High-throughput human sample prep and sequencing on PacBio Revio system

Deborah Moine[1], Gregory Young[1], Jeffrey Burke[2], Julian Rocha[2], Michelle Kim[2], Primo Baybayan[1], Ian McLaughlin[1] and Kelvin J Liu[2]
1. PacBio, 1305 O'Brien Drive, Menlo Park, CA 94025. 2. 701 E. Pratt Street Baltimore, MD 21202 USA

Introduction

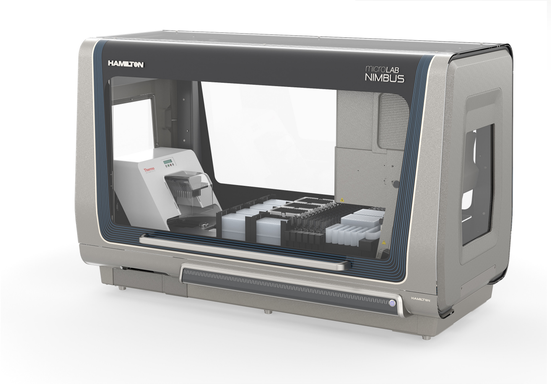
Improved throughput and cost of long-read sequencing, driven by technological advances of the PacBio Revio system, enables investigation of whole human genomes across larger populations.

To support the Revio system, high-throughput (HT) sample and library preparation solutions are needed.

Nanobind HT HMW-DNA extraction, 2h30 → SRE HT size selection, 2h → HT shearing + library prep, 6h

We present a fully automated HT DNA extraction, size-selection, shearing, and library preparation workflow for human whole blood and mammalian cell samples for PacBio HiFi sequencing.

HT high-molecular-weight (HMW) DNA extraction using Nanobind HT kits



HT HMW-DNA extraction is performed utilizing Nanobind magnetic disk technology with automated Hamilton *NIMBUS Presto* or Thermo Fisher *KingFisher Apex* systems.

Nanobind disks feature micro- and nano-structured silica wrinkles to shield bound DNA from damage during extraction.

Extraction result

96 plate format, *HG001* cell and human whole blood are extracted on *Nimbus Presto* and *KingFisher Apex* systems (2h30 run time)

Sample	input	DNA yield*	DNA mode size*
HG001 cell	1×10^6 cells	5.9 µg ±0.6	86 kb ±13
Human blood	200 µl	4.5 µg ±0.6	135 kb ±31

*Replicate of 8 extractions, average yield, and size

Automated HT size selection, pipette shearing, and library preparation

End-to-end automated workflow for 96 samples on Hamilton NGS STAR:

1) **HT size selection** >10 kb using short-read eliminator (SRE) kit on starting DNA (protocol under development)
2) **HT shearing**, HMW DNA shearing to 15–20 kb using robotic pipette shearing (protocol under development)
3) **HT SMRTbell library prep** using SMRTbell prep kit 3

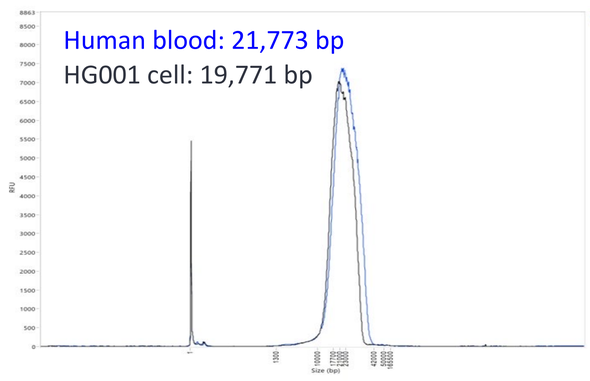


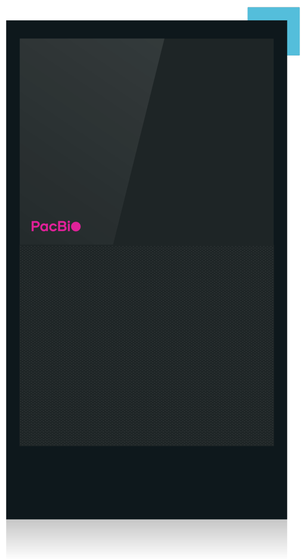
Figure 1. Final library size distribution on Femto *Pulse* system (Agilent Technologies)

Sequencing on PacBio Revio system

One blood and one *HG001* cell library were sequenced on Revio at 225 pM (one SMRT Cell per sample).

Sequencing data

For each sample 30-fold coverage (>90 Gb HiFi yield) per Revio SMRT Cell was obtained.



Sample	# HiFi reads	HiFi yield	Mean HiFi RL	Median QV
HG001 cell	6.0 M	**105 Gb**	17.4 kb	Q30
Human blood	5.1 M	**92 Gb**	18.0 kb	Q30

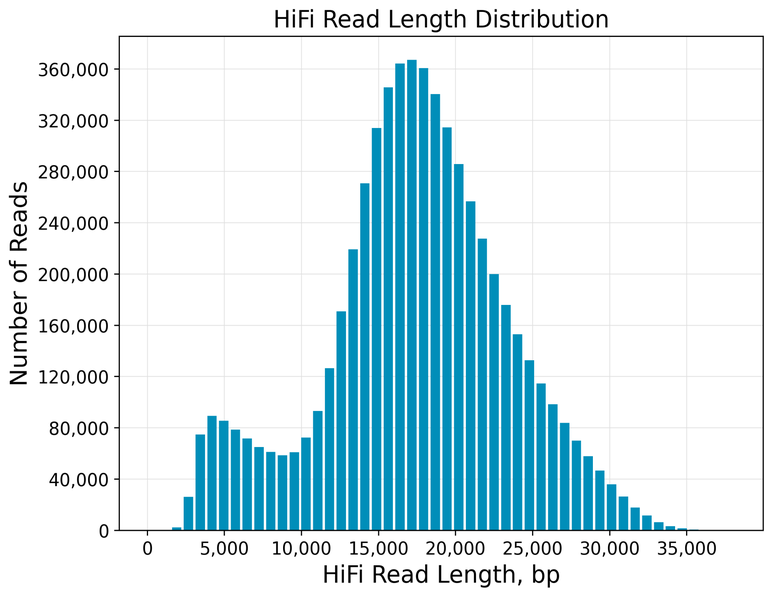


Figure 2. HiFi read length distribution plot for *HG001* cell sample

All analysis done from one cell for *HG001* sample:

De novo assembly using *hifiasm* with default parameter and consensus accuracy assessed with *yak.*

	Total size	N50 Contig length	Consensus accuracy
Hap1	3.01 Gb	44.2 Mb	QV 56.8
Hap2	3.02 Gb	49.7 Mb	QV 56.9

HiFi read mapping, methylation detection, and phased variants

Figure 2. ZIM2 and MIMT1 region (17 kb window, chr19: 56,830,786-56,848,469) reads phased by haplotypes and bases colored by 5mC status in IGV_2.16.1. Red indicates high, while blue indicates a low probability of methylation **(A)**. Same region with bases colored by variant **(B)**.

Structural variants

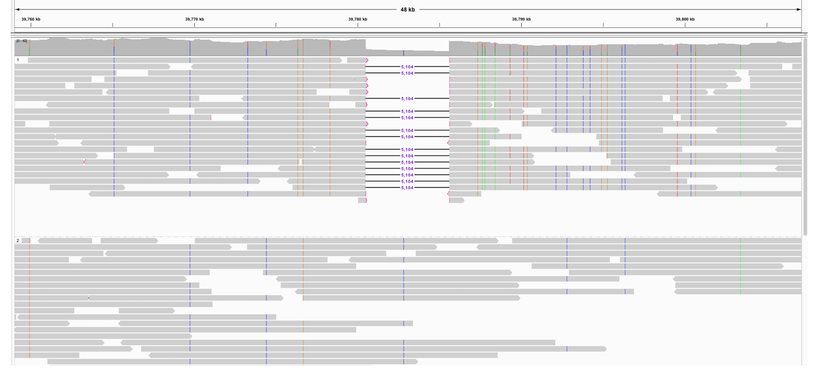


Figure 3. Phased HiFi reads spanning a large SV (5,104 bp deletion) at chr7:39,759,023-39,807,095 as shown in IGV_2.16.1.

Conclusion

We demonstrated a high-throughput, automated workflow for processing human blood samples and mammalian cell from extraction using Nanobind HT kits through HiFi sequencing on the Revio system. All sequenced samples generated ~30× coverage of HiFi data per SMRT Cell, sufficient for analysis including *de novo* assembly, phasing, methylation detection, and variant calling.

Acknowledgements

The authors would like to thank Jackson Mingle for DNA extraction, Aurelie Souppe and Heather Ferrao for sequencing, and Gregory Concepcion for data analysis.

Conflict of interests: authors are employees of PacBio

Research use only. Not for use in diagnostic procedures. © 2023 Pacific Biosciences of California, Inc. ("PacBio"). All rights reserved. Information in this document is subject to change without notice. PacBio assumes no responsibility for any errors or omissions in this document. Certain notices, terms, conditions and/or use restrictions may pertain to your use of PacBio products and/or third-party products. Refer to the applicable PacBio terms and conditions of sale and to the applicable license terms at pacb.com/license. Pacific Biosciences, the PacBio logo, PacBio, Circulomics, Omniome, SMRT, SMRTbell, Iso-Seq, Sequel, Nanobind, SBB, Revio, Onso, Apton, and Kinnex are trademarks of PacBio.