Application of a Novel, Targeted Sequencing-Based Genotyping Approach for Cost Effective Marker Assessment in *O. sativa*.

Andrew Barry¹, Amy B. Emerman², Björn Textor³, Kruti M. Patel³, Sarah K. Bowman³, Scott M. Adams³, Brendan S. Desmond³, Jonathon S. Dunn², Susan E. Corbett², Charles D. Eife², Evan Mauclerc³, and Cynthia L. Hendrickson²

¹New England Biolabs, Inc., Ipswich, MA; ²Directed Genomics, Ipswich, MA; ³New England Biolabs, GmbH, Frankfurt, Germany

Decreases in sequencing costs have increased the availability of public SNP databases while necessitating development of targeted genotyping assays for use in marker assisted genomic selection for a variety of crop species. The NEBNext Direct Genotyping Solution is a novel, hybridization-based target enrichment approach that has been optimized for use in genotyping applications to increase the number of assays that can be performed in a single reaction, while providing sequencing coverage depth suitable for SNP identification. The approach enables high-levels of multiplexing of both isolates and markers, allowing enrichment of hundreds of thousands of SNP targets in a single hybridization reaction, and the protocol is easily completed in a single day. We developed a panel covering the 1,996 single nucleotide polymorphisms previously identified as markers for polymorphism detection in *O. sativa*. Here, we demonstrate the application of this panel to cost-effectively enrich defined SNP markers in a highly specific and uniform manner prior to next-generation sequencing.

**Introduction**

After sequencing, the reads were demultiplexed with a Picard-based workflow. Sequencing reads were aligned to the Osativa, 193 reference genome using BWA-MEM and PCR duplicates were identified using the UMI².

25 ng of 96 individual rice DNA samples were enzymatically fragmented and 5' tagged with an Illumina®-compatible P5 adaptor that incorporates both an inline sample index to tag each sample prior to pooling and an inline UMI to mark each unique DNA fragment within the samples, as shown in the workflow. The 96 samples were pooled and enriched using the NEBNext Direct Genotyping Solution, following removal of PCR duplicates using the UMIs, demonstrating uniform enrichment across SNP marker targets.

**Workflow**

Unique distribution of sequencing reads across all 96 pooled samples demonstrates the comparable performance of each sample during hybridization-based capture and library preparation. The number of reads aligning to each of the 96 samples assayed.

**Methods**

Following removal of PCR duplicates, the 96 pooled samples had an average of 107 unique reads covering the targeted markers in a portion of the NextSeq run, with very little variation in the mean coverage across the samples.

**Uniform Coverage Across 1,996 Markers**

Uniform distribution of sequencing reads across all 96 pooled samples demonstrates the comparable performance of each sample during hybridization-based capture and library preparation. The number of reads aligning to each of the 96 samples assayed.

**Advantages**

- Robust, user-friendly protocol to generate Illumina-compatible, target-enriched libraries within one day
- Multiplexes samples upfront to reduce cost and increase throughput
- Scalable from 100-500 markers or more
- Processes up to 9216 samples in a single 96-well plate
- Flexible multiplexing: Same protocol can pool 4 to 96 samples into a single hybridization
- Maximizes on-target bases by enzymatic removal of off-target sequences
- Column purification of DNA samples is not required for most plants
- Safe workflow stopping points throughout the protocol
- Automation-friendly