Introduction
We previously reported on the use of the Ion PGM next generation sequencing (NGS) platform to genotype HLA class I and class II genes by a super-high resolution, single-molecule, sequence-based typing (SS-SBT) method (Shiina et al. 2012). However, HLA alleles could not be assigned at the field 4 level at some HLA loci such as DQA1, DPA1 and DPB1 because the SNP and indel densities were too low to identify and separate both of the phases. In this regard, we have now added the single molecule, real-time (SMRT®) DNA sequencer PacBio RSII method to our analysis in order to test whether it might determine the HLA allele sequences in some of the loci with which we previously had difficulties.

In this study, we report on sequence-based genotyping of entire HLA gene sequences from the promoter-enhancer region to 3’UTR of the major HLA loci (A, B, C, DRB1, DRB345, DQA1, DQB1, DPA1 and DPB1) using 46 Japanese reference subjects who represented a distribution of more than 99.5% of the HLA alleles at each of the HLA loci and the PacBio RS II and Ion PGM systems.

Experimental design

PCR amplification

<table>
<thead>
<tr>
<th>HLAA</th>
<th>HLA-B</th>
<th>HLA-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-DQA1</td>
<td>HLA-DQB1</td>
<td>HLA-DPA1</td>
</tr>
<tr>
<td>HLA-DPRB1</td>
<td>HLA-DPRB4/5</td>
<td></td>
</tr>
</tbody>
</table>

Sequencing

Library construction (PacBio)

- Library construction (IonPGM)

Data analysis

- PacBio consensus sequence
- Mapping of the IonPGM reads to the PacBio consensus sequences by GS Reference Mapper
- Confirmation of new SNPs and indels by Sanger sequencing
- Determination of the precise HLA allele sequences

Results

HLA alleles for nine HLA loci

Example of validation by Sanger sequencing

Summary of determined alleles at the field 4 level in Japanese population

Summary of new alleles observed in this study

HLA-DQA1 (Aligned length: 7.391 bp)

- NPM: 62.1% (Reduced indels index: 1.716)

- NPM: 6.1% (Reduced indels index: 1.256)

HLA-DQB1 (Aligned length: 7.391 bp)

- NPM: 3.7% (Reduced indels index: 1.086)

Nucleotide diversity plots in HLA class II loci

- DPB1 (Aligned length: 12.395 bp)

- DQB1 (Aligned length: 9.786 bp)

- DPB1 (Aligned length: 12.395 bp)

Summary & Conclusion

- We determined 278 HLA allele sequences at the field 4 level by combining the PacBio RSII and IonPGM NGS systems.
- We collected on average 99.814% (DQA1: 99.219% - DQB1: 100%) of HLA allele sequences at the field 2 level observed in Japanese population.
- The DQA1, DPA1 and DPB1 loci like those for DRB1 and DQB1 are polymorphic throughout the entire gene regions.

NGS HLA genotyping using PacBio RSII and IonPGM together provides data at the field 4 level that precisely detects rare, novel and null alleles in population genetic and disease studies.