Sequence ready library preparation fully automated for microbial and human DNA template

Authors
Jethary Rader and Kinnari Watson

ABSTRACT

The Swift 2S Turbo DNA Library Kit was fully automated using the Opentrons OT-2 for quick, hands-off, high-quality libraries for Illumina Next-Generation Sequencing.

Key Findings

- Eight sequence ready DNA libraries can be constructed in under 3 hours.
- Comparable target insert sizes of 326bp and 330bp were analyzed across libraries of microbial and human genomic DNA.
- Low PCR bias, illustrated by negligible PCR duplicates for both microbial and human libraries at levels of 0.02% and 0.01% respectively.
- High DNA coverage retained throughout the OT-2 run, depicted through similar percent of reads aligned: the microbial libraries with 95.4% and the human libraries with 96.4%.
- Dependable and precise performance of the Opentrons Thermocycler module when compared to a third party thermocycler, demonstrated through similar yields of 5.1 ng/µl and 7.36 ng/µl, and CVs of 13% and 14% respectively.

INTRODUCTION

Next-generation sequencing (NGS) library preparation is a process by which RNA or DNA fragments are enzymatically prepared for sequencing. The usage of NGS library preparation is rapidly increasing, providing researchers in healthcare and life sciences with a variety of workflows to achieve faster, easier, and more cost effective ways to analyze and interpret large genomic datasets (1). The Swift 2S Turbo DNA Library Prep supports the library preparation of samples for various molecular applications. This automation-friendly workflow is condensed, but completing the workflow manually is still subject to pipetting errors and requires several hours to complete. Here we illustrate the full automation of the Swift 2S Turbo library preparation on microbial and human genomic DNA. Complex library preparation can be completely automated with the Opentrons OT-2, modules, labware, and Omega Bio-tek Mag-Bind® TotalPure NGS beads.

MATERIALS AND METHODS

Libraries were prepared for 350bp insert sizes to make ~560bp libraries using human genomic DNA (Coriell NA12878) and E. coli (ER2925) gDNA with 40 and 50% GC content, respectively. Each Opentrons OT-2 run contained 7 replicates and each run included 5 or 10 PCR cycles, depending on starting input. The ligation mastermix used included a 1:10 diluted Reagent W4 to reduce adapter dimers. Each library was sequenced on the Illumina MiSeq Nano v2 2x250 kit. The Opentrons Thermocycler Module and Temperature Module were used for active cooling of the mastermixes, indexes, and samples in between incubations, and the Magnetic Module was used for bead based clean-ups using the Omega Bio-tek Mag-Bind® TotalPure NGS beads (2). These modules collectively allowed for full automation of this compound workflow (Figure 1).

The OT-2 used the P300 GEN2 8-Channel Pipette and the P50 GEN2 Single-Channel Pipette for the run with the E. coli gDNA, and the P300 GEN2 8-Channel Pipette and the P20 GEN2 Single-Channel Pipette for the run with the Human gDNA. However, we recommend the P20 GEN2 Single-Channel Pipette. For each run, 3 tipracks were used (Figure 1). Total run time, including hands-on time and robot run time was around 3 hours (Figure 2). Peak fragment sizes were identified using an Agilent 2100 BioanalyzerTM Instrument. FASTQ files were

analyzed with FASTQC V0.11.9 (3), BOWTIE2 (4), Qualimap-BAMQC (5), and BamCoverage V2.4.1.0 (6) in Galaxy (7) to determine coverage, insert size, reads aligned to genome, GC content, and percent duplicates. BWA Aligner V1.1.4 on Illumina's Basespace Sequencing Hub was used for further analysis.

RESULTS

Library yields were highly reproducible and comparable across samples within the same run and across species. The input was 100 ng for E. coli gDNA and 77 ng for human gDNA. The E. coli libraries workflow included 10 PCR cycles and had an average of 20.34 ng/µl concentration with a 7% coefficient of variation (CV). The human gDNA libraries workflow included 5 PCR cycles and had an average of 5.1 ng/µl concentration with a 13% CV. The libraries were comparable across samples, with an average peak fragment size of 591 bp and a CV of 7% for the E. coli libraries and an average peak fragment size of 569 bp and a CV of 4% for the human libraries. Yield and peak fragment size CV demonstrate both consistency across wells and different species (Figure 3).

Sequencing metrics were also comparable across samples and across species. Average insert sizes for the E. coli and human libraries of 326 and 330 bp, correspond with the target length of DNA that has been incorporated for direct sequencing, according to the reagent specifications. Both experiments showed about 95% reads mapped when aligned to the reference genome sequences and high DNA recovery and size selection were retained throughout each run, illustrated by very low PCR duplicates (Figure 4). Consistent coverage across the E. coli samples (Figure S1) and the human samples (Figure S2) showcases equivalent performance as well as sequence representation. Also, the single nucleotide variation (SNV) rs6061194, was characterized amongst the human libraries, which further demonstrates performance (Figure S3). Finally, the Opentrons Thermocycler Module performed similarly to a manual third party thermocycler (8), with commonalities across yields and CV. The yields were 5.1ng/µl and 7.36ng/µl and the CVs were 13% and 14%, demonstrating the precision and reliability of this end-to-end Opentrons library preparation system (Table 1).

CONCLUSION

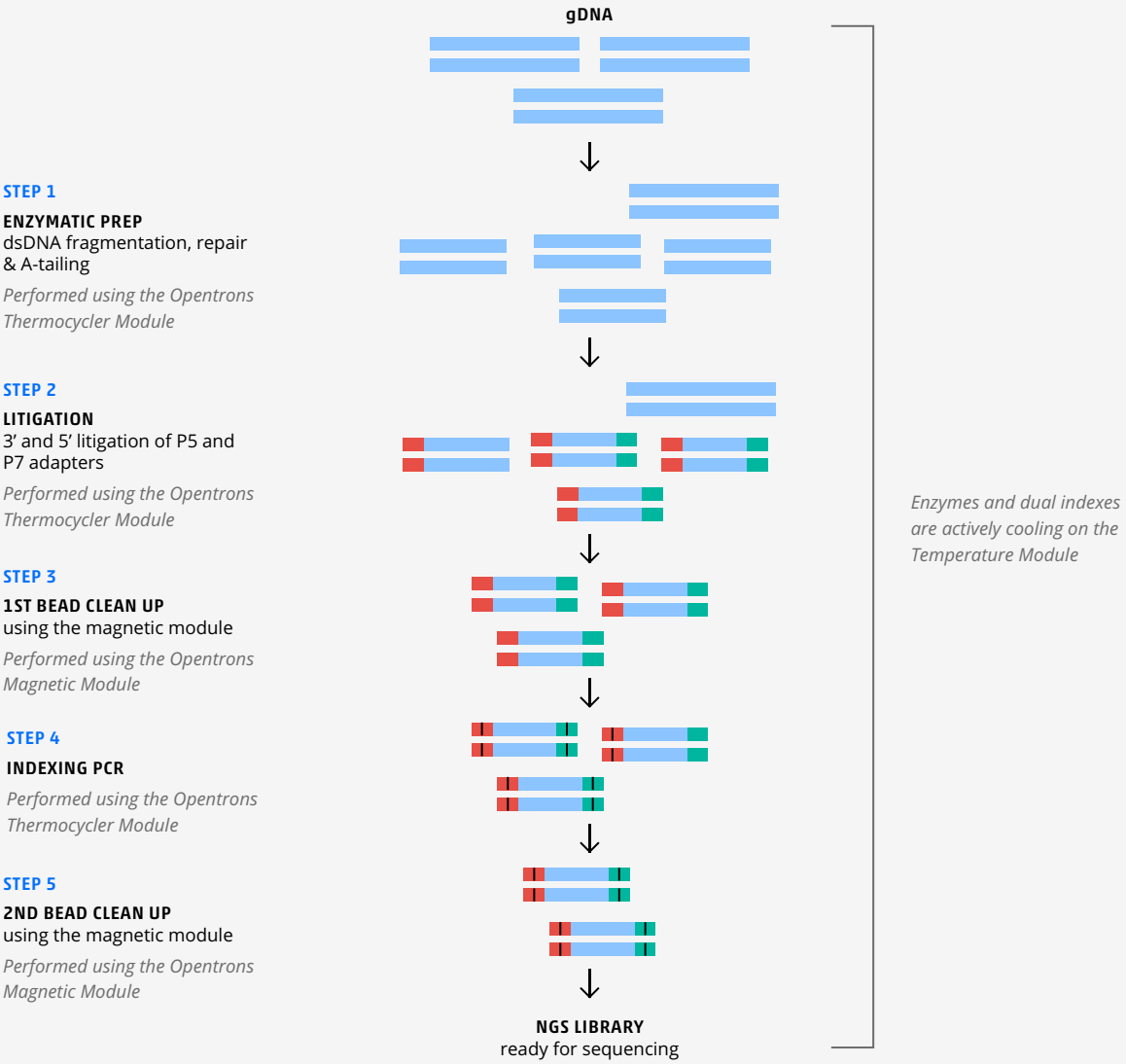
The Opentrons OT-2 enables high quality libraries of eight or more at a time, in a minimum of less than three hours, ultimately reducing the risk of manual pipetting errors and saving the user time.

- Reproducible and consistent sequencing metrics across the microbial and human libraries in terms of yield, CV, insert size, reads aligned, PCR duplicates, and coverage.
- Reliable and precise performance of the Opentrons Thermocycler Module when yield and CV are compared to a third party thermocycler.

REFERENCES

1. Giani, Alice Maria, et al. “Long walk to genomics: History and current approaches to genome sequencing and assembly.” *Computational and Structural Biotechnology Journal (2019).*
2. Lee, Chris, et al. “Fully Automated NGS Clean Up and Size Selection With Omega Bio-Tek Mag-Bind Totalpure NGS On The OT-2.” (2019).
3. Andrews, Simon. “FastQC: a quality control tool for high throughput sequence data.” (2010).
4. Langmead, Ben, et al. “Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.” *Genome biology* 10.3 (2009): R25.
5. Okonechnikov, Konstantin, Ana Conesa, and Fernando García-Alcalde. “Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data.” *Bioinformatics* 32.2 (2016): 292-294.
6. Ramírez, Fidel, et al. “deepTools2: a next generation web server for deep-sequencing data analysis.” *Nucleic acids research* 44.W1 (2016): W160-W165.
7. Afgan, Enis, et al. “The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update.” *Nucleic acids research* 46.W1 (2018): W537-W544.
8. Stedtfeld, Robert, et al. “Rapid high quality next generation sequencing library preparation with Swift 2S™ Turbo DNA Library Kits on the Opentrons OT-2.” (2020).

FIGURE 1



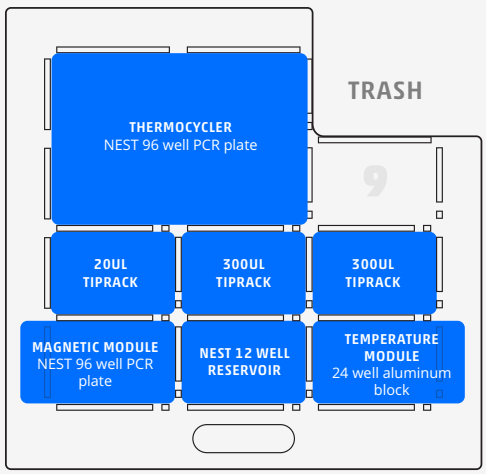


Figure 1. Swift 2S Turbo workflow and OpSwift 2S Turbo DNA Library Kit© protocol. This layout includes the P20 GEN2 Single-Channel Pipette and the P300 GEN2 8-Channel Pipette. Module requirements include the Opentrons Thermocycler Module, Temperature Module, and Magnetic Module. The labware requirements include two NEST 0.1ml 96 Well PCR Plate Full Skirt, one NEST Deep 12 Well Reservoir, three NEST 2ml Screw Cap tubes, and the Opentrons 24 well Aluminum block.

FIGURE 2

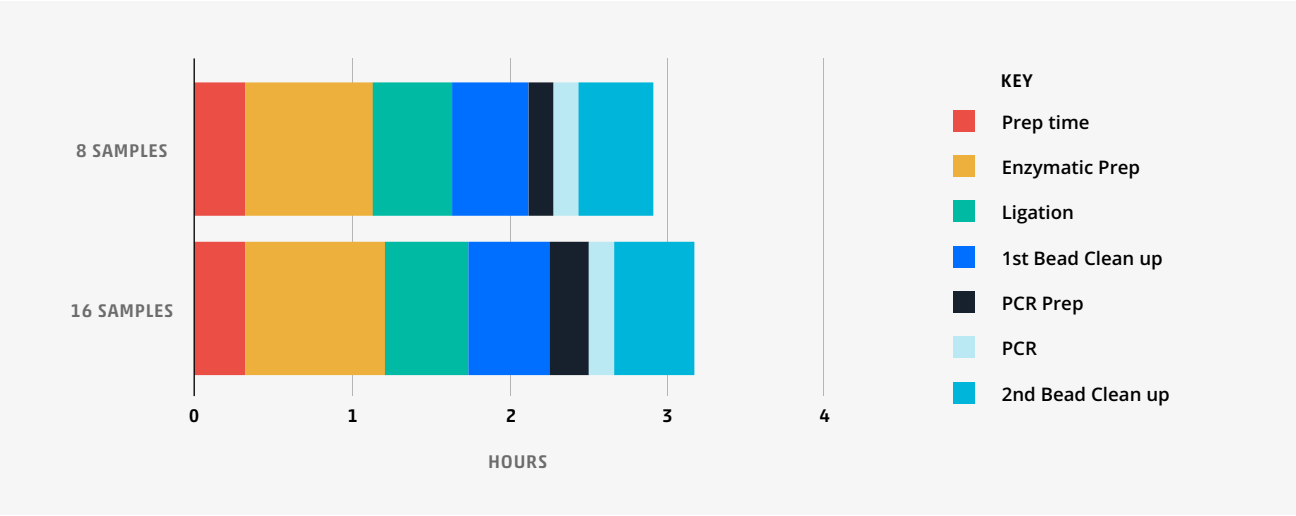


Figure 2. Total workflow break down. The prep time was around 20 minutes and the robot run time was 2 hours and 37 minutes for 8 samples and 2 hours and 48 minutes for 16 samples. This fully automated workflow included a 10 minute fragmentation and 5 PCR cycles.

FIGURE 3

OT-2 SAMPLE NUMBER	E. COLI YIELD (NG/UL)	HUMAN YIELD (NG/UL)
1	21.5	6.24
2	20.7	4.34
3	21.2	4.96
4	22.3	5.22
5	17.7	5.48
6	20.3	4.22
7	18.7	5.28

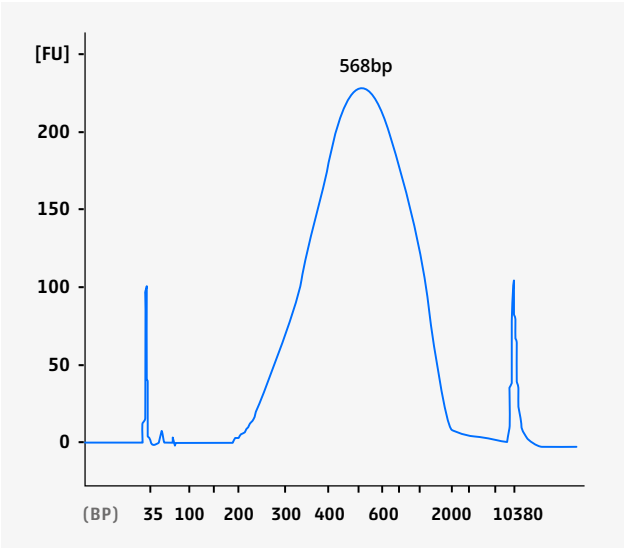


Figure 3. Commensurable yield and fragment size across libraries. The E. coli libraries had 10 PCR cycles and a yield CV of 7% was observed while the Human libraries had 5 PCR cycles and a yield CV of 13% was observed. The E. coli libraries average fragment size was 591bp with a 7% CV and the Human libraries average fragment size was 569bp with a 4% CV.

FIGURE 4

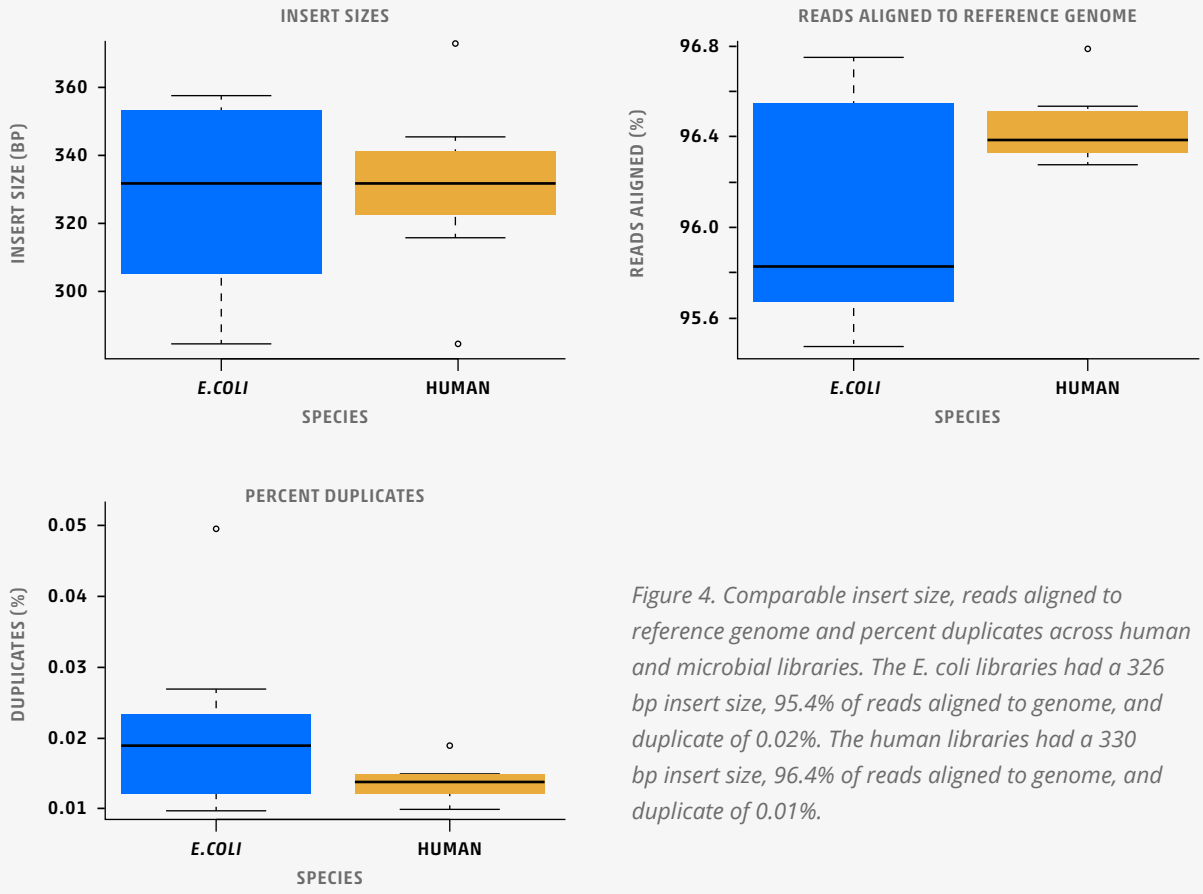


Figure 4. Comparable insert size, reads aligned to reference genome and percent duplicates across human and microbial libraries. The E. coli libraries had a 326 bp insert size, 95.4% of reads aligned to genome, and duplicate of 0.02%. The human libraries had a 330 bp insert size, 96.4% of reads aligned to genome, and duplicate of 0.01%.

TABLE 1

	TYPE OF GDNA	INPUT (NG)	AVERAGE YIELD (NG/UL)	CV (%)
OPENTRONS THERMOCYCLER	Human (Coriell NA12878)	77	5.1	13
THIRD PARTY THERMOCYCLER (4)	Staphylococcus aureus, Escherichia coli, and Streptomyces avermitilis	100	7.36	14

Table 1. Similarly reliable preparation across human and microbial libraries with the OT-2, carried out in separate laboratories. Yields were 5.1 ng/µl and 7.36 ng/µl and the CVs were 13% and 14% respectively, demonstrating the precision and dependability of the Opentrons Thermocycler.

NOTE: SUPPLEMENTAL DATA AVAILABLE UPON REQUEST