Fifty-two Novel HLA Class I Alleles Identified During Routine Bone Marrow Donor Registry Typing

Lihua Hou, Bin Tu, Carly Masaberg, Kanthi Karywasam, Dannah Simbulan, James Han, Daniel Behm, Jennifer Ng and Marcelo A Fernández-Viña

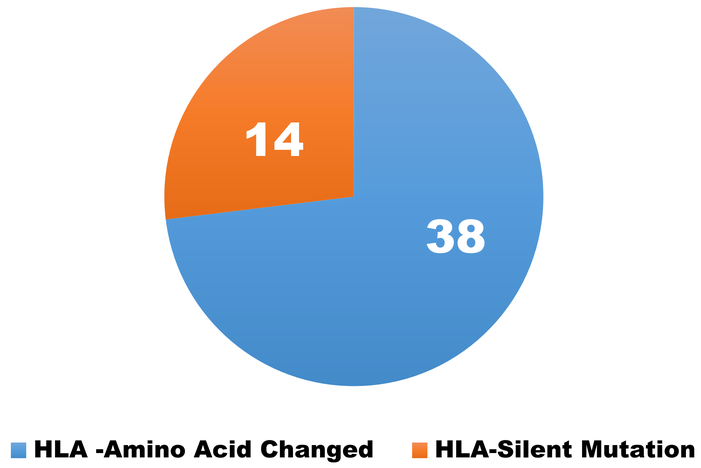
Abstract

Aim: The aim of this study was to characterize and report novel HLA class I alleles through a 6 HLA classical gene NGS strategy used for routine registry typing.

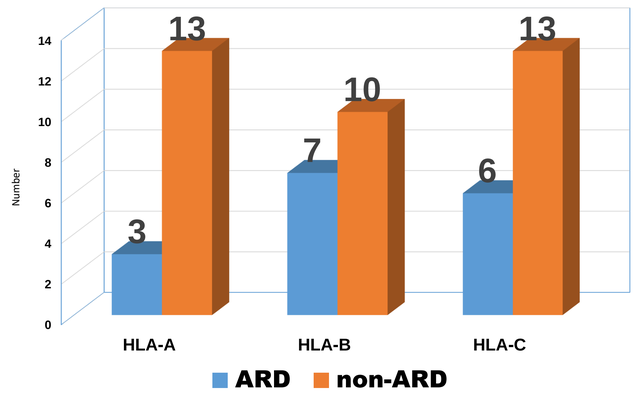
Methods: We performed long-range amplification of full-length HLA class I genes (A, B, and C) and partial HLA class II (DRB1, DQB1, and DPB1) of 6 HLA loci in single tube PCR using DNA from a buccal swab. The amplicons were fragmented using enzymatic digestion. A library was constructed using IDT xGen DNA library kit and the DNA fragments were tagged with one unique index combination. Libraries from 384-768 individuals were pooled followed by simultaneous sequencing in a single 2X300 cycle V3 or P1 Flow cell paired-end run using an Illumina Miseq or NextSeq. Fastq files were generated and analyzed with both Assign TruSight v2.0 and TypeStream Visual (TSV) version 3.0 software. Samples presenting substitutions in HLA Class I alleles suggesting the presence of novel alleles were re-analyzed by TSV software using IPD-IMGT/HLA 3.52.0 version as a reference sequence. NGSengine and Sequencher were used for novel allele isolation, annotation, and alignment. The novel allele sequences were submitted to GenBank (NCBI) and then to the IPD-IMGT/HLA database.

Results: A total of 52 novel HLA class I alleles are described from 54,300 registrants: 16 HLA-A alleles,17 HLA-B alleles, and 19 HLA-C alleles. All 52 novel alleles are single nucleotide variants (SNVs) when compared with the most homologous previously described allele. Thirty-eight (73%) of these SNVs resulted in a single amino acid substitution (SAAS), while 14 (27%) are silent substitutions. Eleven novel alleles (HLA-A:3, B:6, and C:2) differ from previously described alleles by SAAS at residues located in the antigen recognition domains (ARD)-encoding exons 2 or 3. One mutation introducing a premature termination codon resulting in a non-expressed allele (B*39:214N). Twenty-seven novel alleles (HLA-A:9, B:9, and C:9) presented SAAS in non-ARD-encoding exons. The fifty-two novel HLA class I alleles were officially named by the WHO Nomenclature Committee for Factors of the HLA System.

Total 52 Novel HLA ClassI Alleles from 54,300 Registrants



Novel ClassI ARD & non-ARD-encoding Exons Alleles Submitted to IMGT/HLA



Number of Novel Alleles Differ From Antigen Recognition Domains (ARD)Exon2-3

WHO Official Name	Races	New Allele-Locus	Allele1	Allele2	IMGT 3.52 (Closest Know Allele Nucleotide Change)	Exon/Codon/Protein Change	Synonymous (Silent)
A*01:455	Other White	A	A*01:196v	A*03:01:01	A*01:196, E3-892 T to C	Exon3/150 (GTC=>GCC)/V=>A	NO
A*03:486	Western European	A	A*02:01:01	A*03:01:01:86v	A*03:01:01:86, E2-383, G to T	Exon2/61 (GAC=>TAC)/D=>Y	NO
A*32:186	Afrcan American	A	A*02:02:01	A*32:01:01:01v	A*32:01:01:01, E2-386 C to G	Exon2/62 (CAG=>GAG)/Q=>E	NO
B*53:77	Black Caribbean	B	B*53:01:01:01v	B*57:03:01	B*53:01:01:01, E2-465 G to C	Exon2/89 (GAG=>CAG)/ E=>Q	NO
B*13:192	North American	B	B*13:02:01:01v	B*35:01:01	B*13:02:01:01, E3-764 G to A	Exon3/107 (GGG=>AGG)/G=>R	NO
B*39:213	North American	B	B*35:12:01	B*39:01:01:03v	B*39:01:01:03, E2-424 G to A	Exon2/75(CGA=>CAA)/R=>Q	NO
B*07:500	North American	B	B*07:02:01:01v	B*35:03:01	B*07:02:01:01, E3-891 C to A	Exon3/149(GCG=>GAG)/A=>E	NO
B*57:03:07	N/A	B	B*18:01:01	B*57:03:01:02v	B*57:03:01:02, E2-368 G to C	Exon2/56(GGG=>GGC)/G=>G	YES
B*39:214N	Afrcan American	B	B*07:06:01	B*39:10:01v	B*39:10:01, E3-875 C to T	Exon3/144(CAG=>TAG)/Q=> Stop	NO
B*07:501	Eastern European	B	B*07:02:01:01v	B*40:01:02	B*07:02:01:01, E3-986 C to T	Exon3/181(CGC=>TGC)/R-C	NO
C*14:161	N/A	C	C*02:02:02	C*14:02:01:01v	C*14:02:01:01, Exon 3-869, C to G	Exon3/141 (CAG=>GAG)/Q=>E	NO
C*05:01:78	North American	C	C*04:01:01	C*05:01:01:02v	C*05:01:01:02, E2-322 C to T	Exon2/40 (GCC=>GCT)/A=>A	YES
C*06:02:114	North American	C	C*06:02:01:01v	C*12:03	C*06:02:01:01, E3-862 G to T	Exon3/138 (ACG=>ACT)/T=>T	YES
C*03:669	Other White	C	C*03:04:25v	C*05:01:01	C*03:04:25, E2-256 G to A	Exon2/18 (GGG=>GGA)/G=G	YES
C*08:01:39	Chinese	C	C*01:02:01	C*08:01:01:01v	C*08:01:01:01, E3-763 C to A	Exon3/105 (CCC=>CCA/P=>P	YES
C*05:106:03	North American	C	C*05:18:06v	C*07:02:01	C*05:18:06, E3-914 C to T	Exon3/156 (CGG=>TGG)/R=>W	NO

Number of Novel Alleles Differ From non-ARD-encoding Exons

WHO Officially Name	Races	Exon/Codon/Protein Change	WHO Officially Name	Races	Exon/Codon/Protein Change	WHO Officially Name	Races	Exon/Codon/Protein Change
A*01:456	Greman	Exon4/203 TGC=>TGG/C=>W	B*27:273	Noth American	Exon4/ 193 CCC=>CAC/P=>H	C*04:521	Other White	Exon7/341 AAA=>GAA/K=E
A*02:01:219	N/A	Exon4/253 CAG=>CAA/Q=>Q	B*27:274	Northern European	Exon4/ 204 TGG=>TTG/W=>L	C*05:286	Northern European	Exon5/283 CCC=>GCC/P=>A
A*02:01:220	Noth American	Exon4/225 ACC=>ACT/T=>T	B*27:275	Other White	Exon4 /262 CAG=>CGG/Q=>R	C*07:1131	N/A	Exon4/ 220 GAT=>AAT/D=>N
A*01:324:02	Black Caribbean	Exon1 /-3 ACC=>TCC/T=>S	B*18:238	Other Southeast Asian	Exon4/188 CAT=>CTT/H=>L	C*03:668	South Asian	Exon6/319 GGG=>AGG/G=>R
A*03:474	N/A	Exon4/219 CGG=>CCG/R=>P	B*81:13	African	Exon6/324 GCG=>CTG/A=>V	C*12:407	German	Exon6/324 GCT=>GAT/A=>D
A*03:487	Irish	Exn6/340 AAA=>AAT/K=>N	B*44:02:86	Noth American	Exon5/311 AAG=>AAA/K=>K	C*03:04:107	German	Exon4/196 GAC=>GAT/D=>D
A*02:11:14	African American	Exon7/342 TGA=>TAA/Stop=>Stop	B*07:02:99	Noth American	Exon5/281 ACC=>ACT/T=>T	C*12:409	Western European	Exon4/191 CAC=>CAA/H=>Q
A*34:32	Noth American	Exon4/253 CAG=>GAG/Q=>E	B*49:85	Northern European	Exon7/327 GAC=>TAC/D=>Y	C*18:21	N/A	Exon5/290 GGC=>GTC/G=>V
A*03:488	Noth American	Exon4/ 195 TCT=>TTT/S=>F	B*56:98	African American	Exon7 /329 GCC=>GAC/A=>D	C*08:271	Black Caribbean	Exon1/-18 CGA=>CAA/R=>Q
A*31:230	American Indian	Exon4/203 TGC=>TGG/C=>W	B*44:400	American Indian	Exon4 267CCG=>CGG/P=>R	C*02:236	Western European	Exon1/-21 ATG=>GTG/M=>V
A*68:320	American Indian	Exon5/277 TCT=>CCT/S=>P				C*04:01:160	Black Caribbean	Exon4/191 CAC=>CAT/H=>H
A*30:227	Easter European	Exon4/ 183 ACC=>AAC/D=>E				C*17:01:20	African	Exon4/ 241 TTC=>TTT/F=>F
A*11:01:130	North American	Exon6/ 332 TCT=>TCC/S=>S				C*04:01:161	Western European	Exon4/ 234 AGG=>AGA/R=>R

Workflow Step by Step Protocol
gDNA from buccal swab

1:Amplification
HLA-6Loci+ABO
384 samples

NGS-001
NGS-002
NGS-003
NGS-004

2:Libraries

3:Sequencing

V3 Reagents / sequencing Output 15GB
➢ 384 samples for HLA-6Loci +ABO Genes

Flow cell-2x300bp/P1 /Sequencing Output 60GB
➢ 2X384 samples for HLA-6Loci +ABO Genes

Conclusions

The NGS protocol provides accurate and complete full-length, unambiguous high-resolution results for HLA Class I typing. In this study, we identified 52 novel HLA class I alleles using NGS for routine bone marrow donor registry typing from 2021 through 2023.

Acknowledgments

- The work supported by N00014-23-1-2449 and N00014-24-1-2264 grants from Office of Naval Research.
- We are grateful to all members of the Lab. for their dedicated daily work .