P21.055.A

Saliva, a convenient sample type enabling large scale PacBio HiFi sequencing project

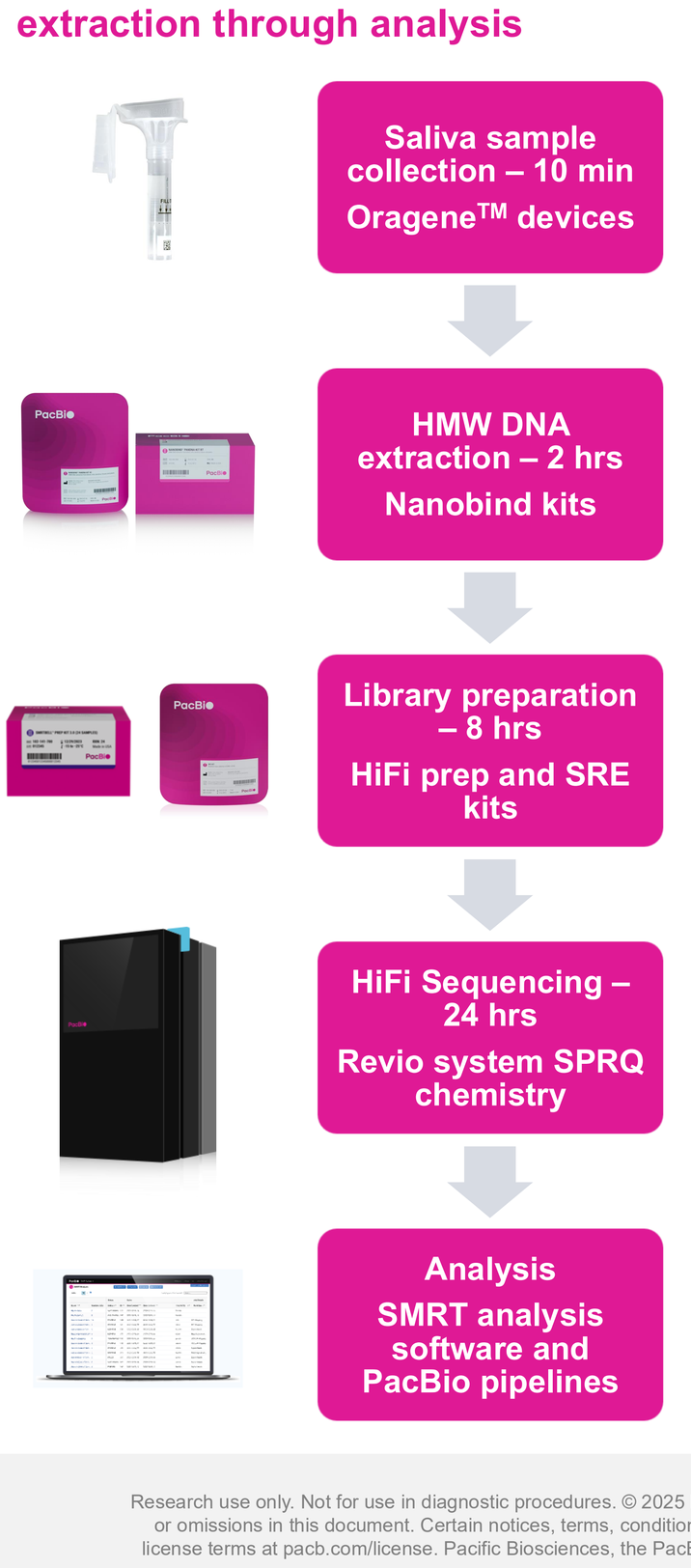
Deborah Moine, Sarah Kingan, Shreyasee Chakraborty, Heather Ferrao, Kristina Weber, Christina Dillane*, Mike Tayeb*, Duncan Kilburn
*DNA Genotek, Inc. Ottawa, ON, Canada

Background

Material and Methods

High Molecular Weight (HMW) DNA was extracted from saliva and blood manually or with automation (KingFisher Apex system) using Nanobind kits. HiFi WGS libraries were prepared on the Hamilton NGS STAR system and sequenced using SPRQ™ chemistry on the Revio system. HiFi reads were mapped against human genome reference (GRCh38). Variant calling and methylation analysis were performed using SMRT® Link analysis pipelines.

Simple workflow for saliva DNA extraction through analysis



Result

1. DNA extraction

High-quality HMW DNA was obtained from all samples using Nanobind extraction kits. Modal DNA size as measured by Femto Pulse is >80 kb for majority of samples.

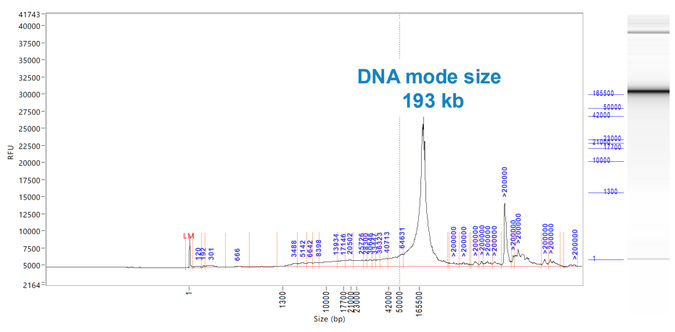


Figure 1. Size distribution of DNA extracted from saliva donor 1 on Femto Pulse System (Agilent technologies).

High throughput (HT) extraction from saliva and blood gave similar yield than manual extraction.

Nanobind workflow	Sample volume	DNA Yield range
Manual saliva	500 µL	0.5-45 µg
HT saliva	500 µL	0.5-45 µg
Manual blood	200 µL	3-10 µg
HT blood	200 µL	3-10 µg

Table 1. Yield for saliva and whole blood DNA extraction. Yield will vary depending on donor white blood cell concentration.

2. Sequencing

Similar HiFi sequencing metrics from saliva versus blood. Paired saliva and blood samples were sequenced on a single SMRT® Cell on the Revio® system. We obtained 98 to 135 Gb of HiFi data resulting in 30- to 39-fold coverage per genome sufficient for comprehensive WGS variant detection and methylation analysis. DNA extracted from saliva is typically ~75−95% human as bacteria can be present in the sample.

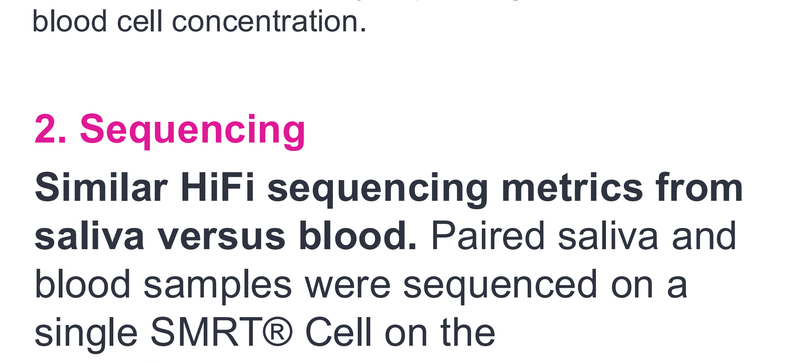


Figure 2. HiFi Read length distribution plot for saliva and blood from donor 6.

Figure 3. Paired saliva and blood samples from same donor show similar metrics for HiFi yield **(A)**, giving 30- to 39-fold human genome coverage **(B)**. 5mC Methylation profile score is on average 97.7% (saliva) and 98.4% (blood) with 74.4% (saliva) and 62.7% (blood) modified sites (Pr > 0.5) **(C)**. This difference can be explained by presence of nonhuman DNA in saliva samples (1.8-7.3%).

3. Analysis

Variant calling with PacBio WGS Variant Pipeline. Paired saliva and blood samples were assessed for concordance in genome coverage quality and variant detection.

Each sample was analyzed as a singleton using v3-a1. Code available at:
github.com/PacificBiosciences/HiFi-human-WGS-WDL/

Variant type	Genotyping	Concordance
Single nucleotide variants (SNVs)	Deep Variant v1.8.0	RTG Tools v3.12.1
Structural variants (SVs)	Sawfish v0.12.7	truvari v5.2.0

Table 2. Software tools and version used for genotyping and concordance analysis.

High variant calling concordance.

Paired **saliva** and **blood** samples for the same donor were assessed for genotype concordance of high-quality (GQ≥20) variants across GRCh38 without masking.

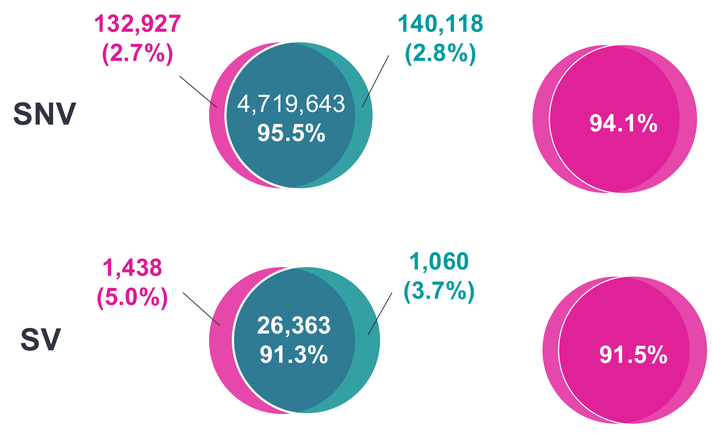


Figure 3. Blood and saliva variant concordance results for donor 6 (left). For comparison, a technical replicate of blood from the same donor is shown on the right (blood/blood). Comparison of blood samples from two different donors had concordance of 21.7% for SNVs and 44.6% for SVs (not shown).

Conclusion

- **Saliva samples collected with Oragene™ devices and DNA extracted using Nanobind kits are a good alternative to blood for HiFi sequencing.**
- Similar sequencing performance and variant calling concordance were obtained from blood and saliva for the same individual.
- High-throughput workflow from extraction using Nanobind HT kit through HiFi sequencing on the Revio system is available for saliva and blood samples.

Acknowledgements

The authors would like to thank Dan Portik and Billy Rowell for data analysis.

Disclaimer: Oragene™ is a trademark of DNA Genotek Inc.