Performance Evaluation of High Resolution 11 Loci HLA-typing Prototype Assay Using NGS Technology

Cassandra Jabara, Marc Domanusa, Dae Hyun Kim, Shiaolan Ho, Jeroen Ademab, Nienke Westerinkb, Natalia Marlowea

*Abbott Molecular, 1300 E. Touhy Ave., Des Plaines, IL USA, bGenDx, Yalelaan 48, 3584 CM Utrecht, The Netherlands

AIM

The aim of this research study was to evaluate the workflow and performance of a NGS-based high resolution HLA typing prototype assay developed by GenDx.

METHODS

48 UCLA DNA Reference panel samples were amplified at 11 loci (HLA-A, B, C, DRB1, DRB3/4/5, DQB1, DQA1, DPB, DPA) using NGSgo®-AmpX primers with Long-Range (LR) PCR and sequenced on the Illumina MiSeq platform. A full gene (5' UTR to 3' UTR) amplification strategy was applied for class I and partial amplification covering all relevant exons and introns for class II (Figure 1). Individual amplicons were generated for each sample using 100 ng of gDNA per LR-PCR reaction (Figure 2). Intrapatient amplicons were pooled, enzymatically fragmented, and purified. Purified fragmented dsDNA was brought through library preparation using NGSgo®-LibX: DNA library prep kit for Illumina and barcoded using NEBNext (New England Biolabs) adaptors for MiSeq sequencing. Due to the limited number of indices (24) in NEBNext, two 2x150 paired-end sequencing runs were completed with each run containing 24 samples. Demultiplexed reads were aligned and typed using GenDx NGSeTengine® v1.4.0 (IMGT v 3.15.0) and Omixon Target v1.8.0 (IMGT v 3.10.0) software.

CONCLUSION

HLA typing by NGS has the potential to provide high-resolution unambiguous phased results across multiple loci and in a high throughput, multi-sample manner. In this research study we have demonstrated that prototype NGS HLA typing assay has high accuracy, throughput, and provided unambiguous high resolution results for 47/48 UCLA DNA Reference panel samples. This assay delivered full genomic sequences for all targeted HLA loci and had the capacity to analyze 11 loci for 24 samples per MiSeq run.

RESULTS

47/48 UCLA samples successfully amplified. A single sample failed due to gDNA integrity (Figure 4). For HLA-A, B, C, DRB1, DRB3/4/5, DQB1, DQA1, DPB and DPA an average of 5 µg (~200 µg/mL) of dsDNA was produced per amplicon (Figure 3), and 99.7% agreed with the reference SBT allele assignments (Figure 7). In one sample DRB1 04:05/15:02 were agreed at 4 digit allele typing level. Average sequencing output was 3.65 Gb of sequence (12 million paired sequence reads) per run and a mean of 459,573 (median: 433,292) paired-end reads per sample (Table 1, Figure 5).

NGS

<table>
<thead>
<tr>
<th>Cluster density</th>
<th>Lib 1 C111-C124</th>
<th>Lib 2 C201-C224</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Identified</td>
<td>829 K/mm²</td>
<td>606 K/mm²</td>
</tr>
<tr>
<td>Total reads</td>
<td>16.5M</td>
<td>12.0 M</td>
</tr>
<tr>
<td>Pass filtered (PF) reads</td>
<td>13.5M</td>
<td>10.6 M</td>
</tr>
<tr>
<td>Minimum reads/pt</td>
<td>307,064</td>
<td>288,183</td>
</tr>
<tr>
<td>Maximum reads/pt</td>
<td>1,432,963</td>
<td>618,393</td>
</tr>
</tbody>
</table>

Table 1. MiSeq output statistics for two runs

An average of 12 million paired sequence reads (3.65 Gb) were produced per sequencing run. Each run had 24 samples total.

Figure 4. Investigation of a failed sample

For a single sample (C111) 92% of reads did not map (Fig 4A). Quantifying dsDNA produced from the 11 individual LR-PCR reaction revealed little to no amplicon production (red circles) compared to the overall population (black circle, Fig 4B).

Figure 5. Number of total and mapped reads per typed sample

An average of 460k paired-end reads were produced per sample. “83% of total reads within a sample mapped to an HLA loci. Red line – total reads per sample. Blue column - mapped reads after NGSeTengine® alignment

Typing

Figure 6. Whole gene overview and coverage as visualized by NGSeTengine®. NGSeTgo®-AmpX primers provide full-length amplification for class I genes. After assembly, linkage across exons is preserved by sequencing through intronic regions.

Figure 7. Overall concordance frequency per locus. There was 99.7% agreement with the reference SBT typing.

Figure 8. Discordance analysis. Both software packages mistyped DRB1 15:02. The closest match was 15:01, which differed at G86V exon 2 (GGT to GTG)