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

PacBie

It would be great if I could use saliva samples for HiFi projects because collecting and storing are much easier than blood!

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.

41743-		1	
40000-			
37500-			
35000-			
32500-	DNA mode size		
30000-	193 kb	_165500	T
27500-		<u> </u>	
25000-			

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/

Is it possible to extract High Molecular Weight DNA from saliva samples?

Yes! With saliva collected in Oragene[™] devices and DNA extracted with Nanobind kits, we obtain HMW DNA optimal for HiFi and comprehensive variant analysis!



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.

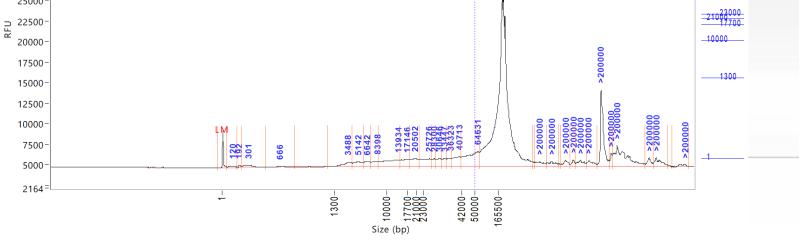


Figure 1. Size distribution of DNA extracted from saliva donor 1 on Femto Pulse System (Agilent technologies).

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 DNAextraction. Yield will vary depending on donor whiteblood cell concentration.

2. Sequencing

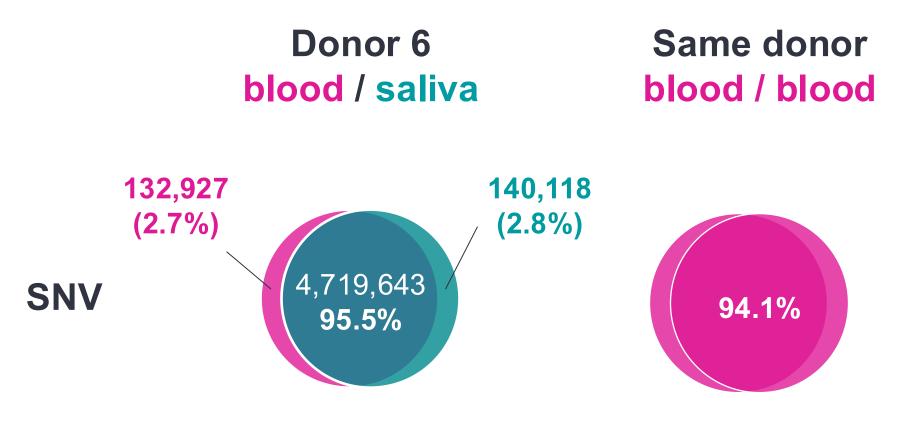
Similar HiFi sequencing metrics from saliva versus blood. Paired saliva and blood samples were sequenced on a single SMRT® Cell on the Revio® system. We obtained 98 to 135 Gb of HiFi data resulting in 30- to 39fold coverage per genome sufficient for comprehensive WGS variant detection and methylation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

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

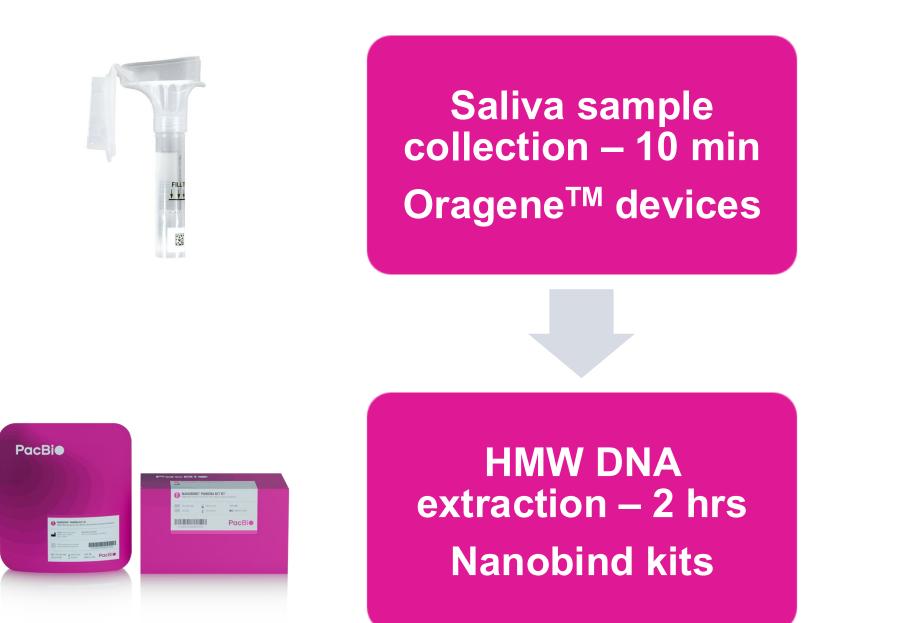
Table 2. Software tools and version used forgenotyping and concordance analysis.

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.



Simple workflow for saliva DNA extraction through analysis



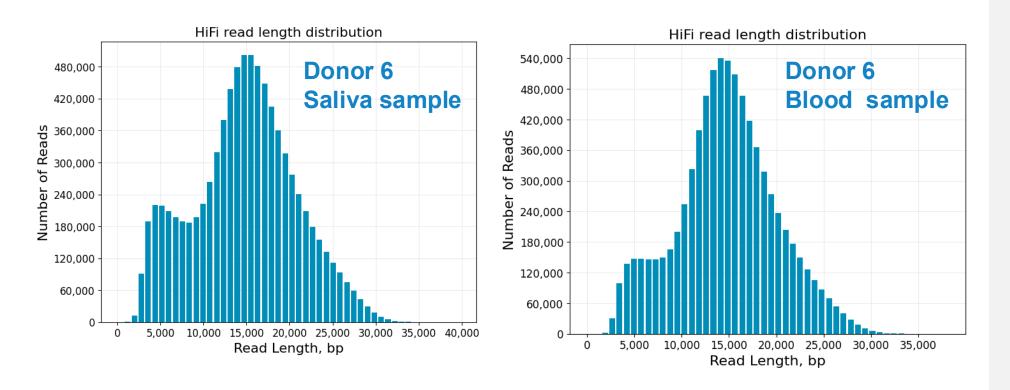
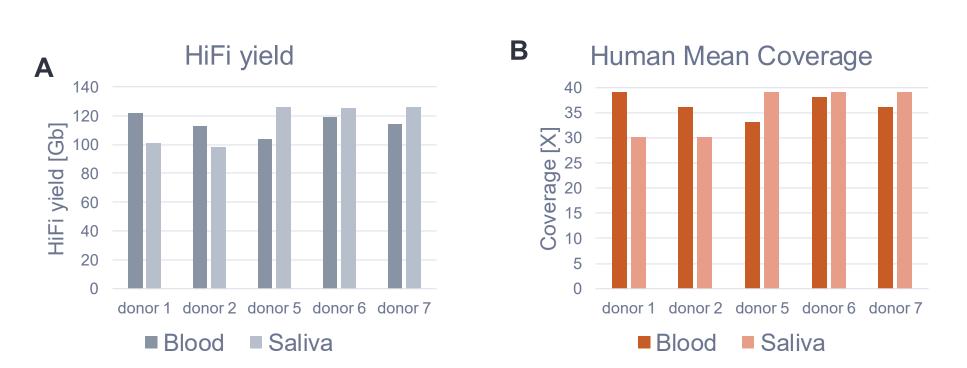


Figure 2. HiFi Read length distribution plot for saliva and blood from donor 6.



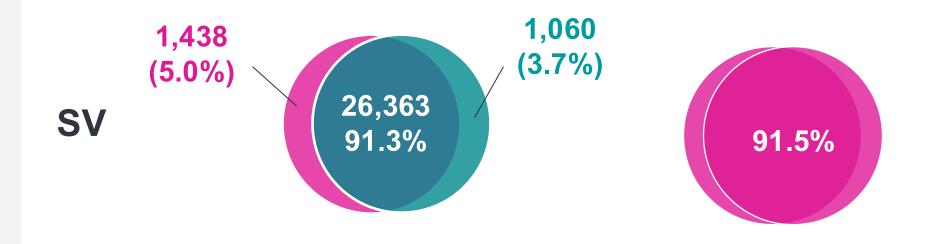


Figure 3. Blood and saliva variant concordance

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).



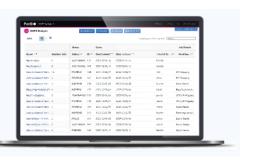
 Saliva samples collected with Oragene[™] devices and DNA extracted using Nanobind kits are a good alternative to blood for HiFi sequencing.

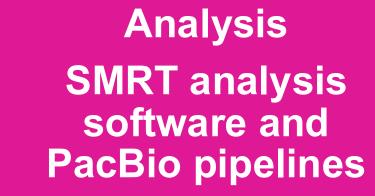


Library preparation – 8 hrs HiFi prep and SRE kits



<section-header><section-header><section-header>





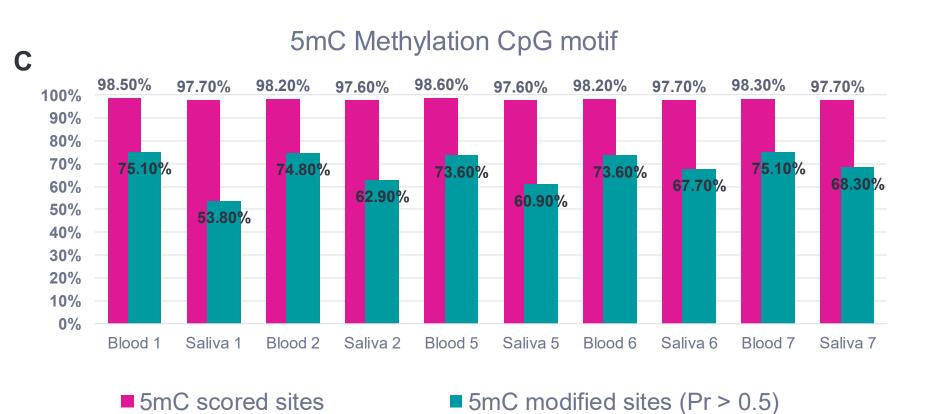


Figure 3. Paired saliva and blood samples from same donor show similar metrics for HiFi yield (**A**), giving 30- to 39-fold human genome coverage (**B**). 5mC Methylation profile score is on average 97.7% (saliva) and 98.4% (blood) with 74.4% (saliva) and 62.7% (blood) modified sites (Pr > 0.5) (**C**).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

- 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.

Research use only. Not for use in diagnostic procedures. © 2025 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, Kinnex, PureTarget, SPRQ, and Vega are trademarks of PacBio.