High throughput multiomic analysis for human genomics on PacBio Revio system

Julian Rocha[1], Jeffrey Burke[1], Renee Fedak[1], Duncan Kilburn[1], Deborah Moine[1], Enrique Bayo Iglesias[2], Dominik Laubscher[2], Birgit Ottenwälder[2], Suzanne Dee[1], Heather Ferrao[1], Kelvin J Liu[1]
1. PacBio, MP, USA; 2. Hamilton, Bonaduz, Switzerland

P16.075.C

Introduction

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

To support the growing capabilities of long-read sequencing, high throughput (HT) sample and library preparation solutions are necessary.

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

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

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

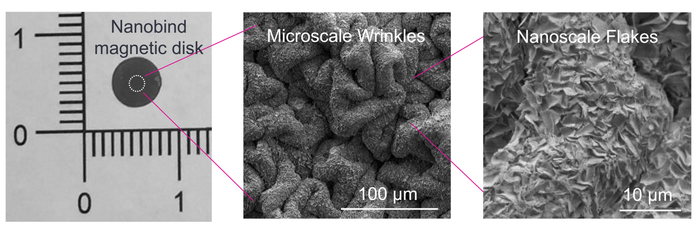


Figure 1. Scanning electron microscopic images of the surface of the Nanobind disk. The magnetic disk is coated with a high density of micro and nanostructured wrinkles coated with nanoscale silica flakes

HMW-DNA extraction is performed utilizing Nanobind magnetic disk technology on the fully automated Hamilton NIMBUS Presto. Alternatively, Thermo Fisher KingFisher instruments provide a semi-automated option with comparable metrics

Extraction result

- 3 to 12 µg of DNA on 200 µL blood sample on 96 well plate in 2.5 hours
- 3 to 70 µg of DNA per 1 mL blood sample on 24 well plate in 2.5 hours

DNA yield is linked to donor White Blood Cell (WBC) count.

Sample	Donor WBC count [10^9/L]	blood volume [mL]	DNA yield [ug]	DNA mode size [bp]
donor 1	6.2	1	24.7	81,501
donor 2	9.7	1	46.0	94,592
donor 3	6.0	1	38.8	85,089
donor 4	6.1	1	28.4	114,330

DNA QC metrics on four different donors showing high yield and HMW DNA. The four samples were used for HT library preparation and HiFi sequencing on Revio system

Automated HT pipette shearing and library preparation

HMW DNA is sheared to 15-20 kb using robotic pipette shearing.

- Protocols under development for Hamilton NGS STAR and Hamilton NIMBUS Presto
- Process 96 samples in 15 min

Automated PacBio library using SPK3 kit on Hamilton NGS STAR in 5 hours for 24 samples (6 hours for 96) automation script are available from Hamilton.

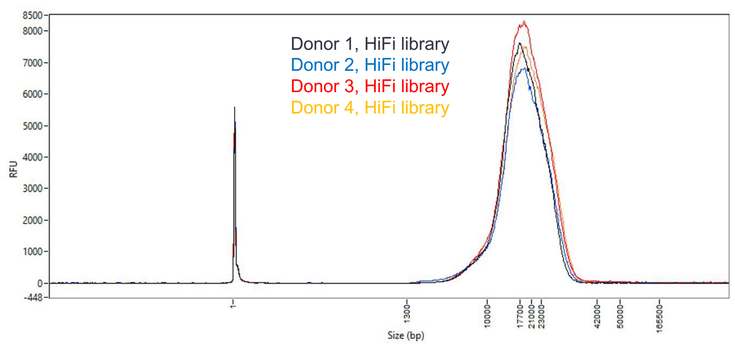
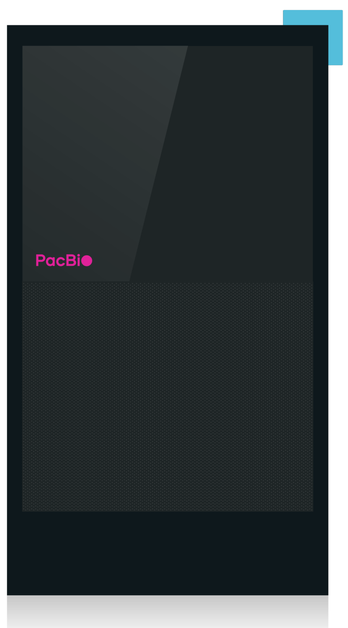


Figure 2. Final library size distribution for four donors

Sequencing on PacBio Revio system

Each library was sequenced on one Revio SMRT Cell (25M) at 225 pM during one sequencing run (24h movie, 4 cells in parallel).



Sequencing result

30-fold coverage was obtained for each donor using one cell per sample.

Sample	# HiFi reads	HiFi yield [Gb]	Mean HiFi RL [bp]	Median QV
donor 1	7.12 M	**106**	14,865	Q31
donor 2	6.35 M	**93**	14,604	Q32
donor 3	5.97 M	**90**	15,059	Q31
donor 4	6.75 M	**102**	15,132	Q32

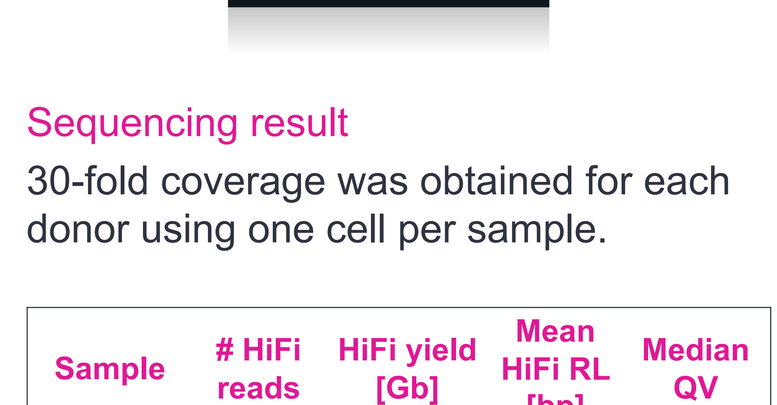


Figure 2. HiFi Read length distribution plot for the four SMRT cell corresponding to donor 1 to 4

All analysis done from one cell for donor 1:

De novo assembly

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

	Total size	N50 Contig length	Consensus accuracy
Hap1	2.95 Gb	43.3 Mb	QV 54.6
Hap2	2.99 Gb	33.8 Mb	QV 54.8

HiFi read mapping, methylation detection and phased variants

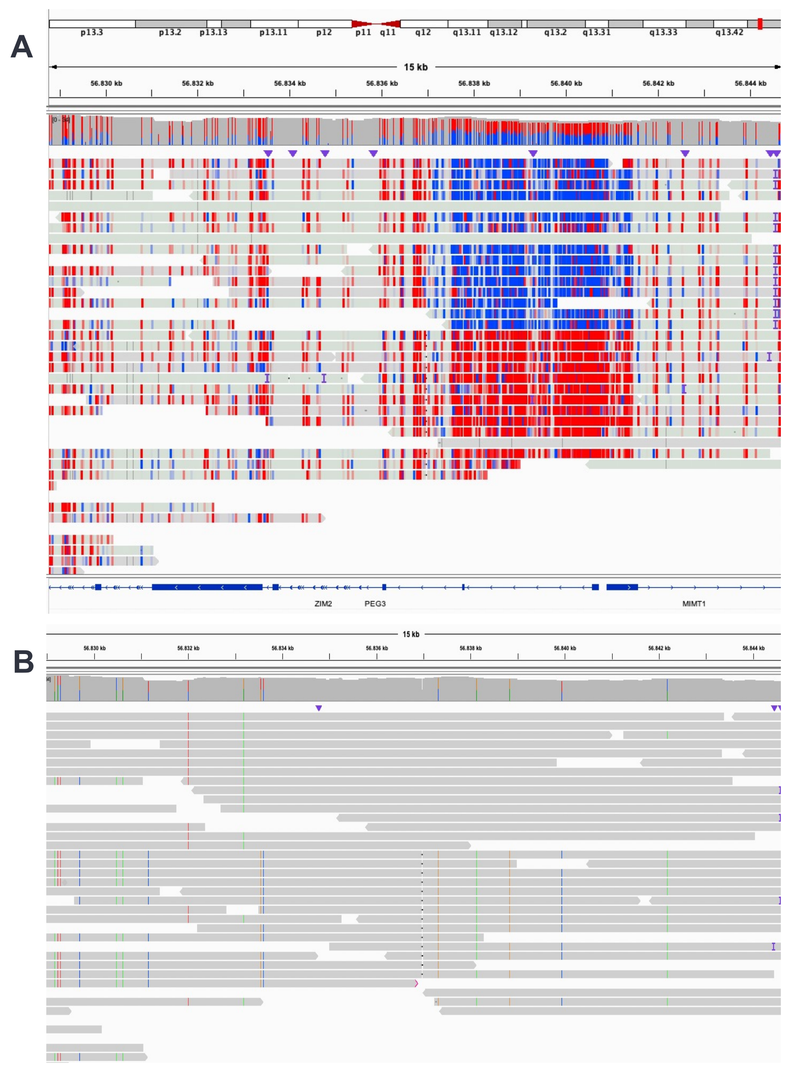


Figure 3. Parental imprinting at PEG3. Donor 1 reads phased by mat and pat haplotypes and bases colored by 5mC status in IGV_2.16.1. Red indicates high, while blue indicates a low probability of methylation **(A)**. Same region with based colored by variant **(B)**

Structural variants

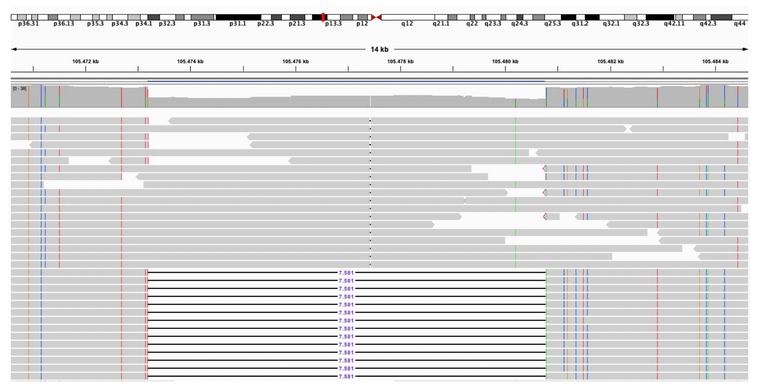


Figure 4. Donor1, Phased HiFi reads spanning a large SV (7581 bp deletion) at chr1:105473192-105480773 as shown in IGV_2.16.1.

Conclusion

We demonstrated a high-throughput, automated workflow for processing human blood samples from extraction using Nanobind HT kits through HiFi sequencing on Revio system.

Four blood samples from different donors were sequenced and each cell generated ~30X coverage of HiFi data, sufficient for analysis including *de novo* assembly, phasing, methylation detection, and variant calling.

Acknowledgements

The authors would like to thank Primo Baybayan and Ian McLaughlin who helped generate data and David Stucki for analysis

Conflict of interests: JR, JB, RF, DK, DM, SD, HF and KL are employees of PacBio. EBI, DL and BO are employees of Hamilton