Single Molecule Real-Time (SMRT®) Sequencing delivers reads that span the lengths of the majority of HLA class I and II genes. Unambiguously phased 4-field HLA types without imputation. With a more accurate and complete picture, gain deeper understanding of immune-related disease causality, graft-versus-host disease in hematopoietic transplantation, and drug hypersensitivity.

- Phase polymorphisms across SNP-poor regions
- Achieve allele-level segregation
- Detect variants in regulatory regions within 5’ UTRs, introns, and 3’ UTRs
- Characterize full-length transcribed minor variants of HLA alleles
- Obtain direct evidence for new HLA alleles through de novo, reference-free consensus generation
- Flexibly adjust amplicon size (e.g., phase through either exons 2, 3 and/or 4, or the entire HLA gene) or scale project size with cost-effective multiplexing solutions

**GENERATE FULL-LENGTH ALLELE SEQUENCE**

PacBio reads

Consensus sequence generated by SMRT Analysis

5,432 bp continuous PacBio consensus read for HLA-A:24:02:01:01 allele.

**VALIDATION OF FULL-LENGTH HLA CLASS I AND EXON 2-4 LONG AMPLICON HLA CLASS II SEQUENCING ACROSS 96 SAMPLES**

<table>
<thead>
<tr>
<th>Locus</th>
<th>No of Expected Alleles</th>
<th>Alleles Discordant to Pre-types</th>
<th>Discordant Alleles Orthogonally Validated</th>
<th>Non-validated Discordant</th>
<th>% Concordance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>175</td>
<td>4</td>
<td>4/4 100%</td>
<td>0</td>
<td>100.0</td>
</tr>
<tr>
<td>B</td>
<td>180</td>
<td>3</td>
<td>3/3 100%</td>
<td>0</td>
<td>100.0</td>
</tr>
<tr>
<td>C</td>
<td>175</td>
<td>4</td>
<td>4/4 100%</td>
<td>0</td>
<td>100.0</td>
</tr>
<tr>
<td>DRB1</td>
<td>161</td>
<td>9</td>
<td>7/7 100%</td>
<td>2</td>
<td>98.8%</td>
</tr>
<tr>
<td>DQB1</td>
<td>177</td>
<td>2</td>
<td>-</td>
<td>3</td>
<td>98.3%</td>
</tr>
<tr>
<td>Total</td>
<td>868</td>
<td>22</td>
<td>18/18 100%</td>
<td>5</td>
<td>99.4%</td>
</tr>
</tbody>
</table>

SMRT Sequencing of 96 samples correctly identified 863 unique allele types and 22 potentially discordant alleles compared to low-res UCLA pre-type data. Orthogonal validation with Sanger Sequencing (of the 18 samples tested) demonstrated 100% of the potential discordant alleles matched the PacBio data. Four remaining alleles have not been validated and one DQB1 allele was miscalled due to PCR-related allele-drop out.
FROM TARGETED REGION TO FULLY PHASED, ALLELE-SPECIFIC HLA TYPES

Library Preparation
- Utilize off-the-shelf enrichment methods or develop your own assays
- Multiplex amplicons that span partial or full-length HLA genes
- Support available for library automation

SMRT Sequencing with PacBio® Systems
- Span the majority of HLA class I and II genes with ~10 kb average read lengths
- Sequence HLA class I and II genes in 2-3 hours
- Identify > 16K full-length alleles per week
- Achieve consensus accuracies > 99.999% by avoiding mapping and systematic errors

Data analysis with SMRT Analysis or SMRT Community tools
- Long Amplicon Analysis (LAA): Generate de novo HLA consensus sequences for 4-field typing
- Analyze data output using commercial HLA typing software from GenDx or Conexio

KEY REFERENCES
1. Hosomichi, K. et al. (2013) Phase-defined complete sequencing of the HLA genes by next-generation sequencing. BMC Genomics. 14, 355