

# Collection of major HLA allele sequences in Japanese population toward the precise NGS based HLA DNA typing at the field 4 level



○Shingo Suzuki<sup>1)</sup>, John Harting<sup>2)</sup>, Primo Baybayan<sup>2)</sup>, Ken Osaki<sup>3)</sup>, Miwako Kitazume<sup>3)</sup>, Junichi Sunaga<sup>3)</sup>, Swati Ranade<sup>2)</sup>, Takashi Shiina<sup>1)</sup>

<sup>1)</sup> Department of Molecular Life Science, Tokai University School of Medicine  
<sup>2)</sup> Molecular Biology Applications, Pacific Biosciences  
<sup>3)</sup> Pacific Biosciences Division, Tomy Digital Biology



## Introduction

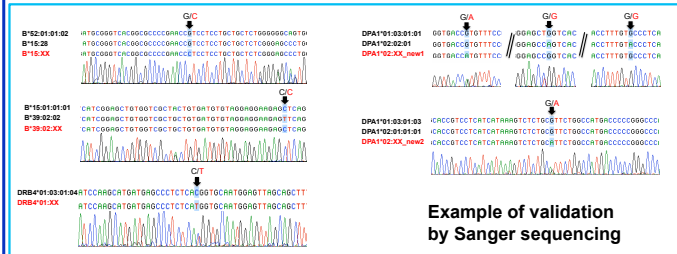
We previously reported on the use of the Ion PGM next generation sequencing (NGS) platform to genotype HLA class I and class II genes by a super-high resolution, single-molecule, sequence-based typing (SS-SBT) method (Shiina *et al.* 2012). However, HLA alleles could not be assigned at the field 4 level 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.

## Results

### HLA alleles for nine HLA loci

• Blue and yellow show new alleles that have variants in exon and/or intron/untranslated regions, respectively.  
 • Green shows the field 4 level sequences determined in this study.



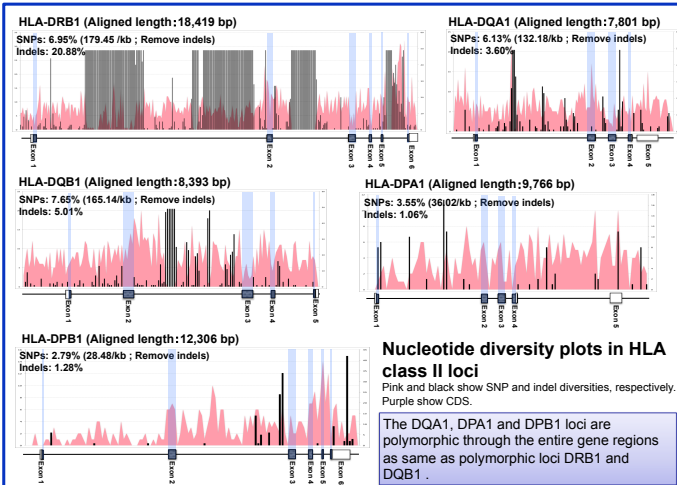
Example of validation by Sanger sequencing

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

	HLA-A	HLA-B	HLA-C	HLA-DRB1	HLA-DRB345	HLA-DQA1	HLA-DQB1	HLA-DPA1	HLA-DPB1	(Total)
From this study	20	45	26	41	26	36	28	16	40	278
From other study	0	3	1	12	0	0	1	0	0	17
Total allele No.	20	48	27	53	26	36	29	16	40	295
Accumulative allele frequency in Japanese population	99.928%	99.837%	99.935%	99.921%	-	99.219%	100%	99.837%	99.933%	99.814%

### Summary of new alleles observed in this study

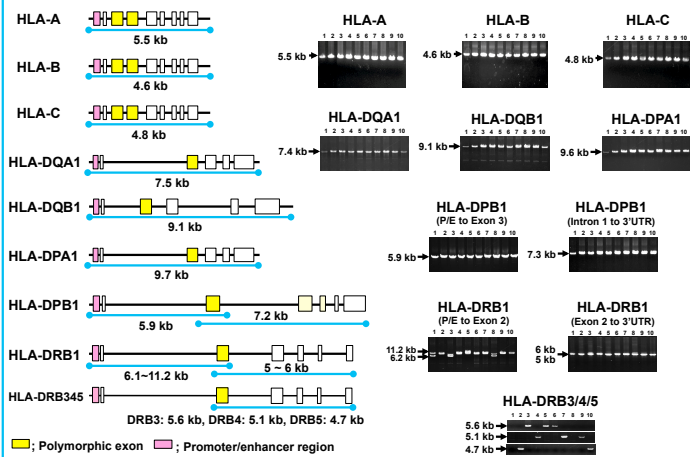
Sample ID	Locus	Allele name	Reference IMGT No	Reference Position (bp)	Location	Nucleotide Reference	Nucleotide Variant	Classification	Amino acid	
34	HLA-B	B*15:XX_new	B*15:28	HLA00191	25	Exon1	G	C	Non Synonymous	V5L
36	HLA-B	B*39:02:XX_new	B*39:02:02	HLA00275	1,008	Exon5	C	T	Synonymous	
04	HLA-DRB4	DRB4*01:XX_new	DRB4*01:03:01:01	HLA00908	13,254	Exon3	C	T	Non Synonymous	T2I4M
04	HLA-DPA1	DPA1*02:XX_new1	DPA1*02:02:01	HLA00508	361	Exon3	A	G	Synonymous	V122M
25	HLA-DPA1	DPA1*02:XX_new2	DPA1*02:01:01:02	HLA14197	4,893	Exon4	A	G	Synonymous	



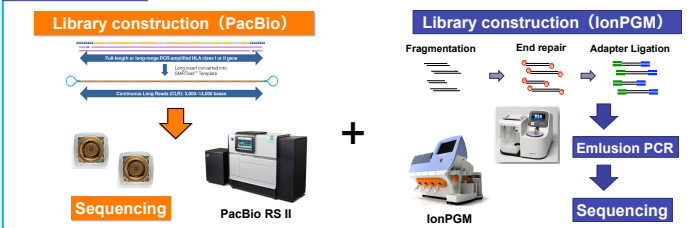
The DQA1, DPA1 and DPB1 loci are polymorphic through the entire gene regions as same as polymorphic loci DRB1 and DQB1.

## Experimental design

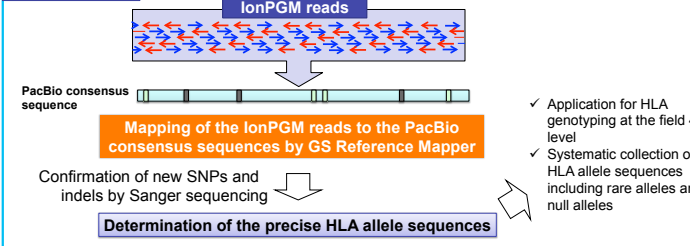
### PCR amplification



### Sequencing



### Data analysis



## 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.