High Quality Automation Libraries Prepared from a Broad Range of DNA Input Amounts

Daniela B Munafó[1], Enrique Neumann[2], Piotr Mieczkowski[3], Ewa Malc[3], Bradley W Langhorst[1], Laurie Mazzola[1], Joanna Bybee[1], Danielle Rivizzigno[1], Barton Slatko[1], Fiona J Stewart[1], Eileen T Dimalanta[1], Caroline Huber[2].

[1]New England Biolabs, Inc. Ipswich MA, USA.
[2]Tecan. Männedorf, Switzerland.
[3] West Virginia State University, Department of Genetics, University of North Carolina. Chapel Hill, NC,USA.

ABSTRACT

Next Generation Sequencing (NGS) is expanding its applications. Laboratories are increasingly implementing NGS and incrementing the number of samples to process. The ability to construct high quality sequencing libraries in a fast turnaround time has become critical. Automating sequencing library preparation reduces bottlenecks, enables higher throughput and minimizes human errors. This work describes the flexible, automated NEBNext® Ultra™ II DNA library preparation protocols on the TECAN® Freedom EVO® NGS workstation. This simple workflow allows for highly reproducible library preparation from a wide range of DNA input (picograms to a microgram), and with variable quality (from intact to heavily degraded FFPE samples). Automation of this protocol on the Freedom EVO NGS workstation enables flexible sample numbers from 1-96, and minimizes hands-on time with minimal user intervention. High input (200ng) and low input (500pg) human and yeast genomic DNA libraries generated on the Freedom EVO NGS workstation have comparable library performance (high yield, absence of adaptor dimer) to those obtained from manual libraries. NEXTSeq® sequencing data shows high quality libraries. The high percentage of aligned reads (>97.7% mapped reads and >99.03% mapped in pairs) and the low percentage of chimeras (<1%) and adaptor-mapping reads (<0.001%) observed indicate that the Tecan automation of the NEBNext Ultra II DNA Library Prep workflow enables high quality sequence data, even with very low input amounts. GC coverage information obtained indicates that automated Ultra II DNA libraries have very uniform coverage across the range of GC content. This automated method provides a much-needed resource for the reliable preparation of DNASeq libraries from a broad range of sample types and input amounts.

NEBNext® Ultra™ II DNA Workflow

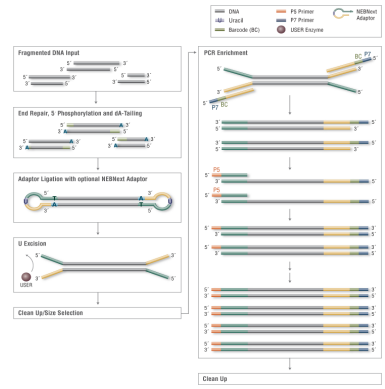


Figure 1: NEBNext Ultra II DNA Library Prep Workflow. The workflow combines the End Repair and dA-Tailing steps and minimizes clean up steps, making it easy to automate. The protocol can accommodate 500 picograms to 1 microgram of input DNA, which can be sheared by either mechanical or enzyme-based methods. The kit can also be used in PCR-free workflows. The protocol is compatible with adaptors and primers from the NEBNext product line ("NEBNext Oligos") or from other sources.

Automation Method for the TECAN EVO NGS Workstation

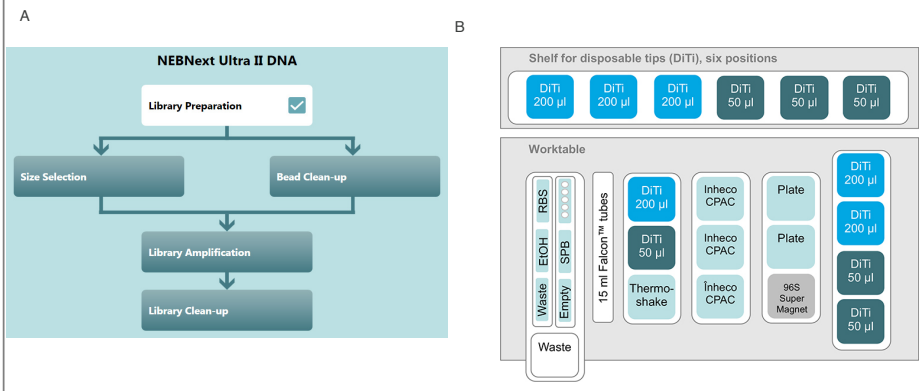
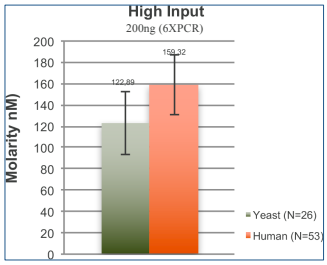


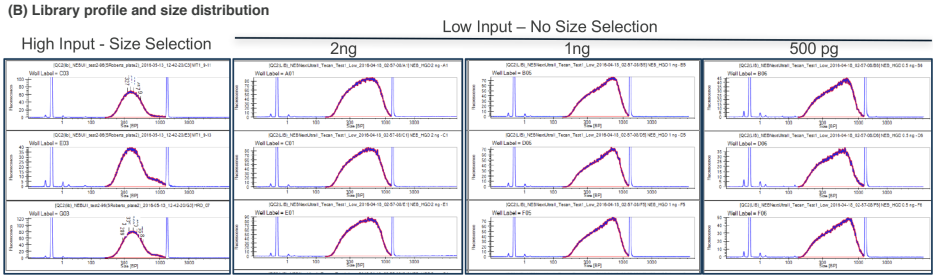
Figure 2: Automation method overview and deck layout description. (A) Simple selection of next step with TouchTools user Interface. This protocol allows highly reproducible library preparation from a wide range of input DNA concentrations, as well as FFPE samples, offering flexible processing of 1 to 96 samples with minimal user intervention. A user friendly TouchTools™ graphical interface guides users through option selection and workable set up, reducing training needs and operator to operator variability. (B) Deck layout of the Freedom EVO NGS workstation set up for NEBNext Ultra II DNA library preparation. The platform uses advanced air displacement pipetting technology, enabling precise eight channel pipetting from 1,000 µl down to just 0.5 µl. It also includes three INHE CPAC thermal devices – allowing reagents to be kept cool and providing optimal conditions for the enzymatic steps – an INHECO Thermoshake heated shaker, a 96 position magnetic plate separator (Alpaqua® 96S Super Magnet) and a Robotic Manipulator Arm. In addition, the compact worktable offers storage space for up to 12 tip boxes, allowing longer unattended runs.

Library Performance and Sequencing Metrics

(A) Library Yield



(B) Library profile and size distribution



(C) Sequencing Metrics

High Input Libraries

Replicate	Total Reads	% Mapped	% Mapped in Pairs	% Chimeras	% Adapter
Replicate 1	376,701,432	97.93%	99.03%	1.00%	0.001%
Replicate 2	383,870,510	97.74%	99.12%	0.47%	0.001%
Replicate 3	399,541,644	97.73%	99.11%	0.46%	0.001%

Low Input Libraries

Input	Total Reads	% Mapped	% Mapped in Pairs	% Chimeras	% Adapter
2ng	635,939,854	97.79%	99.06%	0.96%	0.001%
1ng	734,872,382	97.83%	99.11%	0.83%	0.002%
500pg	707,163,276	97.77%	99.12%	0.78%	0.003%

(D) Uniformity of Coverage

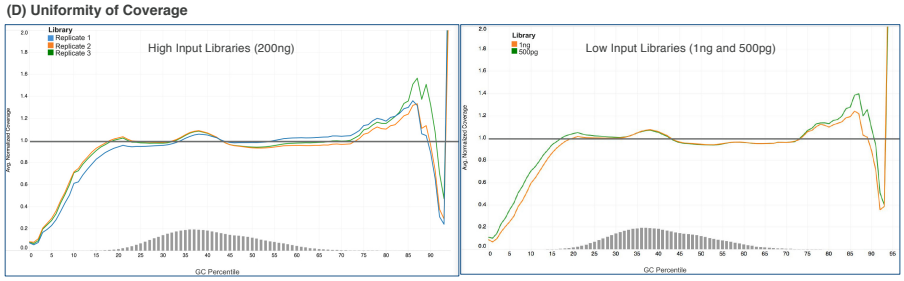


Figure 3: Library Performance and Sequencing Metrics: Genomics DNA library preparation was done according to the kit manufacturer's recommendations. Enzymatic reaction set-up, bead clean-up and size selection were all performed on the work deck. End repair, dA-tailing and PCR amplification steps were performed in an offline thermocycler. High input libraries were prepared from 200 ng of fragmented human genomic DNA in a single plate. Size-selected library preparation (200 bp insert size) followed the recommended DNA:bead ratios, and amplified with six PCR cycles. Low input libraries were prepared from 2 ng, 1 ng and 500 pg fragmented human genomic DNA in a single plate. Library size distribution and yield was assessed on a Caliper LabChip GX System (software v4.1). Automation achieves consistently high library yield (A). As expected, the size selected (high input) libraries have a narrow size distribution with a mean size distribution of 320-340bp, which corresponds to a 200bp fragment insert size. The size distribution of the non-size selected (low input) libraries correlates to the size distribution of the input DNA (B). (C) Sequencing Metrics. Three randomly selected high (200ng) and low input human DNA libraries were sequenced on an Illumina NextSeq® sequencer, generating approximately 400 million (2x75bp) paired-end reads per library, and 700 million (2x75bp) paired-end reads per library, respectively. Reads were mapped to GRCh37 reference using Bowtie 2.2.4 with standard end-to-end settings. The high percentage of aligned reads and low percentage of chimeras and adaptor-mapping reads indicate that the automated protocol enables the generation of high quality sequencing data, even with very low input amounts. (D) Uniformity of Coverage. GC coverage was calculated using Picard's CollectGCBiasMetrics (v1.117). The results show that the automated Ultra II DNA libraries have very uniform coverage across the range of GC content regardless the input amount.

CONCLUSIONS

The results presented in this poster demonstrate that automation of the NEBNext Ultra II DNA library preparation workflow on the Freedom EVO NGS workstation provides a fast and efficient solution for library preparation. This setup enables generation of high quality libraries from a broad range of input samples – from 500 pg to 1 µg – while reducing the number of PCR cycles required. The TouchTools interface ensures a user friendly experience, reducing training needs, minimizing the risk of manual errors and increasing process reproducibility. Combined with flexible processing of up to 96 samples (with or without size selection) and a number of safe stopping points within the protocol, this setup provides highly reproducible, sequencing ready libraries to suit a variety of laboratory workflows.