

Evaluation of performance of gDNA from saliva collected with Oragene®•DNA for the purpose of SNP and CNV analysis on the Affymetrix Genome-Wide Human SNP Array 6.0

Rafal M. Iwaszow¹, Carolyn James¹, Pingzhao Hu², Richard F. Wintle², Steve Scherer²

¹DNA Genotek Inc., Ottawa, Canada ²The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Canada

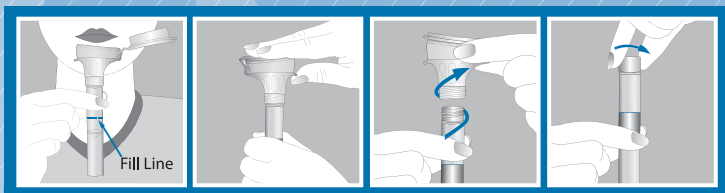
Introduction

For any type of genetic analysis (e.g. population studies, clinical genetic testing, paternity testing, pharmacogenomic testing), non-invasive and easy-to-use sample collection techniques are preferred because they increase compliance rates (1) and reduce costs. Oragene•DNA is a DNA self-collection kit that is completely non-invasive, intuitive to use, and stabilizes DNA at temperatures up to 50°C which facilitates transport and storage by eliminating the need for refrigeration.

For these reasons, Oragene•DNA is often used as a convenient method for collecting high quality genomic DNA (gDNA) from saliva. Numerous studies have shown that DNA from Oragene•DNA/saliva samples gives equivalent results to DNA from blood for applications such as PCR, SNP genotyping and microarrays (2, 3). Furthermore, of all oral collection methods (e.g. swabs, mouthwash, cytobrush), Oragene•DNA provides the highest yield and highest quality of human genomic DNA (4). Genetic research into common and complex human diseases continues to benefit from advances in DNA microarray technology. In recent years, DNA microarrays such as the Affymetrix Genome-Wide Human SNP Array 6.0 (Affy 6.0) have evolved to provide higher resolution through higher density arrays and now also include probes for detection of copy number variation. The Affy 6.0 microarray features more than 1.8 million markers for genetic variation, including more than 906,600 single nucleotide polymorphisms (SNPs) and more than 946,000 additional probes for the detection of copy number variation (CNVs) (5). Previous studies, have shown that genomic DNA from saliva performs just as well as genomic DNA from blood when evaluating SNPs. Similarly, this study reports saliva samples generated call rates >96% and had a >99% concordance with results from blood samples. In the present study we investigated using saliva as a source of gDNA for the evaluation of genotyping and CNV analysis. Furthermore we demonstrated the intra-donor reproducibility of such results.

Materials and Methods

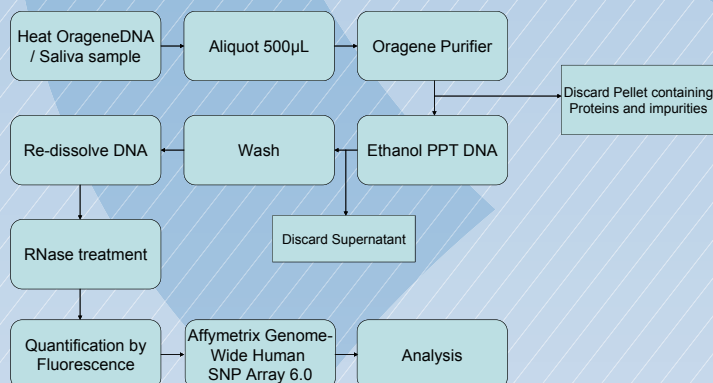
Figure 1 Collection of Saliva using Oragene®•DNA



- Saliva was collected according to DNA Genotek protocol PD-PR-061.
- 3 saliva samples were collected from 4 donors.
- Each saliva sample was collected on a different day.

- Blood samples were collected from the same 4 donors that donated saliva.
- 8 mL of whole blood was collected using EDTA tubes.
- Blood samples were centrifuged and the buffy coat was collected.
- DNA from buffy coat was purified using the Qiagen Blood mini kit.

Figure 2 Schematic representation of processing flow



- OrageneDNA/saliva samples were purified according to DNA Genotek protocol PD-PR-006.
- Purified DNA was RNase treated according to DNA Genotek protocol PD-PR-040.
- DNA was quantified using the Invitrogen Picogreen Quanti-it kit.

- Genotype data was generated using the birdseed algorithm (6) implemented in the Affymetrix genotype console. The concordance of each pair of replicates in a given donor sample was then calculated.

- Two copy number variation analysis methods were used. The first one is the Canary algorithm implemented in the Affymetrix genotype console. The second one is the hidden Markov model (HMM) method implemented in the Partek Genomics Suite software (www.partek.com). The reference set was based on Affy 6.0 data generated by TCGA from 132 healthy control samples. The test (four donor samples in triplicate) and control data were first normalized separately and then the log₂ ratio of case to control data was taken for CNV region detection.

Results

Figure 3 Genotyping QC Call Rates for saliva and blood samples

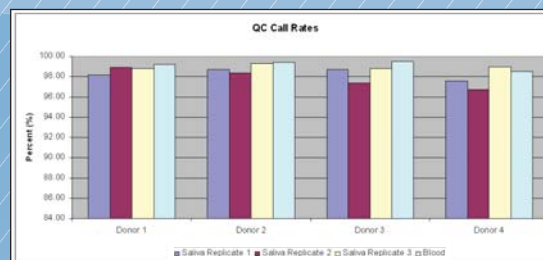


Table 1 QC parameters for saliva and blood samples

Donor	First Run		Second Run			Third Run		
	Saliva Replicate 1		Saliva Replicate 2	Saliva Replicate 3		Blood		
	CQC	MAPD	CQC	CQC	MAPD	CQC	MAPD	
1	2.93	0.21	3.13	0.20	3.32	0.21	1.96	0.23
2	2.71	0.20	3.08	0.21	3.73	0.26	2.75	0.21
3	2.59	0.24	3.01	0.22	3.07	0.23	2.43	0.21
4	2.49	0.23	2.83	0.25	3.10	0.21	1.20	0.30

- The Contrast Quality Control (CQC) algorithm was developed by Affymetrix and used in Genotyping Console for quality control assessment. The CQC cutoff recommended by Affymetrix to indicate whether a sample has good quality is >0.4. All saliva and blood samples are well above this cutoff.
- The Median of the absolute values of all pairwise difference (MAPD) was used to evaluate whether the chip/array produce data that is useful for copy number (CN) analysis. MAPD is defined as the median of the absolute values of all pairwise differences between log₂ ratios for a given chip. Affymetrix recommends that the MAPD value be <0.4 if the data is to be used for CNV analysis. All saliva and blood samples met the above requirements.

Table 2 Genotyping concordance

Saliva Replicate	Donor 1	Donor 2	Donor 3	Donor 4
1 vs 2	99.94	99.90	99.80	99.82
1 vs 3	99.95	99.93	99.85	99.92
2 vs 3	99.98	99.93	99.91	99.85
Saliva vs Blood	99.54	99.83	99.82	99.83

- Intra-donor genotyping concordance for saliva samples was on average 99.9%.
- Saliva / blood genotyping concordance was on average 99.8%.

Table 3 Saliva intra-donor CNV reproducibility and saliva / blood CNV concordance

Donor sample	# CNVs in saliva*			# Common CNVs in at least 2 saliva replicates**	# CNVs in Blood*	# Common CNVs in Saliva and Blood***
	Replicate 1	Replicate 2	Replicate 3			
1	25	24	30	25	25	21
2	18	18	22	17	18	14
3	19	18	17	19	17	15
4	14	14	19	15	11	10

* CNV calls are based on data integrated after running two algorithms

** Saliva CNVs that were called in at least two replicates

*** More than 80% reproducibility of CNVs between saliva replicates as well as between saliva and blood

Conclusions

- Saliva collected using the Oragene•DNA Self-Collection Kit provides genomic DNA of sufficient quality for genotyping and CNV analysis on the Affymetrix Genome-Wide Human SNP Array 6.0.
- DNA from saliva does not vary over time as demonstrated through intra-donor genotyping concordance and CNV reproducibility of samples taken from the same donor on three different days.
- DNA from saliva generates highly concordant data compared with DNA from blood for the same donor. As demonstrated by the genotyping concordance and CNV reproducibility.

References

- (1) Rylander-Rudqvist T, Häkansson N, Tybring G, Wolk A. (2006). Quality and quantity of saliva DNA obtained from the self-administered Oragene method—a pilot study on the cohort of Swedish men. *Cancer Epidemiol Biomarkers Prev.* 15, 1742-1745.
- (2) Ahluwalia N, Kavaslar N, Schackwitz W, Ustaszewska A, Collier J.M, Hebert S, Doelle H, Dent R, Pennacchio, L.A, McPherson, R. (2006). A PIV Q62P variant linked to human obesity. *Human Molecular Genetics.* 15, 387-391.
- (3) Reynolds J, Do T, Hongo D, Kuramoto K, Biggs III, W, Frech, C. (2007). Comparison of high density genotyping results from saliva and blood samples on Affymetrix GeneChip® GenomeWide SNP 6.0 arrays.
- (4) Rogers N, Cole S, Lan H, Crossa A, Demerath E. (2007). New Saliva DNA Collection method compared to buccal cell collection techniques for epidemiological studies. *American Journal of Human Biology.* 19, 319-326.
- (5) Affymetrix, Inc. http://www.affymetrix.com/products_services/arrays/specific/genome_wide_snp6/genome_wide_snp_6 affx
- (6) Joshua M Korn et al., Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nature Genetics* 40:1253-1260, 2008.