiple Displacement Amplification (MDA) and biotinylation of probes

- Mix the following into duplicate - 0.2 ml tubes:
    - 2 µl of denatured/neutralized DNA mixture per tube
    - 4 µl of ϕ29 polymerase buffer, 10X
    - 4 µl of KCl, 0.5 M
    - 4 µl of dNTPs, 10 mM
    - 2 µl of biotinylated random hexamers, 1 nmol/µl
    - 2.5 µl of 0.4 mM biotin-14-dCTP
    - 20.5 µl of dH2O
    - 1 µl of ϕ 29 DNA polymerase
- Incubate reactions for 18 hours at 30°C.
- Pool the products and clean using ethanol precipitation, as the products are too large for QIAgen (10-100kb) and AMPure does not work on such concentrated DNA. Resuspend in 100 µl to 200 µl of EB or 10 mM Tris-HCl.

Note 1: in this reaction, [BIO]dCTP is at a concentration of 0.025 mM and dCTP is at a concentration of 1 mM, so the molar ratio of [biotin]dCTP:dCTP is 1:40. For a target region containing 20% dC, this is adds one biotin per 200 bp.

Note 2: The yield per MDA synthesis can be large, often well over 10 μg per reaction (and up to 25 μg per reaction). For this reason, one labeling reaction can product sufficient probe for 10 – 20 hybridization samples.

1A.5. Probe quantitation

- Estimate the DNA concentration of cleaned, amplified biotinylated probes.
- You need 500 ng (minimum) to 1000 ng (ideal) of purified probe for each hybridization reaction.

STEP 1B

Probe synthesis using single-stranded oligonucleotides. *Use this protocol for single-stranded oligonucleotides. Note that single-stranded molecules do NOT need KOH denaturation prior to MDA.*

1B.1. Kinase Treatment to phosphorylate ends

- Mix: 30 µl of pooled 10 pmol/µl oligos (=300 pmol each), 14 µl dH2O, 5 µl of 10X T4 DNA ligase buffer, and 1 µl T4 PNK.
- Incubate for 1 hour at 37°C.
- Heat inactivate the enzyme at 65°C for 15 minutes.

1B.2. Ethanol precipitation of small, single-stranded phosphorylated oligos.

- Mix Kinase reaction with 2 vols 100% EtOH, 0.1 vol 3M NaOAc, 1.6 µl of 1M $MgCl_2$, and co-precipitant (if desired). Place in the -20°C freezer overnight.
- Centrifuge for 10 min, 12000 x g; remove supernatant and discard.
- Dry pellet at 37°C for 10 minutes.
- Resuspend in 10 µl dH2O.

1B.3. Probe Concatenation (Ligation of single-stranded oligos)

- Mix: 10 µl of phosphorylated oligo sample, 7 µl dH2O, 2 µl of 10X T4 RNA ligase buffer, and 1 µl T4 RNA ligase.
- Incubate for 1 hour at 37°C.
- Heat inactivate the enzyme at 65°C for 15 minutes.

1B.4. Ethanol precipitation of ligated single-stranded oligos.

- Ethanol precipitate the ligated oligos using 2 vols 100% EtOH, 0.1 vol 3M NaOAc, 0.8 µl of 1M $MgCl_2$, and co-precipitant (if desired). Place in -20°C freezer overnight or -80°C freezer for one hour.
- Centrifuge for 10 min, 12000 x g, 4°C; remove supernatant and discard.
- Dry pellet at 37°C for 10 minutes.
- Resuspend in 10 µl dH2O
    - *Be sure to archive this product because it can be used for many hybridization probe labelings*

1B.5. Multiple Displacement Amplification (MDA) and biotinylation of probes

- Mix the following together in a 0.2 ml tube:
  - 1 µl of cleaned, concatenated oligo pool
  - 4 µl of ϕ29 polymerase buffer, 10X
  - 4 µl of KCl, 0.5 M
  - 4 µl of dNTPs, 10 mM
  - 2 µl of biotinylated random hexamers, 1 nmol/µl
  - 2.5 µl of 0.4 mM biotin-14-dCTP
  - 21.5 µl of dH2O
  - 1 µl of ϕ 29 DNA polymerase
- Incubate reactions for 18 hours at 30°C.
- Pool the products and clean using ethanol precipitation, as the products are too large for QIAgen (10-100kb) and AMPure does not work on such concentrated DNA. Resuspend in 100 µl to 200 µl of EB or 10 mM Tris-HCl.

1B.6. Probe quantitation

- Estimate the DNA concentration of cleaned, amplified biotinylated probes.
- You need 500 ng (minimum) to 1000 ng (ideal) of purified probe for each hybridization reaction.

STEP 2 Hybridization

2.1. Hybridization reactions are done using a 40 µl volume. Mix the following reagents in the listed order:

- 12 µl of 20X SSC (final concentration = 6X)
- 0.4 µl of 500 mM EDTA (final = 5 mM)
- 0.4 µl of 10% SDS (final = 0.1%)
- 1.6 µl of 50X Denhardt's solution (final = 2X)
- 4 µl of 0.5 µg/µl Lambda DNA (final = 2 µg)
- 500 – 1000 ng of [BIO]-concatemer probe
- 500 – 1000 ng of prepared Illumina library (or multiple, barcoded libraries that are mixed in an equimolar ratio and total 500 – 1000 ng)
- Add water to a final volume of 40 µl.
- Overlay the reaction with 30 µl of molecular biology grade mineral oil.

2.2. Heat hybridization reaction to 95°C for 10 minutes.

2.3. Cool to 65°C , then incubate at 65°C for 36 - 48 hours.

STEP 3 Mag-bead preparation

1. **Resuspend streptavidin-coated paramagnetic beads, and dispense 50 µl for each library/probe combination.**
2. **Magnetize; remove and discard supernatant.**
   - replace with 200 µl of 2X Casein Blocking Buffer
   - shake for 2 hours.
3. **Magnetize; remove and discard supernatant (bead wash 1).**
   - replace with 200 µl of STE.
   - gently mix until beads are fully resuspended (bead wash 2).
4. **Magnetize; remove and discard supernatant**
   - replace with 200 µl of STE.
   - Gently mix until beads are fully resuspended (bead wash 3).
5. **Magnetize; remove and discard supernatant**
   - replace with 200 µl of STE.
   - Gently mix until beads are fully resuspended (bead suspension).
6. **Store at RT until use.**

STEP 4

Hybrid capture. *We recommend that you save the washes, just in case the products elute at an unpredictable point.*

1. Following hybridization (from step 2), mix 40 µl of hybridization mixture with 200 µl of the bead suspension (from previous step).
   - Incubate for 30 minutes, RT.

2. Magnetize; remove and save supernatant (= wash 1).
   - Resuspend beads in 200 µl of 65°C 1X SSC/0.1% SDS. Mix well.
   - Incubate for 15 min, 65°C.

3. Magnetize; remove supernatant and discard (= wash 2).
   - Replace with 200 µl of 65°C 1X SSC/0.1% SDS. Mix well.
   - Incubate for 10 min, 65°C.

4. Magnetize; remove supernatant and discard (= wash 3).
   - Replace with 200 µl of 0.5X SSC/ 0.1% SDS. Mix well.
   - Incubate for 10 min, 65°C.

5. Magnetize; remove supernatant and discard (= wash 4).
   - Replace with 200 µl of 0.1X SSC/ 0.1% SDS
   - Incubate for 10 min, 65°C.

6. Magnetize; remove supernatant and discard (= wash 5).
   - Add 50 µl of 80°C water. Mix well
   - Incubate for 10 min, 80°C.

8. Magnetize, remove and RETAIN supernatant.
   - This supernatant contains the hybridization-enriched products.

9. Discard beads.

STEP 5

Library Enrichment and Sample Submission

1. Library PCR. Each enriched library should be amplified in a 50 µl PCR reaction using the following components:

- 25 µl 2X Phusion-Flash premix
- 0.5 µl Illumina PCR 1.0 primer (25 µM)
- 0.5 µl Illumina PCR 2.0 primer (25 µM)
- 10 µl of hybridization-enriched library. You want to use a lot because you want minimal amplification of the library.
- Water up to 50 µl.

2. Cycling conditions:

- 98°C, 30 sec
- 98°C, 30 sec; 65°C, 30 sec; 72°C, 30 sec ; Repeat for the minimum number of cycles required to obtain enough product for sequencing (10-12, ideally less).
- 72°C, 5 min
- 4°C hold

3. Product evaluation, purification, quantitation

- 10% of the library should be run on a test gel to verify that amplification worked and that the size distribution matches the original (pre-enriched) library.
- PCR-enriched, selected libraries should be cleaned and quantified, preferably using BioAnalyzer DNA chips or similar products.
- Once cleaned, this product is ready to run on the Illumina GAII or HiSeq sequencer.
- If Library Enrichment (step 5) does NOT yield a PCR product, you can concentrate the enriched sample and run more of the DNA in a PCR reaction.

4. Library Submission

- In our experience, this process enriches conifer chloroplast genomes so that they increase in representation from less than 5% of the total reads, to between 40 - 80% of the reads (average = 50%). For purposes of multiplexing, an estimate of 40% is conservative.
- The 'ideal' multiplexing limit has to be determined experimentally, and it is certain to vary by species, probes, and next-generation sequencing platforms.