Saliva collected using the Oragene® family of products is a reliable source of DNA for HLA typing using Next Generation Sequencing

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

Introduction

Transplant centres and marrow donor registries continue to evolve their processes and technology with the goal of reducing time required to find a match for a patient in need of a hematopoietic stem cell transplant. Critical factors in achieving this goal include facilitating donor recruitment, collecting samples that meet requirements for the HLA typing technologies, integrating cost efficiencies into laboratory procedures and eliminating allele ambiguities. One way to acheive these goals is to examine sample types. Non-invasive saliva sample collection with the Oragene® family of products delivers a high quality and reliable solution for collection, stabilization and transportation of DNA. Saliva samples collected with Oragene are currently being used as a reliable source of DNA for technologies such as SSOP, SSP and SBT in both marrow donor registries and transplant centres. There is significant industry interest in the future use of Next Generation Sequencing as a cost effective solution for HLA typing of high volumes of samples while addressing the challenge of eliminating allele ambiguities.

To evaluate the performance of DNA from Oragene/saliva samples we compared the data against DNA from blood samples collected from the same individuals. Using the HLAseq™ panel (RainDance™ Technologies) to capture the entire HLA super locus (approx 3.8 Mb) we demonstrated the ability to enrich either saliva or blood DNA samples to investigate HLA-related genetic variations using Next Generation Sequencing technologies. Additionally, we demonstrated that accurate HLA calls can be made from the resulting data using the Assign™ MPS software from Conexio Genomics Inc.

Materials and methods

Sample collection and DNA extraction

- Saliva (2 mL) was collected according to the instructions provided in the Oragene self-collection kit (Figure 1).
- Saliva samples were collected from 4 donors.
- Oragene/saliva samples were purified using prepIT®•L2P according to DNA Genotek protocol PD-PR-006.
- DNA was quantified using the Quant-iT™ Picogreen® kit (Invitrogen).
- Blood samples were collected from the same 4 donors that donated saliva.
- Whole blood (8 mL) was collected using EDTA tubes.
- DNA from either whole blood or buffy coat was purified using Qiagen spin-column kits.

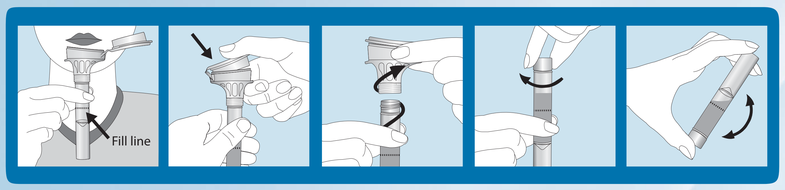

***Figure 1**: Oragene collection instructions.*

Library preparation and sequencing – Ion Torrent™

- Samples were nebulized, cleaned up with a Qiagen MinElute® column and enriched using the RDT 1000 system from RainDance Technologies.
- Libraries were prepared using the Ion Xpress™ Plus gDNA and Amplicon Library Preparation.
- DNA was fragmented followed by ligation of Ion Torrent P1 and Ion Xpress Barcode Adapters.
- Adapter-ligated libraries were SPRI-purified, size selected by agarose gel, nick-translated and PCR amplified.
- Library size and concentration was determined using a Bioanalyzer (Agilent Technologies).
- Samples were pooled and prepared for sequencing using the Ion PGM™ 200 Sequencing Kit (Ion Torrent) protocol.
- Entire pooled sample was loaded on the 318 chip and sequenced on the PGM 200 for 65 cycles.

Library preparation and sequencing – Illumina HiSeq 2000

- Samples were nebulized and enriched using the RDT 1000 system from RainDance Technologies.
- RDT amplicons were end-repaired, phosphorylated and ligated/concatenated overnight and cleaned-up using a QIAquick® spin column.
- Library was prepared according to Illumina® protocol "TruSeq® DNAseq Sample Preparation".
- Custom index (barcode) adapters were ligated via T-A mediated ligation.
- Ligated products were size-selected by gel purification and PCR amplified using Illumina singleton primers.
- Library size and concentration was determined using a Bioanalyzer (Agilent Technologies).
- Libraries were sequenced using the Illumina HiSeq™2000.

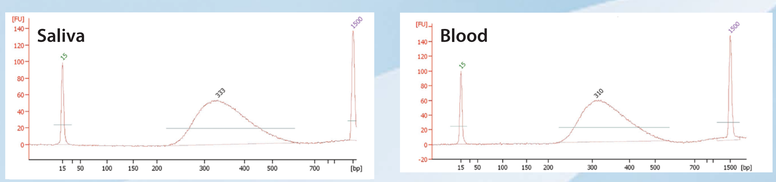

***Figure 2**: Representative Bioanalyzer traces of prepared libraries from saliva and blood DNA.*

Data analysis

Ion Torrent: data processing, filtering and base calling was done using the Ion Torrent server, Torrent Suite v2.2.

Illumina HiSeq2000: Base calling and quality filtering was done using RTA 1.12.4 (HiSeq Control Software 1.4.5). Quality filtering was performed using Illumina CASVA 1.8.2. Raw sequence reads were aligned to human (hg19) and variants were reported.

The complete MHC sequence from the RainDance HLAseq panel was analyzed using Assign MPS v1.0 from Conexio Genomics (Fremantle, Western Australia). Briefly, reads containing sequences with identity to alleles of HLA-A, -B, -C, -DRB1, -DQB1 and -DPB1 were extracted from the complete set of MHC reads. The extracted sequences were aligned to the appropriate reference sequences. Consensus sequences were calculated and compared with HLA libraries from the IMGT HLA database v3.9.

Results

Approximately 400 Mb of data was generated using the Ion Torrent PGM 318 Chip, far below the expected amount of 1 Gb. Both saliva and blood samples performed similarly indicating that the problem was not due to sample type (Table 1). There was insufficient data from the Ion Torrent to proceed to analysis. Sequencing on the Illumina HiSeq2000 was successful, with almost 30 Gb of data generated. There was no statistical difference in the number of bases sequenced in saliva and blood samples. Mean coverage was in excess of 100 for all samples (Table 2).

Table 1: Sequencing metrics from Ion Torrent PGM

Total number of bases (Mb)	402.85
Number of Q20 bases (Mb)	315.34
Total number of reads	4,359,996

	Saliva	Blood
Mean # reads	529,898	514,082
Mean length (bp)	102.78	93.19
Mean yield (Mb)	54.5	47.9
Mean coverage	2.46	2.18

Table 2: Sequencing metrics from Illumina HiSeq2000

Donor	1		2		3		4	
Sample type	Saliva	Blood	Saliva	Blood	Saliva	Blood	Saliva	Blood
Yield (Mb)	3,989	3,810	3,886	3,830	4,005	3,594	3,246	3,550
% > = Q30 bases	86.9	87.4	87.0	87.2	87.5	86.9	87.3	87.4
Mean quality score	34.4	34.5	34.4	34.5	34.6	34.4	34.5	34.6
Mean coverage	133.61	148.96	164.07	189.54	159.85	164.75	108.12	130.21

DNA extracted from saliva performed similarly to DNA from blood as indicated by the similar number of total reads generated and the number of reads for each individual HLA locus (Table 3).

Table 3: Number of total and extracted reads for each HLA locus analyzed

		Number of reads						
		Total	HLA-A	HLA-B	HLA-C	HLA-DPB1	HLA-DQB1	HLA-DRB1
Donor 1	Saliva	39,886,428	28108	29828	31334	2460	1579	9018
	Blood	40,054,288	28350	28202	29505	2737	1825	10872
Donor 2	Saliva	38,100,184	27376	26541	27314	2875	1656	8285
	Blood	35,941,094	29718	27809	29147	2981	1747	8788
Donor 3	Saliva	38,857,472	29165	28443	29006	2631	1648	11565
	Blood	32,458,764	22424	21644	22820	1959	1286	7470
Donor 4	Saliva	38,303,182	33214	32263	31838	3056	2162	11017
	Blood	35,495,834	24173	22608	24254	2224	1532	8740

Using Conexio Assign MPS, we were able to successfully make HLA calls for all 6 HLA loci interrogated (Table 4). Concordance of HLA calls between paired saliva and blood samples was 100%. Additionally, the HLA calls made in this study are 100% concordant with calls made previously using other HLA-typing methods including SSOP, SSP and SBT[1].

Table 4: HLA calls for 6 HLA loci as reported by Conexio Assign MPS

	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2
	HLA-A		HLA-B		HLA-C	
Donor 1	02:01:01:01		15:01:01	44:02:01	03	05
Donor 2	03:01:01	33:01:01	07:02:01	14:02:01	07:02:01	08:02:01
Donor 3	02:01:01:01	23:01:01	39:01:01	49:01:01	07:01:01	12:03:01
Donor 4	01:01:01	02:01:01	08:01:01	39:01:01	07:01:01	12:03:01
	HLA-DPB1		HLA-DQB1		HLA-DRB1	
Donor 1	04:01:01		03:01:01	03:02:01	04:01:01	12:01:01
Donor 2	04:01:01	03:01:01	05:01:01	06:02:01	01:02:01	15:01:01
Donor 3	04:01:01	04:02:01	03:01:01		11:01:01	11:04:01
Donor 4	01:01:01	04:01:01	02:01	05:01:01	01:01:01	03:01:01

Discussion

DNA extracted from saliva collected using the Oragene self-collection kit was successfully enriched for HLA loci using the RainDance HLAseq content panel and sequenced using the Illumina HiSeq2000.

- Prepared saliva and blood libraries were of equivalent quality (Figure 2).
- Samples were successfully barcoded, multiplexed in a single sequencing run.
- Saliva and blood had similar mean quality scores of approximately 34.5.
- Mean coverage for both saliva and blood exceeded 100.
- HLA call concordance between saliva and blood was 100%.
- HLA calls were 100% concordant with previously reported results for these donors using current HLA-typing methodologies.

This study illustrates that DNA from Oragene/saliva samples is a dependable alternative to blood for HLA typing, including Next Generation Sequencing applications. In agreement with previous exome and whole genome sequencing studies[2, 3, 4] we demonstrated that Oragene/saliva samples are a reliable source of DNA for Next Generation Sequencing applications.

Acknowledgments

We would like to thank David Sayer and his colleagues at Conexio Genomics Inc for their kind assistance with the data analysis using their software, Assign MPS.

References

1 HLA typing using saliva DNA collected with Oragene. DNA Genotek. PD-WP-020.
2 Saliva samples collected and stabilized with Oragene•DNA are a reliable source of DNA for Next Generation Sequencing. DNA Genotek. MK-00014.
3 Evaluation of pharmacogenetic markers by exome-sequencing of DNA extracted from saliva samples. DNA Genotek. MK-00048.
4 Quality whole human genome sequencing from saliva samples. Complete Genomics. March 2012.

Superior samples • Proven performance

U.S. Patent Nos. 8,221,381 and 7,482,116;
European Patent No. 1 513 952 and 1 956 969; Patent pending
Canadian Design Nos. 127470; 132896; 132897
U.S. D631,554 S and D640,795 S
Community Design Nos. 001095186-0001; -0002; -0003

DNA genotek

www.dnagenotek.com • support@dnagenotek.com