Application Note

A Practical Protocol for Library Preparation of Samples Sheared in the Covaris® microTUBE-15 using Rubicon Genomics' ThruPLEX® DNA-seq Kit

John Paul Jerome, Sean Carey, Hamid Khoja, Guillaume Durin, and Kamran Shazand

Introduction

Robust library preparation methods that ensure high quality libraries from a wide range of samples are critical for research applications as next generation sequencing progresses to the clinic. In translational genomics where NGS is used extensively with oncological samples, the emphasis is on compatibility with a wide range of input types, quantities, and quality. Also important is the ability to accommodate various levels of sample degradation, especially with FFPE tissue, and overall simplicity of the workflow. In this application note, we describe a library preparation workflow that combines Rubicon Genomics' ThruPLEX® DNA-seq Kit and the Covaris® truSHEAR™ microTUBE-15 for a fully optimized low DNA input solution.

The Covaris truSHEAR mechanical shearing technology is recognized as the gold standard for high quality, robust, and highly reproducible fragmentation of nucleic acids. It employs patented Adaptive Focused Acoustics™ (AFA) technology for the controlled and easily tunable shearing of nucleic acids. Compared to enzymatic-based fragmentation or tagmentation, AFA hydrodynamic shearing generates a tight and completely random fragmentation fundamental to obtaining uniform genome coverage[1, 2] and high complexity libraries[3].

AFA-based hydrodynamic shearing is highly versatile with the unique ability to deliver the same level of performance across a broad combination of DNA input sources, concentrations, and volumes. The AFA microTUBE-15 is optimized for a 15 µl sample volume and is ideal for the ThruPLEX DNA-seq input requirements. The microTUBE-15 incorporates AFA-Beads that enable fully controllable, easy-to-use, and highly reproducible DNA shearing in very low sample volume with high recovery.

The ThruPLEX DNA-seq library preparation kit uses Rubicon Genomics' patented stem-loop technology to repair DNA, reduce background, and generate high complexity libraries. The ThruPLEX DNA-seq Kit can generate DNA libraries from as little as 50 pg of DNA while providing up to 96 indexes for multiplexing. This kit can be used with fragmented DNA from any sample source – biofluids such as cell-free DNA, DNA from FFPE materials, and cDNA[4]. The entire ThruPLEX DNA-seq Kit workflow is performed in a single tube or well in about 2.5 hours and requires no intermediate purification steps or sample transfers. It can be used in a variety of applications including DNA-seq, RNA-seq, and ChIP-seq and offers robust target enrichment performance with all of the leading platforms.

The combination of the AFA-based DNA mechanical shearing and the library preparation with ThruPLEX DNA-seq enables a uniquely simple, fast, and robust workflow optimized not only for routine use, but also for precious and rare samples as DNA input can be as low as 50 pg. Total processing time of the present workflow is less than 3 hours (Figure 1).

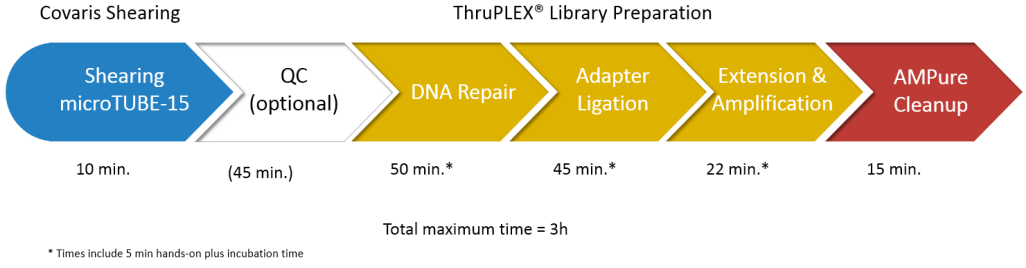

*Figure 1: Complete DNA library preparation workflow beginning with shearing in the Covaris microTUBE-15 followed by the Rubicon Genomics' single tube, low volume, ThruPLEX repair, extension, and amplification workflow. Total workflow time is less than 3 hours. *Each step represents total time including 5 minutes hands-on plus incubation time.*

RUBICON GENOMICS

Covaris

Materials & Methods

DNA Mechanical Shearing

Human genomic DNA NA12878 (Coriell Institute for Medical Research) was quantified by Qubit® using the dsDNA HS Assay Kit (Thermo Fisher Scientific) and diluted to 5 pg/µL, 100 pg/µL, and 1 ng/µL. A 15 µL aliquot of each diluted DNA was loaded into Covaris microTUBE-15 AFA Beads (PN 520145) or 130 µL was loaded into control Covaris microTUBE AFA Fiber Snap-Cap (PN 520045). All mechanical shearing was performed in the Covaris model M220 Focused-ultrasonicator with Holder XTU (PN 500414) using the appropriate insert for each type of shearing volume: M220 Holder XTU Insert microTUBE 130 µL (PN 500489) or M220 Holder XTU Insert microTUBE 15 µL (PN 500420). DNA was sheared to 150 bp and 350 bp in duplicate for each concentration. For processing, pre-programmed settings in the Covaris SonoLab 7.2 software were used.

DNA Quality Control

After shearing, 1 µL of each 1 ng/µL sample was visualized on the Agilent Bioanalyzer® using the High Sensitivity DNA Analysis kit (lower concentrations were below the limit of detection of Bioanalyzer).

Library Preparation

For library preparation, 10 µL of each sheared sample was used as input for the ThruPLEX DNA-seq Kit resulting in starting gDNA inputs of 50 pg, 1 ng, and 10 ng. After library purification, ThruPLEX DNA-seq libraries were quantified by Qubit and qPCR, analyzed for size distribution on the Bioanalyzer, and sequenced using Illumina® MiSeq® V3 flow cell and reagents (paired end 2x75 base reads).

Data Analysis

Data was analyzed in DNANexus. Briefly, reads were mapped to the human genome reference, hg19, using the Burrows-Wheeler Alignment algorithm[5] by BWA-MEM to generate BAM files for each sample. Duplicates were marked and removed using Picard Mark Duplicates[6] and quality of libraries was further analyzed using Picard Tools and BEDTools.

Results & Discussion

Experimental Design

The main purpose was to create a straightforward workflow combining the shearing and library preparation steps without need for additional sample manipulation or concentration adjustment while maintaining comparable results to more tedious protocols.

A summary of the fragment sizes and tubes is indicated in Table 1. For the shearing step, each sample was prepared in duplicate with a total of 24 samples. The library preparation step was then completed in triplicate for each sample.

Fragment size	150 bp	350 bp
50 pg (5 pg/µL)	15 µL microTUBE-15; 130 µL standard vol.	15 µL microTUBE-15; 130 µL standard vol.
1 ng (0.1 ng/µL)	15 µL microTUBE-15; 130 µL standard vol.	15 µL microTUBE-15; 130 µL standard vol.
10 ng (1 ng/µL)	15 µL microTUBE-15; 130 µL standard vol.	15 µL microTUBE-15; 130 µL standard vol.

Table 1: Three different DNA input amounts in two different average sizes were prepared for downstream library preparation.

Quality Control of Sheared DNA

Sheared DNA was analyzed on Bioanalyzer (Agilent Technologies) to measure the size and reproducibilty of shearing. Examples of each category are shown below (Figure 2).

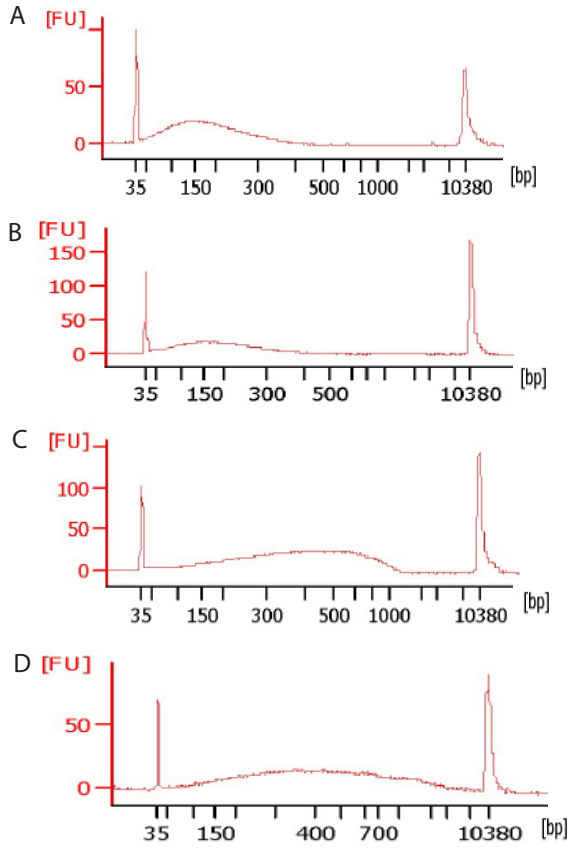

Figure 2: Examples of 1 ng sheared DNA. A: 150 bp in15 µL microTUBE-15; B: 150 bp in 130 µL microTUBE; C: 350 bp in15 µL microTUBE-15; and D: 350 bp in130 µL microTUBE. Peaks at 35 bp and 10380 bp are BioAnalyzer High Sensitivity DNA Markers.

RUBICON GENOMICS

Covaris

Input	50 pg				1 ng				10 ng			
Size	150bp		350bp		150bp		350bp		150bp		350bp	
DNA Shearing volume	130	15	130	15	130	15	130	15	130	15	130	15
Estimated Library Size	2.62E +07	1.65E +07	8.31E +06	7.64E +06	3.12E +08	2.05E +08	2.46E +08	1.16E +08	1.39E +09	1.16E +09	1.07E +09	9.97E +08
% Duplicate	0.6	1.00	1.92	2.17	0.06	0.08	0.08	0.22	0.02	0.02	0.03	0.03
% Unmapped Reads	0.75	0.97	0.85	0.89	0.30	0.29	0.35	0.33	0.27	0.28	0.34	0.33

Table 2: Key sequence metrics of different libraries calculated by the Picard pipeline. Numbers are average values of duplicates for each sample. First line: Estimated Library size reflects the diversity of each library. Percent duplicate: corresponds to the identical reads generated by PCR. Percent Unmapped reads: a quality metric that shows the absence of contamination in the human DNA prep.

All 24 DNA samples were used to prepare libraries with the ThruPLEX DNA-seq Kit. Libraries were then analyzed on the Bioanalyzer. Examples are shown in Figure 3.

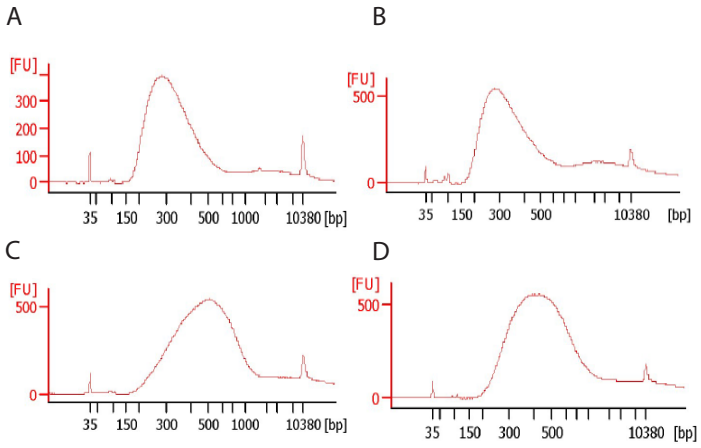

Figure 3: Examples of libraries with 2 insert sizes and different tubes. ThruPLEX DNA-seq library preparation adds approximately 140 bases to the DNA molecules. A: 150 bp in 15 µL microTUBE-15 ; B: 150 bp in 130 µL microTUBE ; C: 350 bp in 15 µL microTUBE-15 ; and D: 350 bp in 130 µL microTUBE.

After quantification, a fraction of each library was pooled, loaded on an Illumina MiSeq instrument, and sequenced following a paired-end 2x75 bp protocol. Data was collected and down-sampled to 156K read pairs for analysis. A comparative summary of some key metrics is shown in Table 2 and Figure 4.

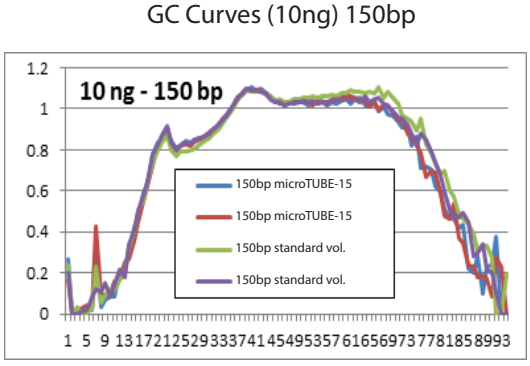

Figure 4: An example of the GC bias calculated over a window of 100 bp reads. Each curve represents the average distribution of two replicates for each sample.

We also evaluated if shearing volume, input amounts, or fragment size had any impact on the sequencing data metrics. To do so, we plotted the estimated library size as a function of read pairs by differentially pooling our data set (Figures 5 and 6).

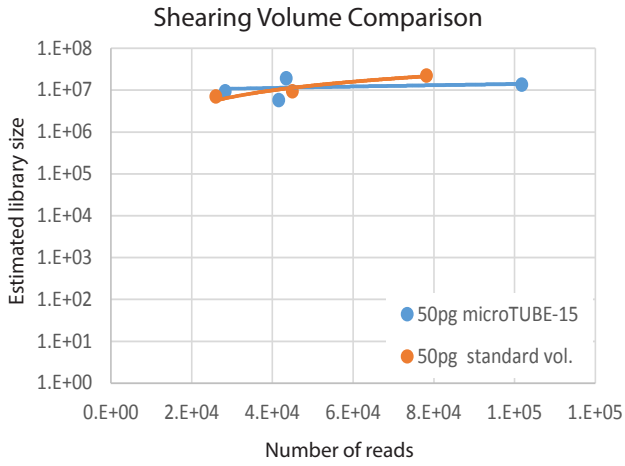

Figure 5: Estimated library size comparison between samples sheared in the micoTUBE-15 and standard 130 µL tube. The plot shown for 50 pg is identical to the plots of both 1 ng and 10 ng (data not shown).

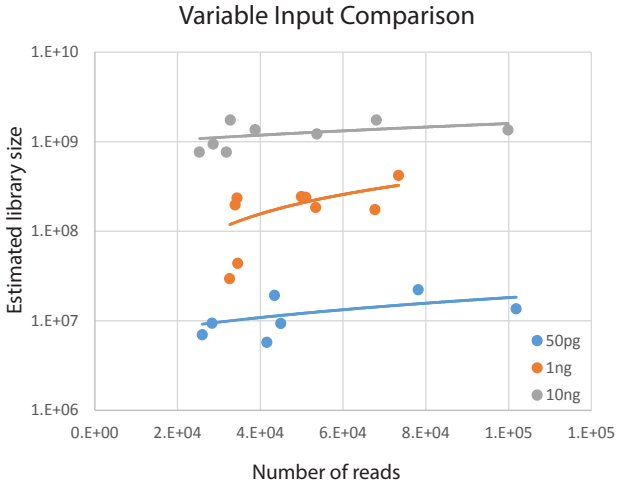

Figure 6: Comparison of estimated library sizes using a pool of both shearing volumes (standard vol. and microTUBE-15) and various input amounts (50 pg, 1 ng, and 10 ng)

RUBICON GENOMICS

Covaris

The superimposed lines shows equivalency between the two types of tubes used to shear the DNA indicating shearing volume does not affect the estimated library size (Figure 5). As expected, varying the input amount results in a difference in library size indicating that libraries are more complex as the input amounts increase, as indicated by the comparable flat slopes (Figure 6). Together, Figures 5 and 6 strongly suggest that in this range of read pairs, the estimated library size is stable regardless of the shearing volumes used.

In consideration of these data, it is clear that the use of the microTUBE-15 in combination with ThruPLEX DNA-seq Kit presents a much simpler and more practical low sample volume workflow to prepare DNA libraries without compromising library quality, quantity, or complexity. The workflow described here is fully functional for any application requiring sheared DNA for Illumina library preparation and is especially useful for applications where the DNA is of limited amount and subject to loss and/or contamination.

References

1. Illumina sequencing of 15 deafness genes using fragmented amplicons. Van Nieuwerburgh et al. BMC Res Notes. 2014 Aug 9; 7:509. doi: 10.1186.
2. Impact of three Illumina library construction methods on GC bias and HLA genotype calling. Lan et al Hum Immunol. 2015 Mar; 76(2-3):166-75. doi: 10.1016.
3. Comparison of Custom Capture for Targeted Next-Generation DNA Sequencing Samorodnitsky, Eric et al. The Journal of Molecular Diagnostics, Volume 17, Issue 1, 64-75.
4. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Murtaza M et al. Nature. 2013 May 2;497(7447): 108-12. doi: 10.1038/nature12065.
5. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009, 25, 1754-1760.
6. Picard - A set of tools (in Java) for working with next generation sequencing data in the BAM format. http://broadinstitute.github.io/picard.

Trademarks