Miniaturized, fully automated NEBNext Direct® Genotyping Solution using the Biomek i7 Hybrid Automated Workstation in conjunction with the Echo 525 Acoustic Liquid Handler for marker-assisted breeding

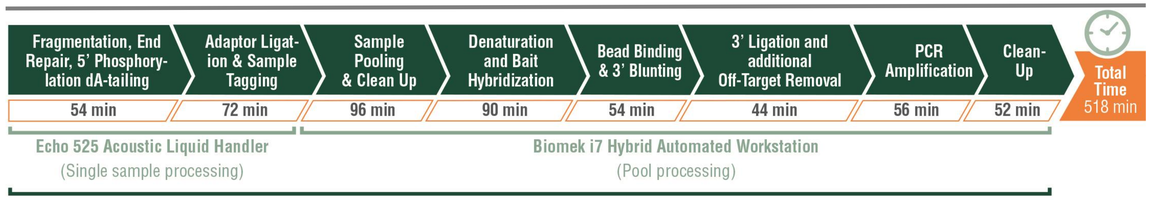
Michael Petridis[1], Juergen Zimmermann[2], Ferris Jung[2], Martin Giner[3], Dominik Zahr[3], Evan Mauceli[4], Brendan Desmond[4], Jonathan Landry[2], Bjoern Textor[1], Cynthia Hendrickson[4], Andrew Barry[5], Vladimir Benes[2], and Zach Smith[6]

(1) New England Biolabs GmbH, Frankfurt am Main, Germany, (2) GeneCore, EMBL, Heidelberg, Germany, (3) Beckman Coulter GmbH, Krefeld, Germany, (4) Directed Genomics, Ipswich, MA, (5) New England Biolabs, Ipswich, MA, (6) Beckman Coulter Life Sciences, Indianapolis, USA

Introduction

The NEBNext Direct® Genotyping Solution delivers cost-effective, high-throughput, NGS-based targeted genotyping for plant and animal applications that allows efficient genotyping of thousands of markers with high target coverage and uniformity. The intrinsic features of the NEBNext Direct® Genotyping Solution facilitate high-throughput sample processing including pre-capture barcoding, pooling of up to 96 samples per capture, and dual barcoding, allowing up to 9,216 samples to be run in parallel on Illumina sequencers. As genotyping for molecular breeding applications require solutions that enable robust and cost-efficient processes to implement at scale, we incorporated the Echo 525 Acoustic Liquid Handler in a fully automated workflow to enable miniaturization of certain reaction volumes and to further reduce the operational costs per sample, while preserving data quality. Here we demonstrate the development and implementation of an automated protocol for the NEBNext Direct® Genotyping Solution on the Biomek i7 Hybrid Automated Workstation including the Echo 525 Acoustic Liquid Handler to address the high throughput demands of marker-assisted plant breeding applications.

Workflow



The complete workflow of the NEBNext Direct® Genotyping Solution. Indicated below each process is the duration of the individual steps for the processing of 384 samples distributed over four 96-well plates. The Echo 525 Acoustic Liquid Handler enables the miniaturization of the *Fragmentation and End Prep* as well as the *5' Adaptor Ligation with Sample Indexing steps* by a factor of 4x. Subsequent steps of the workflow are performed on the Biomek i7 Hybrid Automated Workstation.

Implementation of Echo 525 Acoustic Liquid Handler

The method was carried out using the NEBNext Direct® Genotyping Solution with 25 ng tomato DNA input per sample in a 384-plex distributed over four 96-well-plates, using 96 pre-capture sample indexes and one post-capture pool index per 96-well plate. The Echo 525 Acoustic Liquid Handler was used for miniaturization of the Fragmentation and End Prep as well as the 5' Adapter Ligation with Sample Indexing steps by a factor of 4x. Subsequent steps of the workflow were performed on the Biomek i7 Hybrid Automated Workstation.

Biomek gripper loading an Echo Qualified 384-well Polypropylene Source Microplate onto the Echo source stage.

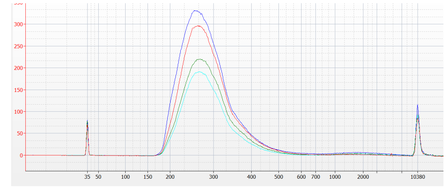
	Plastics	384 samples in four pools Biomek i7 only	384 samples in four pools Echo 525 & Biomek i7
Pre-Pool Consumables	BC1025F	144 tips	116 tips
	BC190F	32 tips	44 tips
	BC50F	1,560 tips	0 tips
	Midi plates (AB_1127)	1	1
	Echo Qualified 384-well PP Plate	0	1
	PCR plates	4	4
Post-Pool Consumables	BC1025F	37 tips	37 tips
	BC190F	276 tips	276 tips
	BC50F	28 tips	28 tips
	Midi plates (AB_1127)	1	1
	PCR plates	2	2

Consumables required to prepare 384 samples in four post-capture pools, with and without use of the Echo 525 Acoustic Liquid Handler.

Library QC Metrics and Sequencing of Highly Multiplexed Libraries

Tomato DNA was fragmented for 15 minutes, the samples were indexed with a full 96 Indexed 5' adaptor plate and pooled into four individual pools. Hybridization was performed using baits derived against a set of 500 SolCAP markers from tomato. The pools were indexed with a pool index, amplified with 19 PCR cycles and a double bead cleanup was performed. Final library yields and quality were assessed with an Agilent Bioanalyzer High Sensitivity DNA chip. The concentration of the libraries was 16.2 ng/µL ± 4 ng/µL, resulting in a total yield of 452.9 ng ± 112.1 ng, with an average fragment size of 294 bp ± 3 bp peaking at 266 bp ± 4 bp

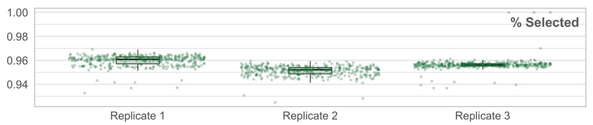
The three replicates of a 384-plex were loaded at 8 pM and sequenced on an Illumina MiSeq with the MiSeq v3 kit and a running-mode 20-8-100. Read 1 captures the inline UMI and sample barcode, the i7 index read captures the pool index that is added to all samples of one pool during PCR enrichment and Read 2 captures the target sequence. Replicate 1 was sequenced at a higher depth than the pool of replicates 2 and 3, that is reflected in the box plots shown on the right. After sequencing, the reads were demultiplexed with a Picard-based workflow. The sequence reads were aligned to the reference genome using bwa mem[1] and PCR duplicates were filtered using the UMIs and fgbio[2]. Sequencing metrics were calculated using Picard Tools[3]. Key parameters extracted from three replicates of a 384-plex highlight the high data quality. The results of the miniaturization and automation of the NEBNext Direct® Genotyping Solution further emphasize the high process stability with a dropout rate below 0.6%



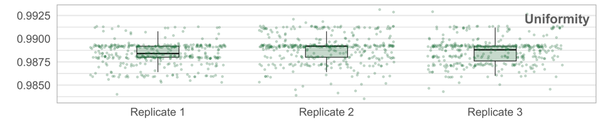
Electropherograms of final libraries from replicate 3 assessed with an Agilent Bioanalyzer High Sensitivity DNA chip.

Box plot showing the number of reads passing Illumina's filter for each of the 384 samples per replicate as measured by the Picard Alignment Summary Metrics tool. Sequencing on an Illumina MiSeq yielded counts between 21,114-109,577 with an average of 54,992.78 (Replicate 1) and 6,106-61,930 with an average of 27,085.95 (Replicates 2 and 3).

Box plot showing the mean target coverage for each of the 384 samples per replicate as calculated by the Picard HS Metrics tool. Analysis of sequencing data from the Illumina MiSeq showed a mean target coverage between 30-158.34 with an average of 80.14 (Replicate 1) and 9.40-98.74 with an average of 42.80 (Replicates 2 and 3).



Box plot showing the percentage of bases located on or near a target for each of the 384 samples per replicate calculated by the Picard HS Metrics tool. Analysis of sequencing data from the Illumina MiSeq showed percent selected is >92.49% across all samples. Two dropouts in Replicate 1 and four in Replicate 3 are not shown for scaling purposes.



Box plot showing the sequencing uniformity for each of the 384 samples per replicate as calculated by the Picard HS Metrics tool. Uniformity is defined as the percentage of targets with coverage greater than 20% of the mean target coverage. Analysis of the sequencing data highlights the uniformity of >98.36% across all samples. Two dropouts in Replicate 1 and four in Replicate 3 are not shown for scaling purposes.

Conclusions

- Integration of the Echo 525 Acoustic Liquid Handler enables the miniaturization of the single sample processing steps by a factor of 4x, reducing the operational costs through reduction of plastic consumables.
- Miniaturization and automation of the NEBNext Direct® Genotyping Solution generate highly consistent data to address the high throughput demands of marker-assisted plant breeding applications.

References

[1] Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 [q-bio.GN].
[2]Fulcrum Genomics, https://github.com/fulcrumgenomics/fgbio.
[3] "Picard Toolkit." 2019. Broad Institute, GitHub Repository. https://broadinstitute.github.io/picard/; Broad Institute