Automating Library Preparation for PacBio® HiFi Sequencing on Opentrons Flex®

Written by
D'Andra Howell,[1] Matthew Akana,[1] Christine Lambert,[2]
Gregory Young,[2] and Kinnari Watson, PhD[1]
[1]Opentrons Labworks, Inc., [2]Pacific Biosciences

ABSTRACT

Long-read sequencing has been a game changer for scientists assembling whole genomes. One method for long-read sequencing is HiFi sequencing from PacBio, which delivers short read like accuracy but at multi-kilobase read lengths. Preparing samples for HiFi sequencing on the Opentrons Flex liquid handler involves two methods: (1) eliminating short DNA via their Short Read Eliminator (SRE) kit and mechanical fragmentation on the remaining DNA, and (2) constructing the DNA library for HiFi sequencing. In this application note, we demonstrate how automating this workflow on the Opentrons Flex liquid handler produces high quality libraries with less user intervention and hands-on time.

Key features:

- Automated SRE, DNA shearing, and library prep in eight hours with only one hour of hands-on time.
- Automated library prep delivers similar high-quality results as manual methods, meeting PacBio benchmarks for yield, barcoding, and base quality.

INTRODUCTION

PacBio long-read HiFi sequencing technology offers accurate and reliable sequencing of DNA libraries ranging from 0.5 kb to over 25 kb in length. In addition, the technology simultaneously detects base modifications enabling characterization of both the genome and epigenome. This empowers researchers to deliver a *de novo* genome assembly, phase haplotypes, detect complex variation, and classify the epigenetic states from methylation information and chromatin accessibility from the Fiber-seq assay – all from a single experiment.

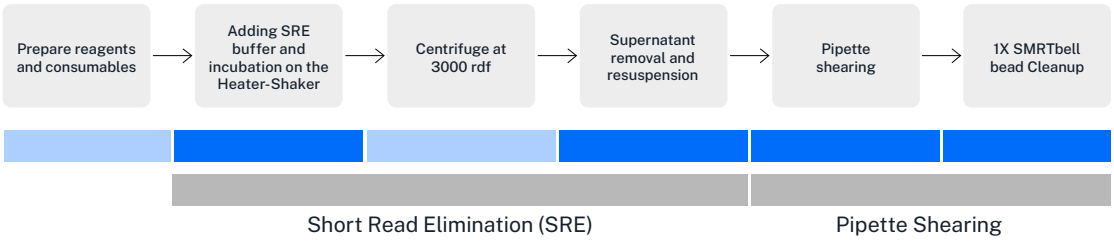
The HiFi library preparation workflow begins with a step to improve sample quality by depleting degraded, short DNA fragments from the sample using SRE. This step can be optional if starting with high quality DNA or working with limited DNA amounts. Next a fragmentation step (i.e. shearing) is performed on the remaining high molecular weight DNA to optimize the size distribution for HiFi sequencing. The workflow then goes on to library prep involving a series of enzymatic reactions like traditional NGS workflows: end repair, A-tailing, and adapter ligation. After ligation a nuclease reaction is employed to remove unligated fragments. The final product is a circular molecule called a SMRTbell® library. To facilitate the process, we have adapted this workflow on the Opentrons Flex using two sequential protocols: (1) eliminating short reads and shearing DNA to the appropriate size, and (2) library preparation of the sheared DNA.

Automating this workflow with liquid handlers such as the Opentrons Flex liquid handler has many advantages including: (1) reducing hands-on time for scientists, (2) reducing user errors, and (3) eliminating the need for specialized DNA shearing equipment by using automated pipette shearing.

Here, we demonstrate how automation can be leveraged to generate HiFi sequencing libraries in 8 hours with only 1 hour of hands-on time (Figure 1).

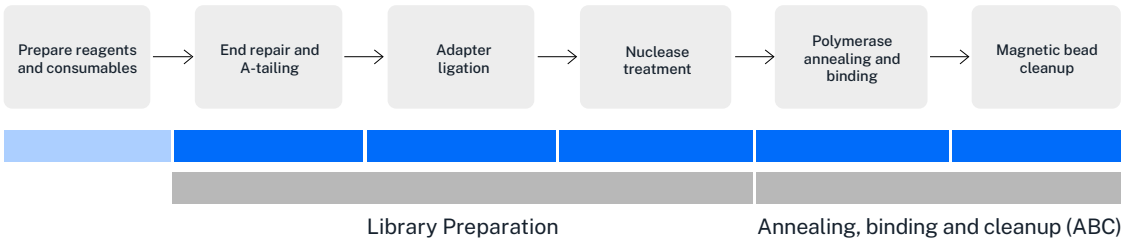
METHODS

SRE and Pipette Shearing Workflow



8x: 3.5 hours total time
48x: 6.5 hours total time
~20 minutes hands-on time

HiFi Library Preparation Workflow



8x: 4.5 hours total time
48x: 8.5 hours total time
~30 minutes hands-on time

Figure 1. Workflow steps for the short read elimination and DNA fragmentation protocol (top) and the HiFi library preparation protocol (bottom). Manual steps (in light blue) and automated steps on the Opentrons Flex (in dark blue) are shown below each stage of the workflow.

Sample preparation: HMW DNA was sourced from the Coriel Institute for Medical Research (cat. #: HH01891), and from human saliva and whole blood using the PacBio Nanobind PanDNA kit (cat. #: 103-260-000).

Short read elimination (SRE) and DNA fragmentation: We used the PacBio SRE HT kit (cat. #: 103-124-500) to eliminate short reads and pipette shearing to fragment HMW DNA to 15-20 kb on the Opentrons Flex liquid handling robot using the PacBio Short Read Elimination and Pipette Shearing protocol (library.opentrons.com/p/pacbio-sre-shearing).

Analysis of fragmented DNA: DNA before and after shearing was analyzed on the Agilent Femto Pulse system (cat. #: M5330AA) using the Genomic DNA 165 kb kit (cat. #: FP-1002-0275).

WGS HiFi library preparation: Fragmented DNA (≤ 3 µg) was used to generate sequence-ready HiFi libraries with the PacBio HiFi prep kit 96 (cat. 103-381-200) and Revio SPRQ polymerase kit 96 (cat. #:103-497-000). These reagents can be purchased together as the Revio® SPRQ™ HiFi prep kit 96 (cat. # 103-522-600) This protocol was automated on the Opentrons Flex NGS Workstation (cat. # 991-00354) using the PacBio WGS HiFi Library Preparation protocol (library.opentrons.com/p/pacbio-hifi-prep).

HiFi sequencing: HiFi libraries were sequenced using the PacBio Revio system, with one library complex loaded per SMRT® Cell. Sequencing movies ran for approximately 24 hours using a Revio SPRQ sequencing plate and polymerase kit. Libraries were loaded at 250 pM concentration with adaptive loading.

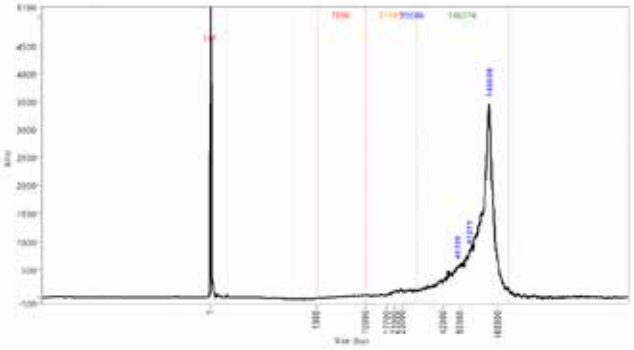
Analysis of HiFi sequencing library: DNA was quantified using the Qubit™ 1X dsDNA High Sensitivity Assay Kit (cat. #: Q33231) and the Qubit 4 Fluorometer to assess input DNA mass, HiFi library mass, and library-complex mass after annealing, binding, and cleanup steps. Raw sequencing data was processed using SMRT® Link analysis Software to generate quality control reports and determine percent barcoded reads detected, barcode/sample cross-contamination, and base quality.

RESULTS AND DISCUSSION

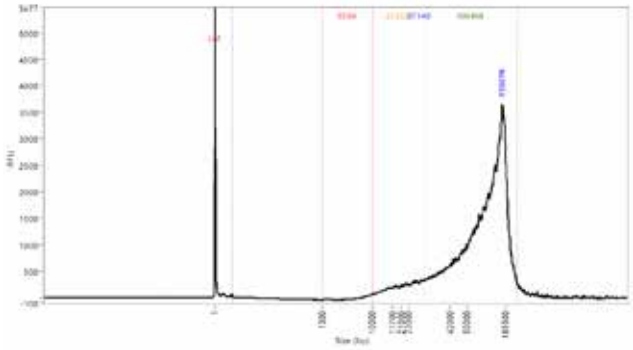
Short Read Elimination and DNA Shearing

To prepare HMW DNA for PacBio HiFi library preparation and sequencing, we used the PacBio Short Read Elimination and Pipette Shearing workflow (library.opentrons.com/p/pacbio-sre-shearing) to automate the workflow on the Opentrons Flex. This protocol first progressively depletes fragments up to 25 kb, which improves sample quality particularly in cases where the sample is partially degraded. After short fragments are removed, the high molecular weight DNA is then pipette sheared to 15-20 kb and cleaned up with SMRTbell cleanup beads. We analyzed the fragments using the Agilent Femto Pulse system. Figure 2A shows the input DNA and Figure 2B displays the output after SRE. Figure 2C displays the size profile after pipette shearing. As demonstrated, the DNA after this workflow is around the ~15 kb (average 15,472 bp, peak at 14,739 bp), the ideal size needed as input for PacBio HiFi library preparation.

(A) Input HMW DNA



(B) HMW DNA after short read elimination



(C) HMW DNA post-shearing and 1x SMRTbell bead cleanup

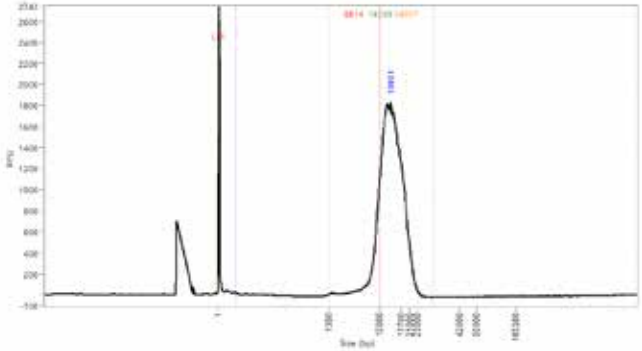


Figure 2. Analysis of HMW fragment sizes during SRE and pipette shearing. (A) Input HMW DNA before beginning the workflow. (B) HMW DNA after SRE. (C) HMW DNA after shearing and 1x SMRTbell bead cleanup.

An advantage of automating this process on the Opentrons Flex is that it avoids the need to purchase specialized equipment to shear DNA. Instead, we show that we were able to generate fragmented DNA using pipette shearing, streamlining the process and reducing costs. In total, we were able to complete SRE and shearing of five samples in 3.5 hours.

WGS HiFi Library Prep

The next step in the workflow is to prepare HiFi DNA libraries using the PacBio WGS HiFi Library Preparation workflow (library.opentrons.com/p/pacbio-hifi-prep). Key metrics to analyze the success of the library prep would need to compare the amount of DNA recovered after preparing the libraries (library mass) and loading the polymerase (ABC complex mass) to the amount of input DNA. We found that the total recovery, which is the percentage of input DNA remaining after loading the polymerase, was in line with the expected recovery of 10-25% (Table 1).

Table 1. DNA mass and recovery % after the WGS HiFi library prep protocol. Input mass represents the DNA mass after fragmentation. Library mass is the amount of DNA recovered after preparing the library and before complexing with polymerase. ABC (annealing, binding, cleanup) mass is the amount of DNA complexed to polymerase after final clean up. Total recovery % = (ABC complex mass) / (input mass) x 100.

Sample ID	Input Mass (ng)	Library Mass (ng)	ABC Complex Mass (ng)	Total Recovery (%)
Coriel_1	1070	178.5	141	13.2
Coriel_2	1120	180.7	146	13
WB_1	1049.6	224.25	207.25	19.7
WB_2	1090.6	286.5	280	25.7

After preparing the libraries, the next step was to sequence the libraries and access quality control metrics. To do this, we first looked at the barcodes. Each sample is assigned a unique barcode so that the sample can be identified. We found that of our five replicates, each sample contained one barcode indicating that there was no cross-contamination between samples (Table 2). On average, 93.4% of reads contained a barcode.

Table 2. HiFi sequencing metrics

Sample ID	HiFi Yield (Gb)	P1 Loading	HiFi Read Length N50 (bp)	Base Quality ≥Q30 (%)	Fold Coverage
Coriel_1	98.33	59.17	17.18	93.88	30.7
Coriel_2	113.71	64.49	15.14	95.14	35.5
WB_1	116.25	72.51	15.76	93.79	36.3
WB_2	116.29	72.14	15.52	93.7	36.3

Next, we turned to sequencing metrics such as HiFi yield, Q30 quality score, and read length to assess how well the libraries were sequenced. The yield is the number of bases sequenced and ideally sits between 100 Gb and 120 Gb. We found this is generally the case in our data as our replicates gave a yield ranging from 98.33 Gb to 115.29 Gb. Our data shows the read length N50 as 15.9 kb, within the recommended 15-20 kb size range for best WGS results. For our replicates, we found that 93-95% of bases were above Q30, in line with PacBio's standard success criteria of ≥90% Q30.

CONCLUSIONS

We found that automating WGS library prep for PacBio HiFi sequencing produced high quality libraries and sequencing results. The Opentrons protocols for these workflows can be completed in 8 hours with just 1 hour of hands-on time.

ACKNOWLEDGEMENTS

The authors would like to thank Cristian Bravo, Alice Jung, and Elly Lee from Psomagen for valuable feedback on method development.