Evaluation of pharmacogenetic markers by exome-sequencing of DNA extracted from saliva samples

Rafal M. Iwasiow and Mike Tayeb
DNA Genotek, Inc. Ottawa, Ontario, Canada

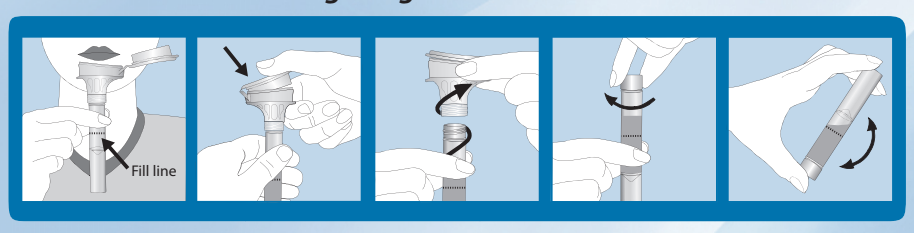
Introduction

With the rapid technological advancement and decreasing cost of next generation sequencing that has occurred over the last few years, it has become more practical to sequence larger numbers of samples. The Oragene®•DNA self-collection kit facilitates the non-invasive collection of reliable, high quality genomic DNA and results in samples that are stable at ambient temperature for years. The extracted DNA is suitable for next generation sequencing applications. In this study we investigated DNA extracted from saliva collected using Oragene•DNA and paired-end sequenced on the Illumina Genome Analyzer II for the purpose of identifying pharmacogenetic markers for drug metabolism.

To evaluate the performance of the saliva samples collected using Oragene•DNA we compared the data against blood samples collected from the same donors.

Materials and methods

Collection of saliva using Oragene•DNA



Using an Oragene•DNA self-collection kit, 2 mL of saliva was collected from each of 3 donors. DNA was extracted from the Oragene•DNA/saliva samples using prepIT•L2P (protocol PD-PR-006, DNA Genotek). EDTA tubes were used to collect 8 mL of whole blood from the same three donors. The buffy coat was collected via centrifugation and the DNA extracted using the QIAamp DNA Blood Mini Kit (Qiagen). All extracted DNA was quantified using the Picogreen® Quant-IT™ kit (Invitrogen).

DNA samples were enriched for the human exome using the Agilent SureSelect All Exon kit and sequenced using paired-end sequencing on the Illumina Genome Analyzer II. One lane per sample was used.

For sample analysis, 1854 SNPs involved in drug metabolism and metabolic disorders were identified. Of these 1854 SNPs, 1323 were present in sequences captured by SureSelect. The data for those 1323 SNPs was analyzed.

Results

All samples generated in excess of 6 x 10^9 bases per lane on the Illumina Genome Analyzer II. There was no statistical difference in the number of bases sequenced between blood and saliva samples. Coverage of the exome was in excess of 95% and >49% of the bases sequenced map to the exome captured by Agilent SureSelect. A mean sequencing depth of 76x was achieved across all blood and saliva samples.

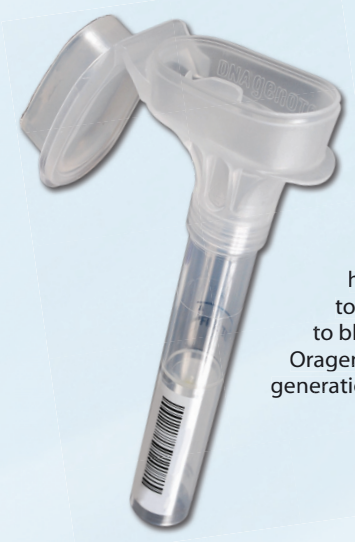
	Donor 1		Donor 2		Donor 3	
	Saliva	Blood	Saliva	Blood	Saliva	Blood
Lane yield (kb)	6,268,330	6,447,860	6,434,248	6,217,788	6,072,430	6,676,662
Proportion in exome	65.4%	45.6%	45.1%	46.2%	45.6%	47.0%
Exome coverage (%)	99.1%	99.1%	95.0%	99.3%	99.2%	98.9%
Mean depth	81.8x	76.4x	72.1x	74.5x	69.8x	81.8x

Concordance between saliva and blood was examined for the SNPs of interest. When all genotypes were included in the analysis the average concordance across three donors was 99.0%. The concordance increased to 100% when the calls were filtered based on a quality score of greater than or equal to 20.

		All genotypes			QV >= 10			QV >= 20		
		# of SNPs called	# of common SNPs	Concordance	# of SNPs called	# of common SNPs	Concordance	# of SNPs called	# of common SNPs	Concordance
Donor 1	Saliva	1268	1262	98.9%	1125	1110	99.6%	1086	1076	100%
	Blood	1278			1138			1098		
Donor 2	Saliva	1235	1224	98.6%	1167	1086	98.7%	1093	1043	100%
	Blood	1290			1140			1096		
Donor 3	Saliva	1279	1266	99.4%	1140	1123	99.6%	1101	1083	100%
	Blood	1282			1140			1105		

Both saliva and blood samples can be used to investigate genetic variations using next generation sequencing technologies. We were able to successfully interrogate the samples for SNPs in genes involved in drug metabolism and observed perfect correlation between blood and saliva.

Gene	rsID	Donor 1		Donor 2		Donor 3		Clinical significance
		Saliva	Blood	Saliva	Blood	Saliva	Blood	
CYP2C9	rs1057910	A	A	A	A	A	A	Warfarin sensitivity
CYP2C19	rs17882687 rs17878459 rs41291556	A G T	A G T	A G T	A G T	A G T	A G T	Clopidogrel efficacy
CYP2D6	rs1065852 rs5030862	G C	G C	G C	G C	G C	G C	Codeine sensitivity
ABCB1	rs10276036 rs28381867	C C	C C	C/T C	C/T C	C/T C	C/T C	Colchicine resistance
DPYD	rs3918289 rs17376848	G A	G A	G A	G A	G A	G A	5-Fluorouracil toxicity



Conclusions

DNA extracted from saliva collected using Oragene•DNA was successfully sequenced using the Agilent SureSelect Human All Exon kit and the Illumina Genome Analyzer II. The high coverage and sequencing depth allowed us to confidently mine the data for SNPs of interest, in this case, several pharmacogenetic markers. Concordance between paired blood and saliva samples was extremely high, exceeding 98.5% on unfiltered data and increasing to 100% with a QV-cutoff of 20. Given the similar performance to blood, these results indicate that saliva collected using Oragene•DNA is a suitable alternative to blood for next generation sequencing applications.

DNA genotek

Superior samples
Proven performance

www.dnagenotek.com • support@dnagenotek.com