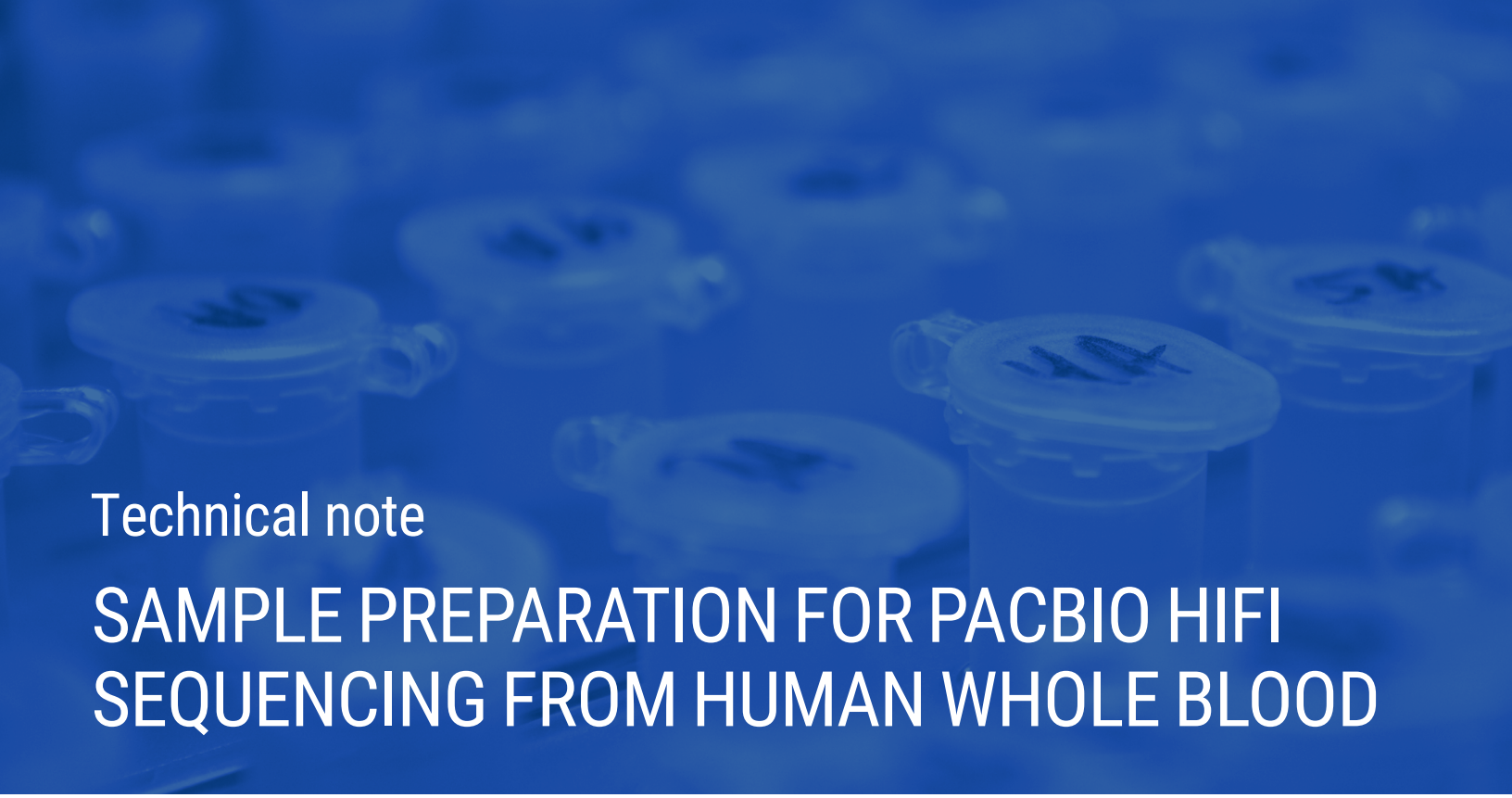

Technical note

SAMPLE PREPARATION FOR PACBIO HIFI SEQUENCING FROM HUMAN WHOLE BLOOD

Introduction

Whole blood is a common and easily accessible source of DNA that — with proper handling — provides high-quality input for PacBio HiFi sequencing. To define the best practices for handling human whole blood samples, we tested the effect of anticoagulant, sample storage time, storage conditions, and white blood cell count on the sequencing performance of DNA extracted using the Nanobind® CBB big DNA kit[1].

Summary

Stage	Variable	Best practice for PacBio HiFi sequencing
Before DNA extraction	Sample type	Human whole blood
	Anticoagulant	Potassium EDTA (K_2EDTA)
	Sample storage temperature	4 ± 3°C
	Sample storage time	≤ 2 days from collection to extraction
DNA extraction	Volume of whole blood	200 µL
	White blood cell (WBC) count	≥ 4 × 10^6 cells/mL for ≥ 3 µg of DNA
	DNA extraction method	Nanobind CBB big DNA kit
After DNA extraction	DNA storage	Rest 1 day at ambient temperature, then store at 4 ± 3°C
	DNA size distribution	• 90% of DNA ≥ 10 kb (genomic quality number at 10 kb ≥ 9.0) • 50% of DNA ≥ 30 kb (genomic quality number at 30 kb ≥ 5.0)
	UV absorbance	• A260/280 nm ≥ 1.7 • A260/230 nm ≥ 1.5

PacBio

Evaluation of sample preparation conditions

Human whole blood draw with either K_2EDTA, NaHep, or NaCit anticoagulant

↓

Sample storage at 4°C or ambient (21°C)

↓

DNA extracted with Nanobind CBB big DNA kit using 200 µL of whole blood[2]

↓

DNA quality control (Qubit and Femto Pulse systems)

↓

DNA shearing on Megaruptor 3 system and library prep with SMRTbell® prep kit 3.0[3]

↓

Sequencing on the Sequel® IIe system

Anticoagulant

HiFi sequencing performance was best for samples stored with potassium EDTA (K_2EDTA) as an anticoagulant. Samples stored in sodium heparin (NaHep) and citrate (NaCit) also performed well in very limited testing.

Anticoagulant	HiFi yield	HiFi read length	
		Average	N50
K_2EDTA ✓	31.0 Gb	14.2 kb	16.6 kb
NaHep	22.5 Gb	13.5 kb	16.9 kb
NaCit	30.3 Gb	12.0 kb	14.5 kb

Table 1. Average sequencing performance for samples stored with different anticoagulants.

Sample storage temperature and time

The amount of DNA extracted and the HiFi sequencing performance (yield, read length) declined as the days from the blood draw to DNA extraction increased. Storing samples at 4°C resulted in higher DNA recovery and better sequencing performance compared to ambient temperature (21°C).

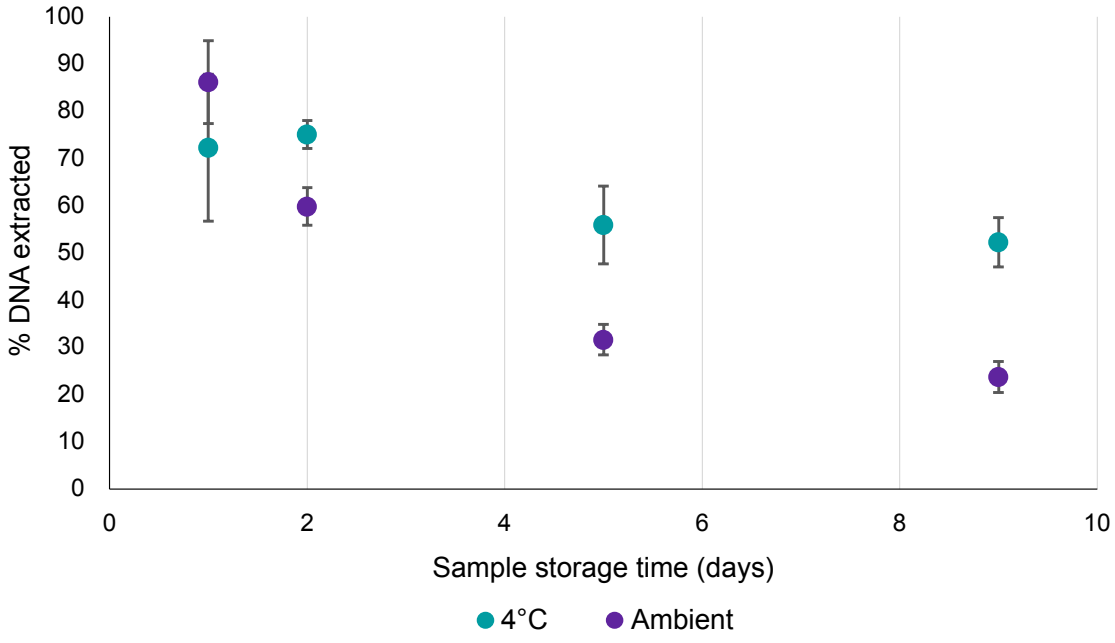

Figure 1. **Percent of DNA extracted across sample storage temperatures and times, normalized to an expectation of 6 pg of DNA per white blood cell (WBC).** Day 1, 4°C had 9 replicates; all other conditions had 3 replicates. Bars show standard deviation. Yield was measured on days 1, 2, 5, and 9 after collection. All samples used K_2EDTA as an anticoagulant.

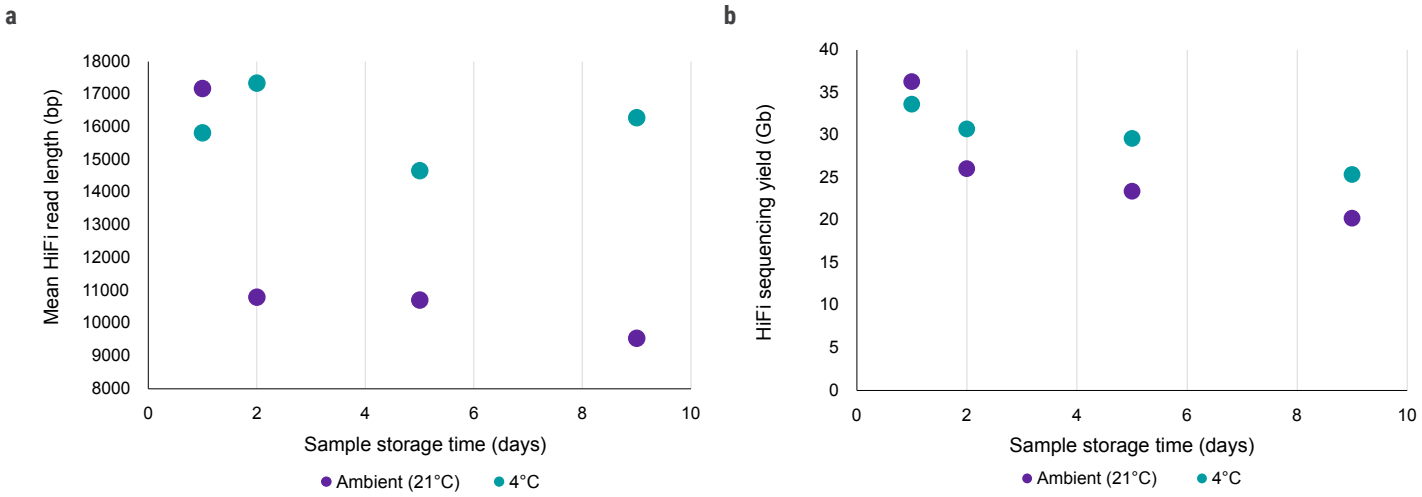

Figure 2. **HiFi sequencing performance across sample storage times and temperatures.** HiFi yield **(a)** and mean HiFi read length **(b)** for different storage conditions.

White blood cell (WBC) count

The amount of DNA extracted increased proportionally to WBC count. WBC counts above 4 × 10^6 cell/mL yielded more than 3,000 ng of DNA, which is enough for most human whole genome sequencing projects using SMRTbell prep kit 3.0 on the Sequel IIe system.

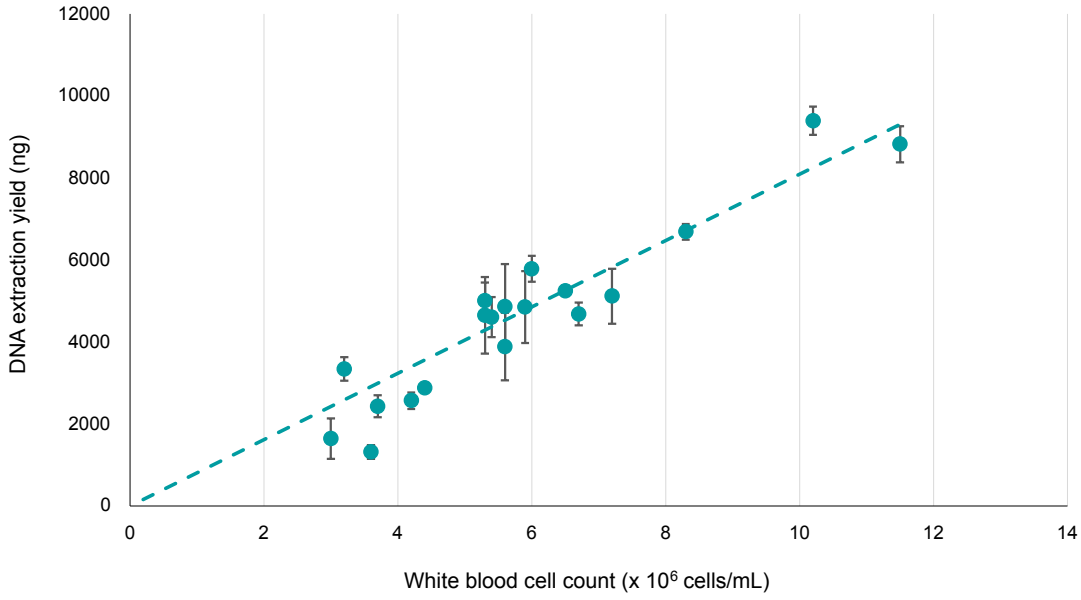

Figure 3. **DNA extraction yield from 200 µL of whole blood versus WBC count using the Nanobind CBB big DNA kit.** All samples shown were stored for 1 day at 4°C and used K_2EDTA anticoagulant. 3 replicates per sample with bars showing the standard deviation.

UV absorbance

No correlation was observed between UV absorbance ratios at 260/280 or 260/230 nm and sequencing performance when using the Nanobind CBB big DNA kit. All samples tested had A260/280 nm values above 1.7 and A260/230 nm above 1.5.

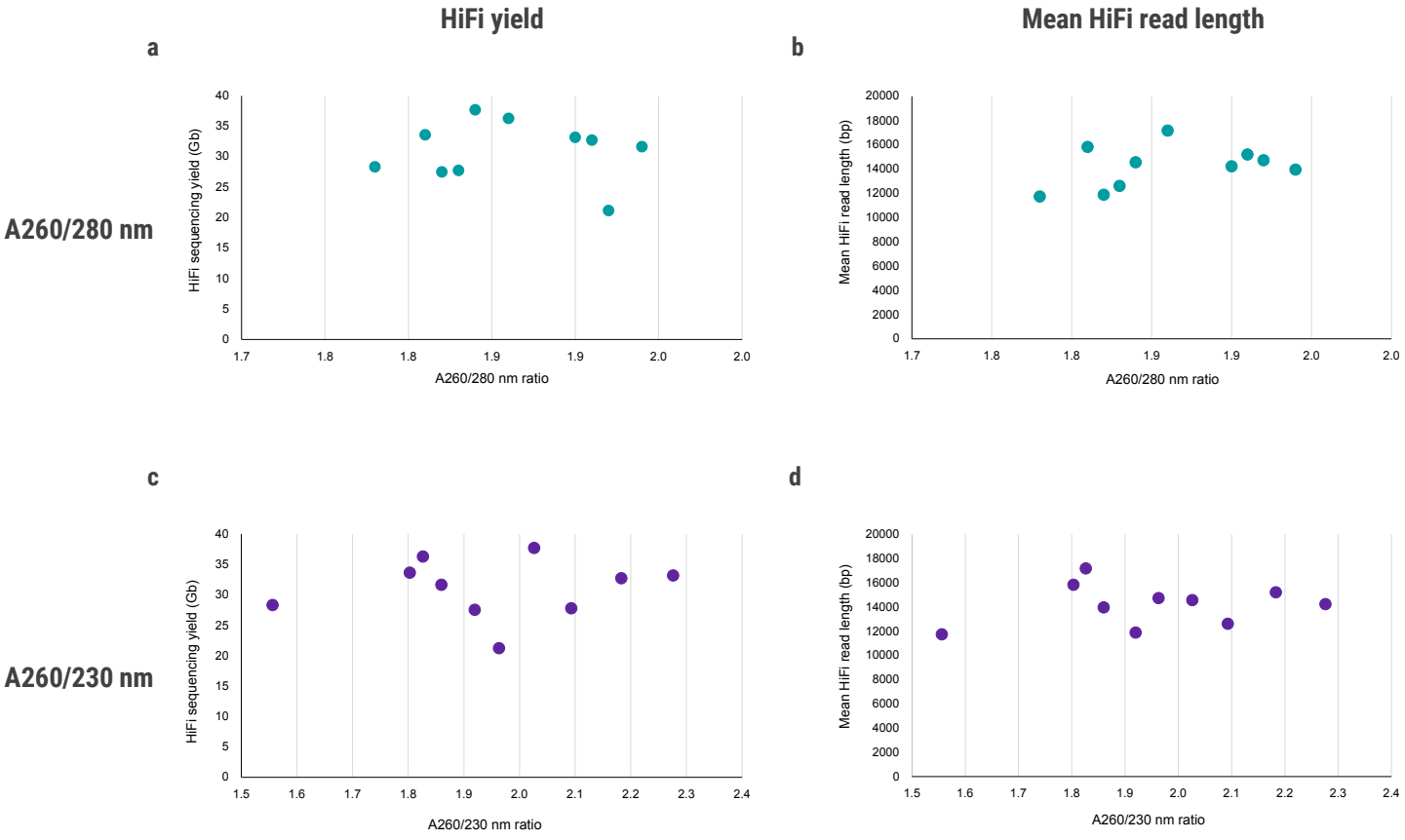

Figure 4. **UV absorbance vs HiFi sequencing yield and read length.** HiFi sequencing yield **(a,c)** and mean read length **(b,d)** by A260/280 nm **(a,b)** and A260/230 nm **(c,d)** UV absorbance ratios. Samples were stored for 1 day using K_2EDTA.

Typical results when following best practices

The human whole blood sample *Circ1207_12* illustrates typical results when following the sample extraction best practices described in this document. The sample was stored for one day after extraction at 4°C with K_2EDTA as anticoagulant. DNA was extracted using the Nanobind CBB big DNA kit and following the DNA extraction protocol for mammalian whole blood, yielding 9.0 µg with 87% of DNA longer than 30 kb and 96% longer than 10 kb. DNA was sheared using Megaruptor 3 system, a SMRTbell library was prepared using SMRTbell prep kit 3.0 (PacBio 102-182-700), and the library was sequenced on the Sequel II system. HiFi yield was more than 37 Gb with an average read length of 14.5 kb and median quality score of Q34.

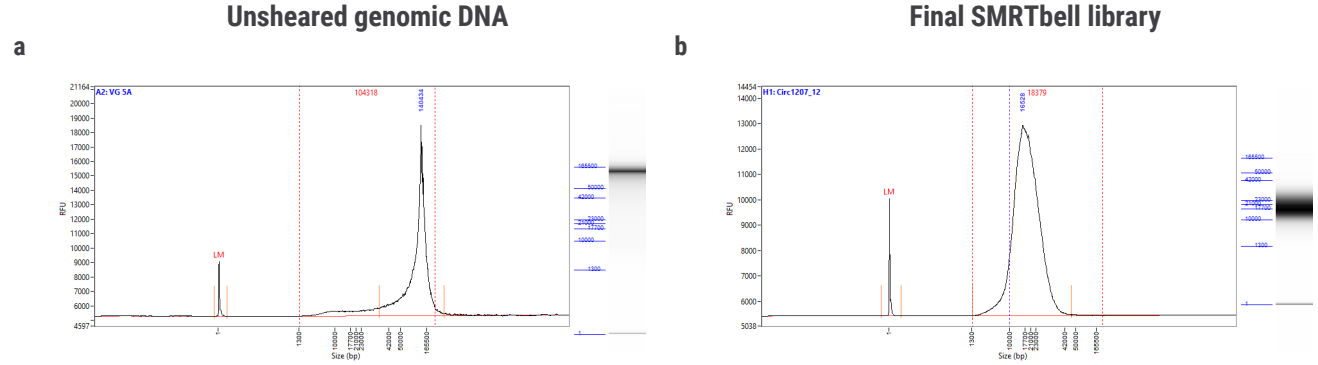

Figure 5. **DNA size distribution of the unsheared gDNA (a) and the final SMRTbell library (b) for sample Circ1207_12 on the Femto Pulse system (Agilent Technologies).**The modal length for the library is 16.5 kb. The average is 18.4 kb.

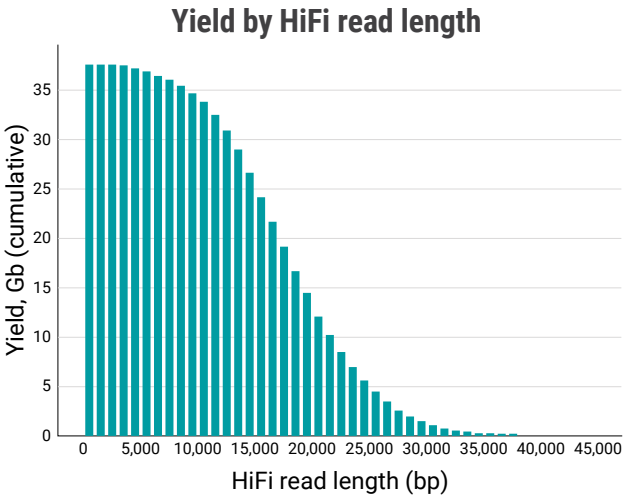

Figure 6. Cumulative yield (Gb) for all HiFi reads at a given read length for sample Circ1207_12. The total yield was 37.7 Gb with over 24 Gb (64%) in reads longer than 15 kb.

Metric	Value
HiFi reads	2,590,469
HiFi sequencing yield	37.7 Gb
Mean HiFi read length	14,551 bp
Median HiFi read quality	Q34

Table 2. HiFi sequencing metrics for Circ1207_12.

Discussion

Sample preparation is a critical factor for whole genome sequencing (WGS) projects that impacts sequencing yield, read length, and ultimately performance for variant calling and genome assembly. Best practices to optimize performance for PacBio HiFi sequencing are to collect whole human blood, store with potassium EDTA (K_2EDTA) as the anticoagulant for fewer than 2 days at 4°C, and extract DNA from 200 µL of blood using the Nanobind CBB Big DNA kit. When it is not possible to extract from fresh samples, storing samples at 4°C maximizes extraction yield and minimizes degradation for at least 9 days (figures 1 and 2).

For anticoagulants, K_2EDTA was extensively tested and showed good and consistent sequencing performance. Very limited testing was performed with blood samples stored in sodium heparin (NaHep) and citrate (NaCit) and therefore it is not clear how consistent their performance would be. We recommend using K_2EDTA when collecting blood samples for successful HiFi sequencing (table 1).

Samples with white blood cell (WBC) counts higher than 4×10^6 cell/mL typically yield more than 3 µg of genomic DNA, which is sufficient for most human WGS projects. Samples with lower WBC counts might require higher blood volumes or multiple extractions (figure 3).

No correlation was observed between HiFi sequencing performance and UV absorbance ratios at 260/280 or 260/230 nm (figure 4), which suggests that it is not necessary to use UV spectrophotometry as a quality control checkpoint when following best practices. In addition, good UV absorbance values are not a guarantee of good sequencing performance because not all inhibitors absorb at the wavelengths of 230, 260, and 280 nm. Nevertheless, all tested samples had A260/280 greater than 1.7 and A260/230 greater than 1.5. For samples with absorbance ratios outside this range, it is recommended to do a 1× AMPure® PB bead cleanup to remove potential contaminants.

Following these best practices for sample preparation and the SMRTbell prep kit 3.0 protocol for library preparation is key to limiting variability, reducing costs, and maximizing yield and data quality from HiFi sequencing. In turn, high quality HiFi sequencing will provide comprehensive detection of genetic and epigenetic variation in your samples (figure 6).

KEY REFERENCES

1. **Nanobind CBB big DNA kit (NB-900-001-01)**
2. **Nanobind HMW DNA extraction – Mammalian whole blood (EXT-BLH-001)**
3. **Procedure + checklist — Preparing whole genome and metagenome libraries using SMRTbell® prep kit 3.0 (102-166-600)**