Evaluation of performance of genomic DNA from saliva collected with Oragene®•DNA for the purpose of SNP discovery and CNV analysis on Illumina® BeadChip technology

Mike Tayeb and Rafal Iwasiow (DNA Genotek Inc., Ottawa, Canada)
Pingzhao Hu, Richard F. Wintle and Steve Scherer (The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Canada)

Introduction

Single nucleotide polymorphisms (SNPs) and structural variation (such as copy number variants, CNVs) form the basis of genetic diversity among individuals. The Illumina Infinium® Human610-Quad and the Human1M-Duo BeadChip microarrays include more than 600,000 and 1.1 million markers for SNPs/CNVs, respectively. These technologies enable genome-wide association studies by providing a high-density, high-throughput method of SNP discovery and CNV analysis.

Traditionally, genomic DNA from blood was used for GWAS. However, genomic DNA from saliva is increasingly being used as an alternative to DNA from blood. Oragene•DNA is a DNA self-collection kit that provides an alternative to blood. Oragene•DNA simplifies sample collection and transport by eliminating phlebotomy costs and stabilizing DNA at ambient temperature. The collection procedure is intuitive and non-invasive which leads to higher donor compliance rates[1] and the kits yield large quantities of high molecular weight gDNA.

In this study, we investigate the use of genomic DNA extracted from saliva collected using Oragene•DNA self-collection kits for SNP and CNV analysis on Illumina BeadChip technologies. We compare the performance of paired blood and saliva samples and demonstrate the intra-donor reproducibility of the results.

Materials and methods

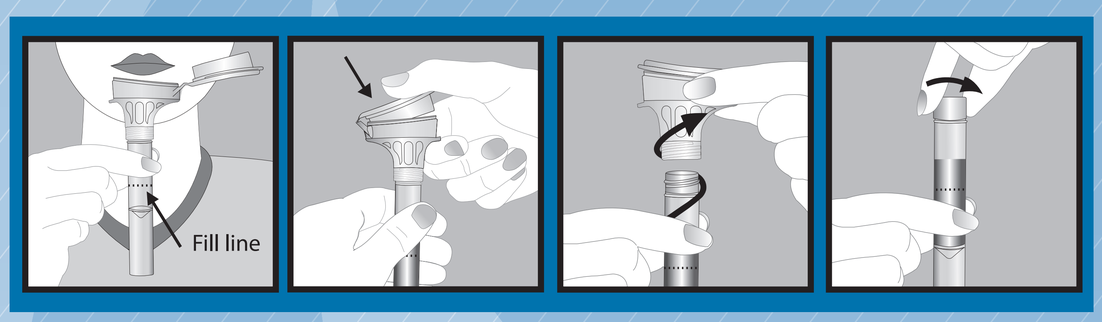

Figure 1: Collection of saliva using Oragene•DNA

Sample Collection:
- Saliva was collected according to the instructions provided in the kit.
- Samples were collected from nine donors, with four donors providing a second sample on a different day.
- 8 mL of whole blood was collected from the same donors using EDTA tubes.

DNA Extraction:
- Blood samples were centrifuged and the buffy coat was collected.
- DNA from buffy coat was purified using the Qiagen QIAamp Blood Mini Kit.
- Oragene•DNA/saliva samples were purified according to DNA Genotek protocol PD-PR-006.
- DNA was quantified using the Invitrogen Picogreen® Quant-iT™.

Data Analysis – Human610-Quad:
- Genotype calls were made using the algorithm contained in the Illumina GenomeStudio Software. Normalized X and Y values were also drawn from GenomeStudio.

Data Analysis – Human1MDuo:
- Genotype calls were made using the algorithm contained in the Illumina BeadStudio Software.
- QuantiSNP[2] and PennCNV[3] were used for copy-number analysis. A minimum of 5 markers was required for a CN call. CNV results from each analysis method were integrated. The reference set was based on HapMap samples.

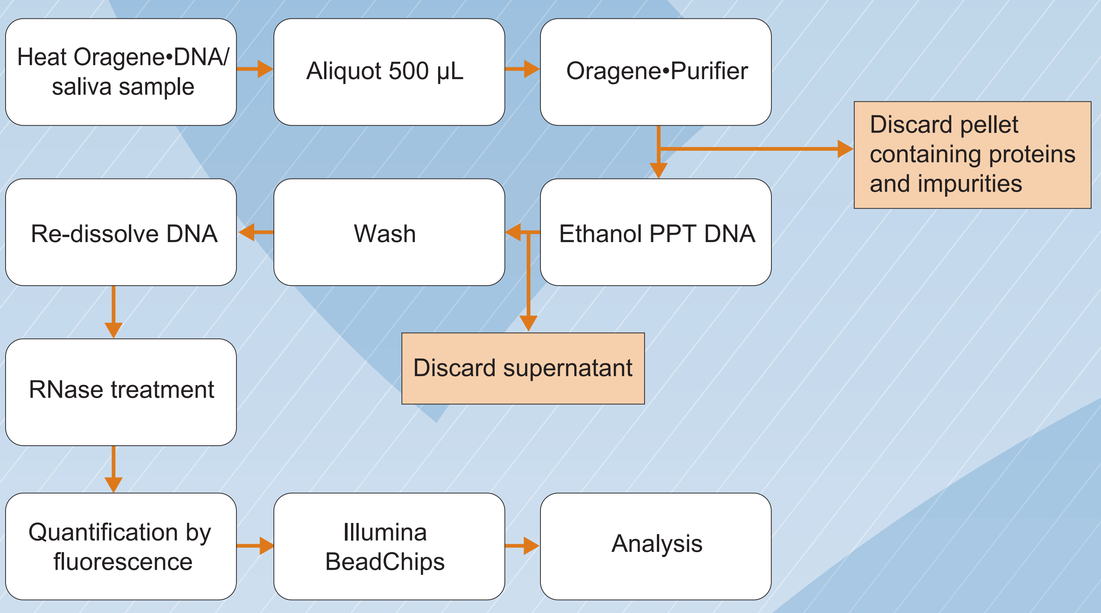

Figure 2: Schematic representation of processing flow of Oragene•DNA/saliva samples

Results

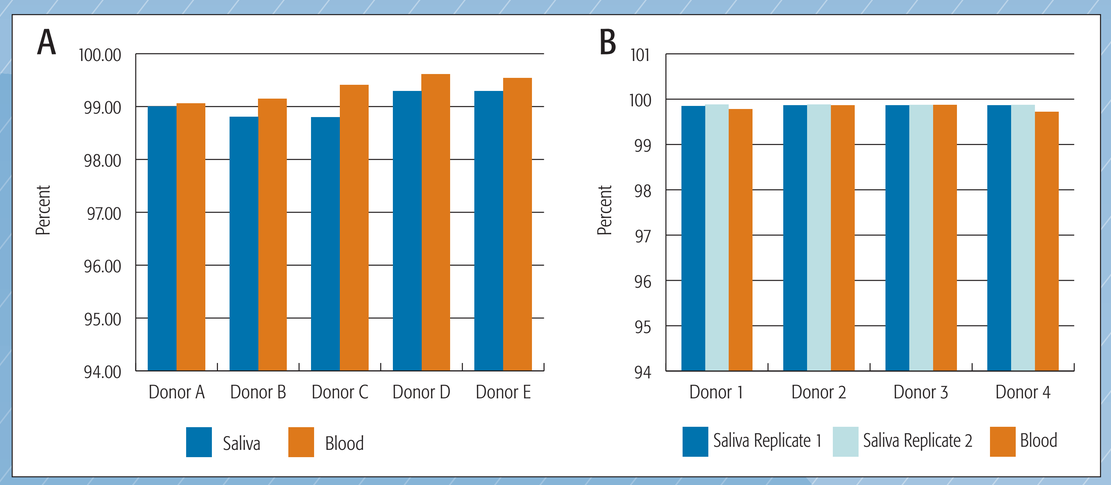

Figure 3: Genotyping QC call rates for saliva and blood samples on the Human 610-Quad (A) and the Human1M-Duo (B)

SDLRR – Human610-Quad		
Donor	Saliva	Blood
A	0.3136	0.3013
B	0.3521	0.3317
C	0.3958	0.3050
D	0.2985	0.2495
E	0.3007	0.2550

SDLRR – Human1M-Duo			
	Saliva		
Donor	Replicate 1	Replicate 2	Blood
1	0.1593	0.1568	0.2343
2	0.1563	0.1484	0.2009
3	0.1648	0.1562	0.1901
4	0.1577	0.1527	0.2375

Table 1: QC parameters for saliva and blood samples

- The log *r* ratio (LRR) is defined as the logged ratio of the observed probe intensity to the expected intensity (determined from the reference sample set). The standard deviation of the LRR (SDLRR) is a measure of background noise on the array and low SDLRR values are indicative of high quality DNA samples.
- For the Human610-Quad BeadChip, all SDLRR values were <0.4. For the Human1M-Duo all SDLRR values for saliva were <0.2. The values for saliva were slightly lower than the values for blood.

Human610-Quad	
Donor	% Concordance: saliva vs blood
A	99.99
B	99.99
C	99.99
D	99.99
E	100.00

Human1M-Duo		
Donor	% Concordance: Saliva rep. 1 vs Saliva rep. 2	% Concordance: saliva vs blood
1	100.00	99.99
2	100.00	100.00
3	100.00	100.00
4	100.00	99.99

Table 2: Genotyping concordance

- Intra-donor genotyping concordance for saliva samples was 100% for all samples (Human 1M-Duo). Saliva/blood genotyping concordance was on average 99.99%.

Donor	# CNVs in saliva		# Common CNVs in saliva replicates	# CNVs in blood	# Common CNVs in saliva and blood
	Replicate 1	Replicate 2			
1	26	24	23	13	13
2	16	21	15	9	7
3	17	18	15	10	9
4	18	21	16	8	7

Table 3: Saliva intra-donor CNV reproducibility and saliva/blood CNV concordance on the Illumina Human1M-Duo

- Greater than 90% reproducibility between saliva replicates.
- Greater than 80% reproducibility between paired blood and saliva.

Conclusions

- Saliva collected using the Oragene•DNA self-collection kit provides genomic DNA of sufficient quality for genotyping on the Illumina Human610-Quad and both genotyping and CNV analysis on the Human1M-Duo BeadChip arrays.
- Both saliva and blood samples performed better on the Human1M-Duo.
- 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 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, Håkansson N, Tybring G, Wolk A. (2006). Quality and quantity of saliva DNA obtained from the self-administrated Oragene method–a pilot study on the cohort of Swedish men. *Cancer Epidemiol Biomarkers Prev.* 9, 1742-1745.
2 Colella S et al. (2007) QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res.* 35(6):2013-25.
3 Wang, K. et al. (2007) PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* 17:1665-1674.

DNA genotek *Superior Samples Proven Performance*

www.tcag.ca