Analysis of concordance of genetic variants across sample types

Introduction

Recent developments in nucleic acid extraction procedures from diverse sample types have opened the possibility of diagnostic and research applications using various kinds of tissues and body fluids. Currently, several techniques and sample types are being used for nucleic acid extraction, including bronchoalveolar lavage (BAL) fluid, nasal swabs, brush biopsies, pharyngeal swabs, feces, blood, and saliva. Of these, saliva has emerged as a promising sample type for diagnosis of infectious diseases and genetic applications, including genotyping for personalized medicine and rare disease diagnosis. Thermo Fisher Scientific has been at the forefront of testing and has released highly sensitive kits, including the Applied Biosystems™ MagMAX™ Viral/Pathogen II Nucleic Acid Isolation Kit, which can be used for sample types in both saliva and nasopharyngeal viral transport medium (NPVTM) to extract viral nucleic acid from SARS-CoV-2 and a variety of other pathogens.

The MagMAX Viral/Pathogen kit has been a frontline solution for SARS-CoV-2 RNA extraction and research. The Applied Biosystems™ MagMAX™ DNA Multi-Sample Ultra 2.0 Kit is another product that can be used to extract nucleic acid containing genomic DNA (gDNA) from multiple sample types, including whole blood, saliva, buffy coat, and buccal swabs. The extracted samples can be used for various downstream purposes such as genotyping by sequencing and quantitative PCR (qPCR). The extraction can be automated by pairing the MagMAX DNA kit with the magnetic bead–based Thermo Scientific™ KingFisher™ Flex Purification System, allowing for a single streamlined and efficient workflow for multiple sample types. The KingFisher Flex system provides high-throughput sample extraction with consistent nucleic acid quality and quantity. Saliva has been established as one of the top sample type options both by the Clinical Laboratories Improvement Amendments (CLIA)-certified labs for diagnostic research and by consumers in the biotechnology industry and academic research.

Extracted gDNA can be used in many downstream applications such as genotyping to find out genetic disorders, ancestry prediction, and research on oral cancers and other diseases. Blood, buccal swabs, and saliva are the three main sample types used for extraction of gDNA to be used for next-generation sequencing (NGS) or qPCR. Blood has historically been the common sample choice for gDNA extraction due to the high quality and quantity of DNA yield [1].

Recently it has become clear that there is an increasing need to utilize noninvasive sample types such as buccal swabs or saliva. Moreover, saliva and buccal swab samples are more economical in their shipping and storage options than other media such as whole blood, serum, and plasma. Saliva samples also require less manipulation compared to blood, serum, plasma, and other sample types [2]. However, data for the concordance between different sample types used for diagnostic purposes, especially using the same extraction chemistry and workflows, is limited. To determine if common human genetic variants in a population can be detected consistently across sample types, DNA was extracted from 25 different donors with matched whole blood, buccal swab, and saliva samples. Nucleic acid was extracted and analyzed for quality and quantity of extracted DNA. Concordance among genetic variants across donors was observed using the Ion AmpliSeq™ Exome panel sequencing workflow and qPCR.

Materials and methods

Sample preparation and nucleic acid quality

Whole blood, buccal swabs, and saliva samples were collected from 25 male and female donors ranging in age from 21 to 67 years. Race and ethnicities were recorded across all randomized donors as white and black as well as Hispanic and non-Hispanic. DNA extractions from whole blood and saliva were performed using 400 µL sample inputs, and one buccal swab per donor was used for DNA extractions from buccal swabs. All DNA extractions were performed using the MagMAX DNA Multi-Sample Ultra 2.0 Kit and high-throughput DNA extraction protocols specific for each sample type. The quality and purity of the extracted DNA were analyzed using the Thermo Scientific™ NanoDrop™ 8000 Spectrophotometer. The purity of the extracted DNA was determined by absorbance ratios (A_{260}/A_{280} and A_{260}/A_{230}), along with the total nucleic acid yield. The total DNA yield of each sample was measured using the Invitrogen™ Qubit™ dsDNA HS Assay Kit on the Invitrogen™ Qubit™ 4 Fluorometer. Although both instruments can detect nucleic acids, the NanoDrop 8000 Spectrophotometer is unable to distinguish between DNA and RNA. In contrast, the Qubit dsDNA HS Assay Kit is specific to DNA and has high sensitivity, and therefore is more reliable at quantifying low DNA concentrations. However, the NanoDrop 8000 Spectrophotometer provides insight into the quality of the nucleic acid, so using both instruments in tandem is common practice.

Genotyping by sequencing and qPCR

All extractions were normalized to 5 ng/µL and sequencing libraries were prepared with the Ion AmpliSeq™ Exome RDY Kit on the Ion Chef™ Instrument. Prepared libraries were sequenced on the Ion GeneStudio™ S5 System. Variant calling concordance was analyzed using Ion Reporter™ Software 5.6 with the Ion AmpliSeq Exome RDY Kit's paired sample workflow.

To confirm the integrity of DNA across sample types, three different Applied Biosystems™ TaqMan® Assays were utilized for DNA and copy number quantification and concordance between matched sample types (ACTB_g1, GAPDH_g1, and RNaseP_Copy Number Reference) on the Applied Biosystems™ QuantStudio™ 12K Flex Real-Time PCR System.

Genotyping concordance was analyzed with qPCR for 16 of 25 donors across whole blood, buccal swab, and saliva sample types. The extracted nucleic acid from each sample was diluted, and 2.25 µL of the diluted sample was analyzed by qPCR, in duplicate, using Applied Biosystems™ TaqMan® Genotyping Master Mix on the QuantStudio 12K Flex Real-Time PCR System following the standard PCR cycling conditions. Genotyping was performed for three SNPs, the TaqMan Assay IDs of which are C_27102425_10 (SNP ID: rs16947), C_34816116_20 (SNP ID: rs28371725) and C_1202882_20 (SNP ID: rs1801133). rs16947 and rs28371725 were chosen because they are related to genes of drug metabolism and can be applied to various application needs. Applied Biosystems™ QuantStudio™ 12K Flex Real-Time PCR Software was used to call genotypes after 40 cycles.

Results and discussion

gDNA extraction workflow and sample quality

Genomic DNA was isolated from all three sample types—whole blood, buccal swabs, and saliva—using the MagMAX DNA Multi-Sample Ultra 2.0 Kit. Samples from 22 of the 25 donors produced enough high-quality DNA across all matched sample types to move forward with sequencing and evaluate variant concordance. Saliva and whole blood samples utilize the same high-throughput workflow and script when using the MagMAX DNA kit, and thus both types of samples could be extracted together on one plate in 50 minutes. Figure 1 details the total nucleic acid yield from each of the 25 donors across blood, buccal swab, and saliva extractions.

DNA yield was evaluated across all 25 donors with the Qubit dsDNA HS Assay Kit. The Qubit 4 Fluorometer with the Qubit dsDNA HS Assay Kit provides superior sensitivity for DNA concentration compared to the NanoDrop 8000 Spectrophotometer for sequencing applications. The average DNA yields across whole blood, buccal swabs, and saliva samples were found to be within one standard deviation of each other. Figure 2 shows the average DNA yield with matched blood, buccal swabs, and saliva samples across 25 donors.

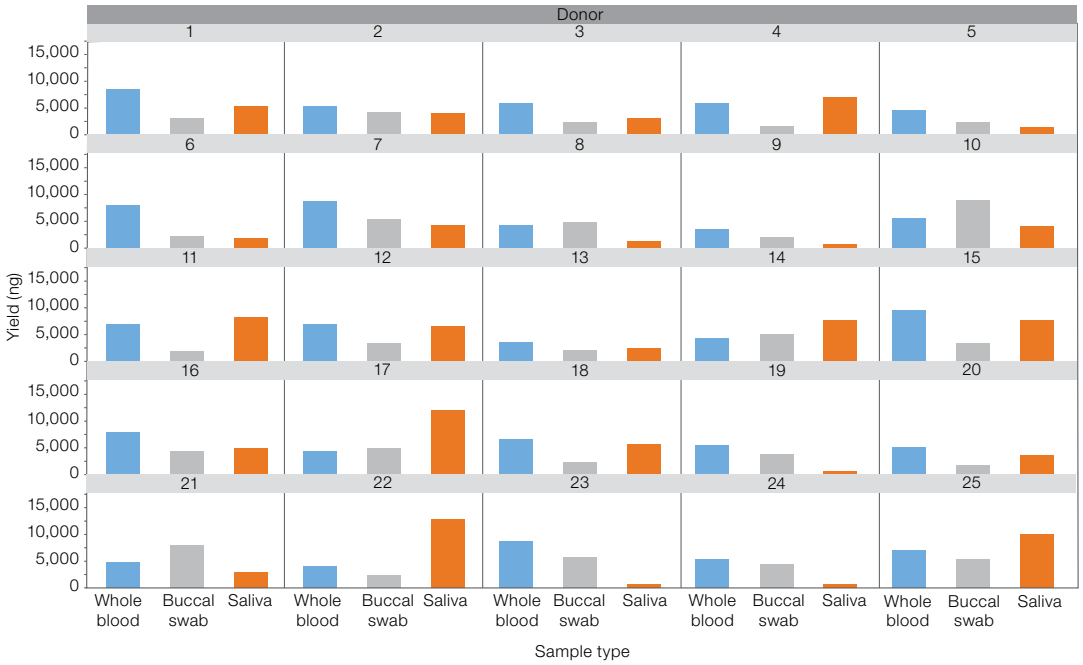

Figure 1. Total nucleic acid yields from whole blood, buccal swabs, and saliva samples across 25 donors, determined using the NanoDrop 8000 Spectrophotometer. The yield was evaluated across all 25 donors. The yields across whole blood, buccal swabs, and saliva samples were found to be within one standard deviation of each other.

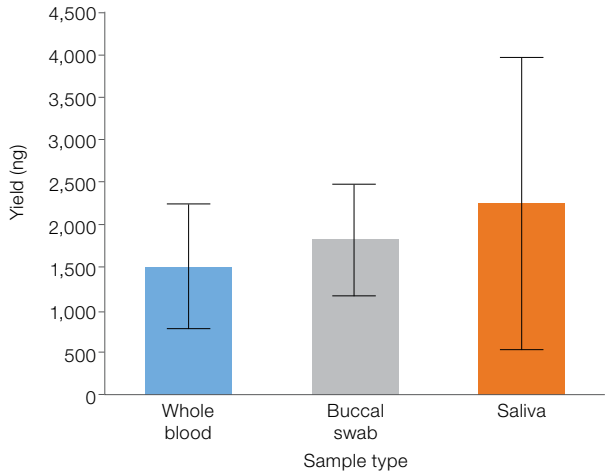

Figure 2. Average DNA yields from whole blood, buccal swabs, and saliva samples from 25 donors, determined using the Qubit dsDNA HS Assay Kit and the Qubit 4 Fluorometer.

Sample purity was evaluated across matched blood, buccal swabs, and saliva samples across all 25 donors by analyzing the absorbance at 260 nm, 280 nm, and 230 nm with the NanoDrop 8000 Spectrophotometer. The ratio of absorbances at 260 nm and 280 nm (A_{260}/A_{280}) is widely used to assess the purity of DNA, and the ratio at 260 nm and 230 nm (A_{260}/A_{230}) is a secondary measure of nucleic acid purity. The A_{260}/A_{280} averaged 2.0 for all sample types, indicating pure DNA across most samples (Figure 3B). One exception was a female donor's (donor 24) saliva sample that had a high A_{260}/A_{280} but was accepted to move forward with sequencing (Figure 3A). There was no variation in A_{260}/A_{280} observed between female and male donors (Figure 3A). Similarly, A_{260}/A_{230} was determined for all samples for nucleic acid purity; the average A_{260}/A_{230} for blood and buccal swabs were similar, while the A_{260}/A_{230} ratios for saliva samples were marginally lower (Figure 3A). Samples from 22 donors yielded enough for further sequencing studies.

qPCR and genotyping by qPCR

To confirm that the quality and quantity of the DNA from sample types is suitable for qPCR-based applications, qPCR was performed for two well-known housekeeping genes: ACTB and GAPDH. The Applied Biosystems™ TaqMan® Copy Number Reference RNase P Assay was used for quantification of gDNA before performing genotyping by qPCR. 16 of 24 donors had enough eluate (post-NGS) to perform both qPCR and genotyping by qPCR for all three genes, with an exception of one buccal swab sample from donor 5 (Figure 4B). All three samples from donor 5 were tested for GAPDH and RNase P copy number targets; however, there was not enough eluate to perform qPCR for ACTB target. As shown in Figure 4, the amplification of ACTB, GAPDH, and RNase P from all sample types was comparable across the three matched sample types for 16 donors. We used TaqMan Assays for genotyping of three SNPs, two of which are found within genes associated with drug metabolism.

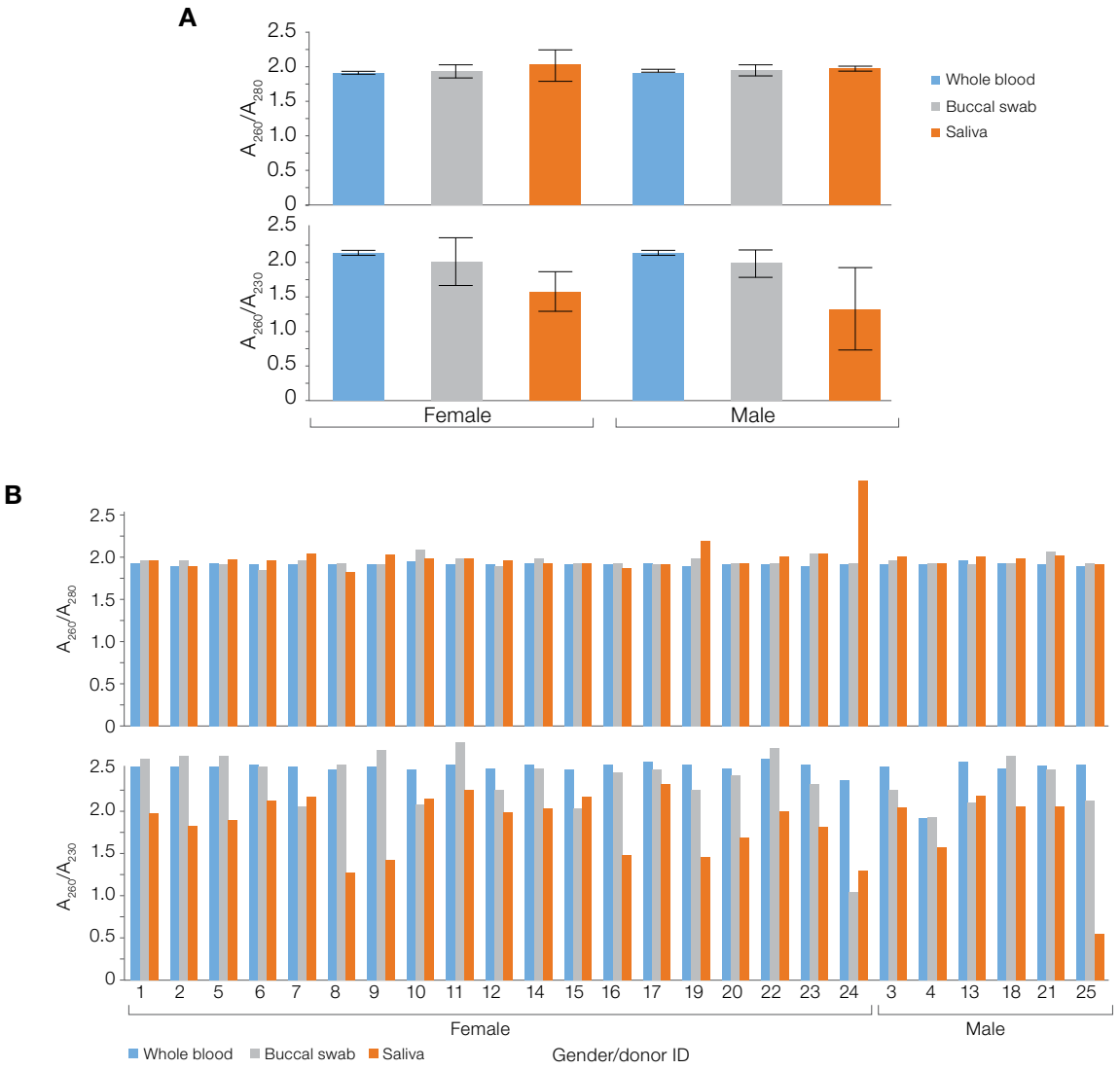

Figure 3. A_{260}/A_{280} and A_{260}/A_{230} values obtained with the NanoDrop 8000 Spectrophotometer across whole blood, buccal swabs, and saliva samples. (A) Average values by gender and **(B)** individual values for all samples.

Figure 5 shows consistent genotype calling across all three sample types. This outcome supports the result from genetic variant calling by sequencing shown in this paper, and suggests that all three sample types extracted with MagMAX DNA Multi-Sample Ultra 2.0 Kit can be used with sensitive molecular biology techniques.

Variant concordance with exome paired sample analysis

Sufficient amounts of DNA were extracted from 22 of the 25 donor samples to perform exome panel sequencing using the Ion AmpliSeq Exome RDY Kit with the Exome v1 panel for matched extracted blood, buccal swabs, and saliva samples. Matched eluates were analyzed for variant call concordance with next-generation sequencing (NGS) utilizing the Ion AmpliSeq™ workflow. The Ion AmpliSeq™ Library Kit Plus was used for the preparation of amplicon libraries. Libraries were prepared using the Ion Chef system and sequenced on the Ion GeneStudio S5 System with an Ion 550™ Chip. Figure 6 illustrates variant calling concordance >98% across all sequenced samples for shared loci, which indicates consistent calling functionality for all samples. On average, less than 2% of shared loci had one inconsistent variant between three matched donor samples, and less than 0.05% of shared loci had three inconsistent variants.

When observing variant calling across sample types, sequencing results showed buccal swabs and saliva with 94% matched variant calls across 22 donors (Figure 7). By comparing matched saliva to whole blood and whole blood to buccal swabs, >93% matched variant calls were observed. There is high concordance between matched sample types when extracted with the MagMAX DNA 2.0 kit.

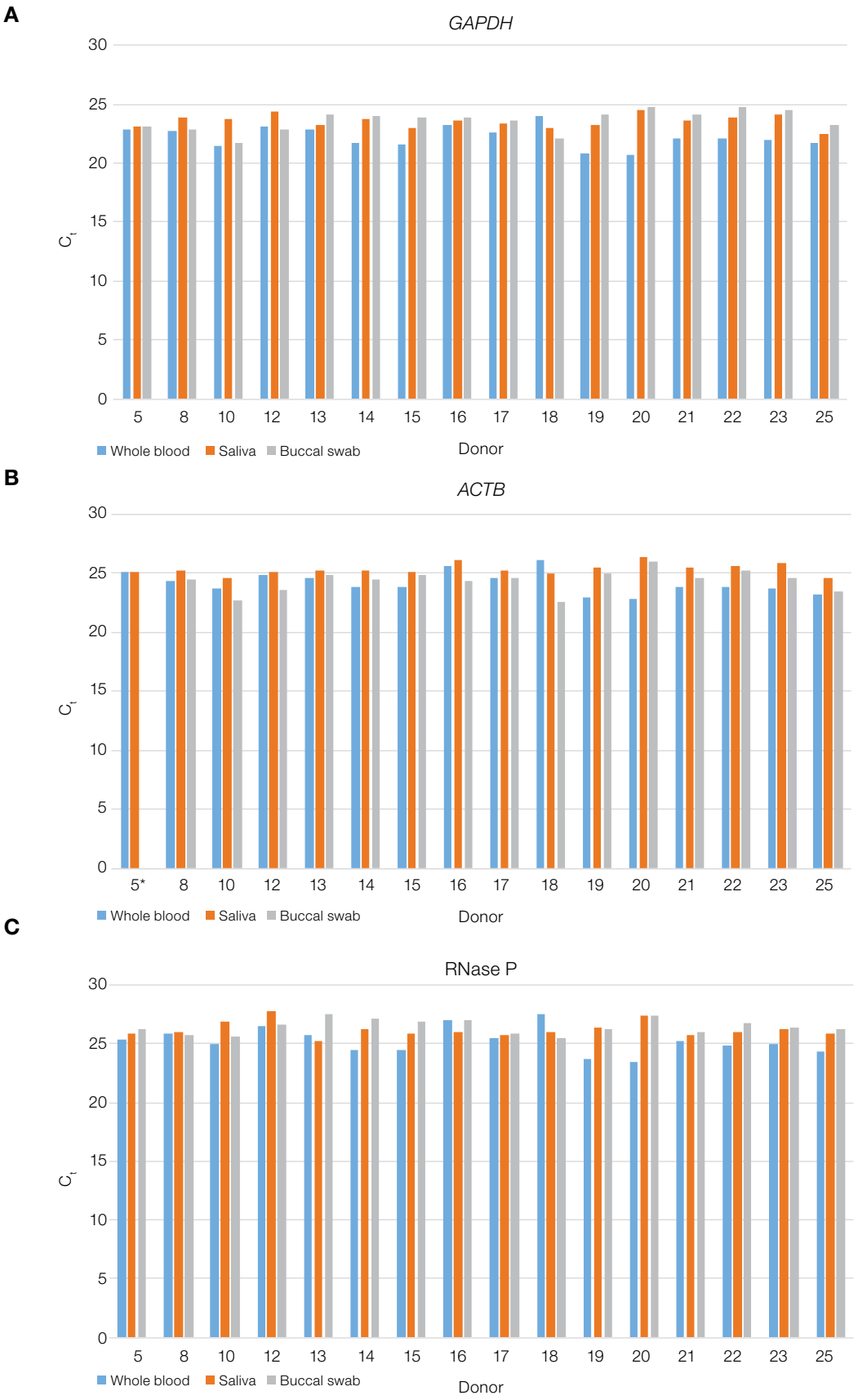

Figure 4. C_t values obtained by qPCR with two housekeeping genes, (A) *GAPDH*, (B) *ACTB*, and (C) RNase P. C_t values obtained using TaqMan Assays and reagents were consistent across whole blood, saliva, and buccal swab samples for all donors. (*There was not enough sample from the buccal swab of donor 5 to run the assay for *ACTB*.)

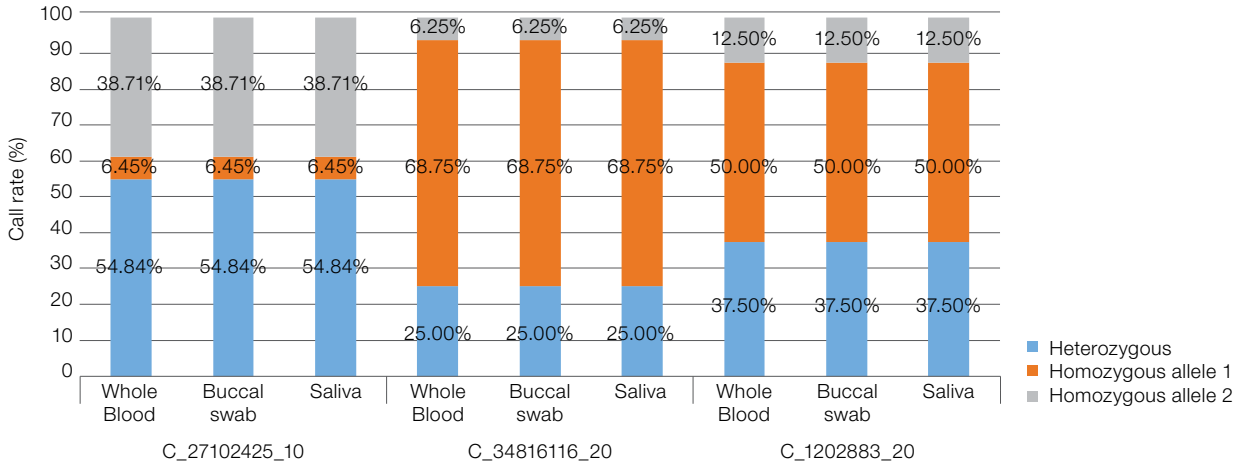

Figure 5. Genotyping percent call rates for heterozygous (blue), homozygous allele 1 (orange), and homozygous allele 2 (gray), detected with three SNP assays (C_27102425_10, C_34816116_20, and C_1202883_20) across matched blood, buccal swab, and saliva samples from 16 donors.

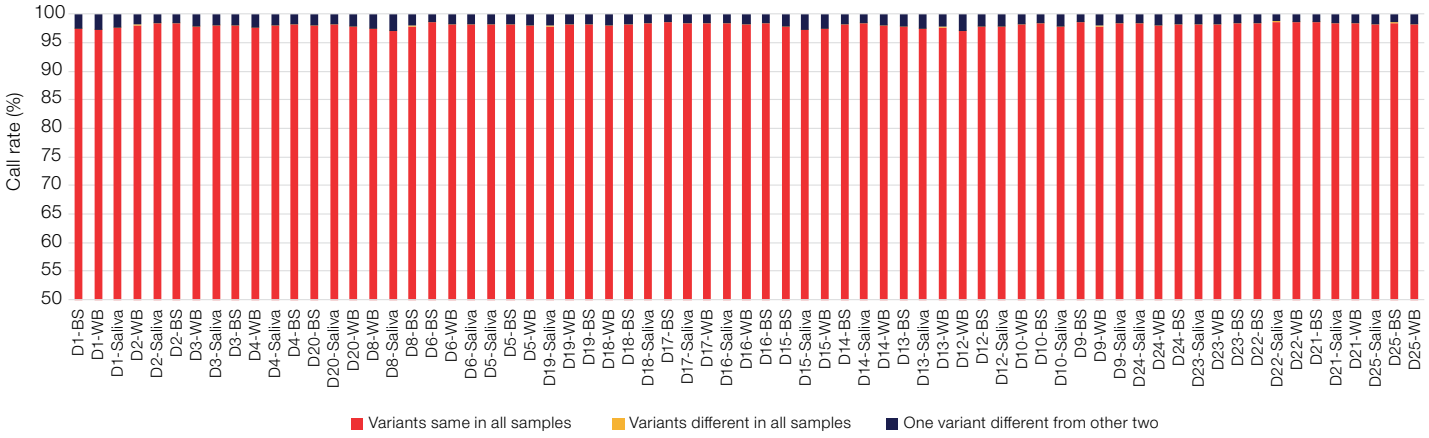

Figure 6. Variant calling concordance at shared loci across whole blood, buccal swab, and saliva samples from 22 donors, determined via exome sequencing using an Ion AmpliSeq workflow. WB: whole blood; BS: buccal swab.

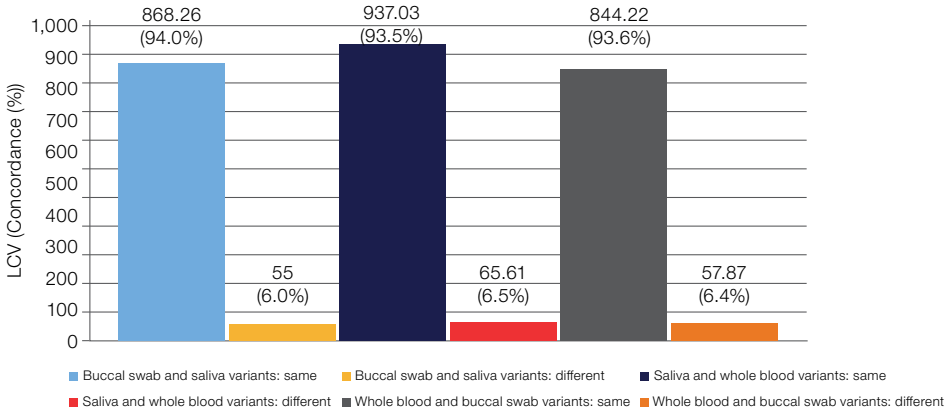

Figure 7. Loci containing variant (LCV) and variant calling concordance across matched whole blood, buccal swab, and saliva samples, determined using the Ion AmpliSeq workflow.

Conclusions

Extraction from three sample types (whole blood, buccal swabs, and saliva) using the MagMAX DNA Multi-Sample Ultra 2.0 Kit produced high-quality eluates yielding enough DNA for downstream qPCR and NGS applications. When analyzed on the NanoDrop 8000 Spectrophotometer, eluates showed high purity by consistent A_{260}/A_{280} and A_{260}/A_{230} absorbance ratios. Extracted DNA showed consistent genotyping concordance with qPCR, as well as variant calling concordance with whole-exome sequencing across matched sample types. Results from this study show that saliva collection is a cost-effective and noninvasive sampling method that can be used for high-sensitivity downstream applications such as genomic variant calling via qPCR, and NGS with Ion AmpliSeq exome sequencing workflows.

References

1. Cozier YC, Palmer JR, Rosenberg L (2004) Comparison of methods for collection of DNA samples by mail in the Black Women's Health Study. *Ann Epidemiol* 14(2):117–122.
2. Javaid MA, Ahmed AS, Durand R, et al. (2016) Saliva as a diagnostic tool for oral and systemic diseases. *J Oral Biol Craniofac Res 6(1):66–75.*

Ordering information

Product	Cat. No.
MagMAX DNA Multi-Sample Ultra 2.0 Kit	A36570
NanoDrop 8000 Spectrophotometer	ND-8000-GL
Qubit 4 Fluorometer	Q33238
Qubit dsDNA HS Assay Kit	Q32851
MagMAX Viral/Pathogen II (MVP II) Kit	A48383
TaqPath COVID-19 Combo Kit	A47813
KingFisher Flex Purification System	24074431
Ion GeneStudio S5 System	A38194
Ion Chef Instrument	4484177
Ion Reporter Server System	4487118
Ion AmpliSeq Exome RDY Kit 1x8	A38262
TaqMan Assays (various)	4331182
TaqMan Copy Number Reference Assay, human, RNase P	4403328
QuantStudio 12K Flex Real-Time PCR System	4471087
TaqMan Genotyping Master Mix	4381657
QuantStudio 3 Real-Time PCR System, 96-well, 0.1 mL, with laptop	A28566
QuantStudio 5 Real-Time PCR System with laptop	A28570

Find out more at **thermofisher.com/magmaxultra**

ThermoFisher SCIENTIFIC