Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies

Cédric Mariać1, Marie Coudersc1, Nora Scarcostc1, Claire Billoz2, Adeline Bamaud3, Adama Faye4,5, Bonaventure Sonké6, Jean-Christophe Pintaud1, Yves Vigouroux1 and Thomas L. P. Couveur1,5
(1) IRD, Montpellier, France; (2) CIRAD, Montpellier, France; (3) IRD/ISRA, Dakar, Senegal; (4) IRD Yaoundé, Cameroon; (5) Ecole Normale Supérieure, University of Yaoundé I, Cameroon

Biodiversity, phylogeography and population genetics studies are expected to be revolutionized by access to large datasets thanks to new sequencing methods. We develop an easy and cost-effective protocol for in-solution enrichment hybridization capture of complete chloroplast genomes. The protocol uses cheap in-house species-specific probes developed via long range PCR of the entire chloroplast. Barcoded libraries are constructed and in-solution enrichment of the chloroplasts is done using the probe. Our protocol allows whole chloroplast genomes to be sequenced at a modest cost for large samples. This will allow unprecedented resolution for closely related species in phylogeography and population genetics studies using chloroplast sequences.

**Protocol**

Protocol undertaken on four crop species: Oryza glaberrima (African rice), Pennisetum glaucum (Pearl millet), Dioscorea spp. (African yams) and Digitaria exilis (Fonio).

To capture chloroplast sequences directly from the genomic DNA libraries, biotinylated probes were produced in house.

Libraries were then constructed using 6-bp barcodes to allow for multiplexing at different degrees.

**Results**

**Validation of the probes in Oryza glaberrima**

![Graph showing number of chloroplast reads](image)

Number of reads mapped to the chloroplastic reference genome of O. sativa (NC_001320):

(A) Sequencing the probes only: confirms the probe is chloroplastic.

(B) Sequencing a library enriched using the probes: confirms good sequencing of chloroplast.

**Enrichment success for 4 crop species**

![Graph showing enrichment success](image)

Enriched versus non enriched libraries for the four species using the enrichment protocol. Specific probes were designed for each species.

**Enrichment success of 95 multiplexed pearl millet individuals in a single MiSeq run**

![Graph showing enrichment success](image)

(A) % of reads mapping to the reference; (B) Average coverage: 99% across the 94 individuals. 90 out of 95 with coverage > 95%; (C) The number of reads mapping to their reference plotted against % cp genome coverage.

This study was financed by

Grant number CHLORODIV 1202-040