Original Article
Diagnostic Immunology

Ann Lab Med 2021;41:310-317
https://doi.org/10.3343/alm.2021.41.3.310
ISSN 2234-3806 • eISSN 2234-3814

ANNALS OF LABORATORY MEDICINE

Identification of 8-Digit HLA-A, -B, -C, and -DRB1 Allele and Haplotype Frequencies in Koreans Using the One Lambda AllType Next-Generation Sequencing Kit

Wonho Choe, M.D., Ph.D.[1], Jeong-Don Chae, M.D., Ph.D.[1], John Jeongseok Yang, M.D.[2], Sang-Hyun Hwang, M.D., Ph.D.[2], Sung-Eun Choi, Ph.D.[3], and Heung-Bum Oh, M.D., Ph.D.[2]

[1]Department of Laboratory Medicine, Eulji University School of Medicine, Seoul, Korea; [2]Department of Laboratory Medicine, University of Ulsan College of Medicine and Asan Medical Center, Seoul, Korea; [3]Department of Statistics, Dongguk University, Seoul, Korea

Background: Recent studies have successfully implemented next-generation sequencing (NGS) in HLA typing. We performed HLA NGS in a Korean population to estimate HLA-A, -B, -C, and -DRB1 allele and haplotype frequencies up to an 8-digit resolution, which might be useful for an extended application of HLA results.

Methods: A total of 128 samples collected from healthy unrelated Korean adults, previously subjected to Sanger sequencing for 6-digit HLA analysis, were used. NGS was performed for HLA-A, -B, -C, and -DRB1 using the AllType NGS kit (One Lambda, West Hills, CA, USA), Ion Torrent S5 platform (Thermo Fisher Scientific, Waltham, MA, USA), and Type Steam Visual NGS analysis software (One Lambda).

Results: Eight HLA alleles showed frequencies of ≥10% in the Korean population, namely, *A*24:02:01:01* (19.5%), *A*33:03:01* (15.6%), *A*02:01:01:01* (14.5%), *A*11:01:01:01* (13.3%), *B*15:01:01:01* (10.2%), *C*01:02:01* (19.9%), *C*03:04:01:02* (11.3%), and *DRB1*09:01:02* (10.2%). Nine previous 6-digit HLA alleles were further identified as two or more 8-digit HLA alleles. Of these, eight alleles (*A*24:02:01, B*35:01:01, B*40:01:02, B*55:02:01, B*58:01:01, C*03:02:02, C*07:02:01,* and *DRB1*07:01:01*) were identified as two 8-digit HLA alleles, and one allele (*B*51:01:01*) was identified as three 8-digit HLA alleles. The most frequent four-loci haplotype was *HLA-A*33:03:01-B*44:03:01:01-C*14:03-DRB1*13:02:01*.

Conclusions: We identified 8-digit HLA-A, -B, -C, and -DRB1 allele and haplotype frequencies in a healthy Korean population using NGS. These new data can be used as a representative Korean data for further disease-related HLA type analysis.

Key Words: Next-generation sequencing, Human leukocyte antigen (HLA), allele frequency, haplotype frequency, 8-digit HLA alleles, Korean population

Received: February 21, 2020
Revision received: May 13, 2020
Accepted: November 26, 2020

Corresponding author:
Heung-Bum Oh, M.D., Ph.D.
Department of Laboratory Medicine, University of Ulsan College of Medicine and Asan Medical Center, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea
Tel: +82-2-3010-4505
Fax: +82-2-478-0884
E-mail: hboh@amc.seoul.kr

INTRODUCTION

Human leukocyte antigen (HLA) is the most polymorphic gene among the known functional human genes [1]. Accurate identification of HLAs is important in solid organ and hematopoietic transplantations. Molecular HLA typing methods, including sequence-specific oligonucleotide probe, sequence-specific primers, and sequence-based typing methods have been widely used [2-4]. However, owing to the increasing number of HLA alleles, the problem of HLA typing ambiguity cannot be resolved, as these methods only analyze one or two exons where these variants are mainly found [5]. Recently, there have been several re-

ports on successful implementation of next-generation sequencing (NGS) in HLA typing [6-8]. NGS has been reported to reduce ambiguity largely arising from heterozygotes [9]. By sequencing both exons and introns, NGS can reduce HLA typing ambiguity arising from sequencing only the specified regions [10, 11].

For high-resolution 8-digit HLA typing, we used the One Lambda AllType NGS Amplification kit (One Lambda, West Hills, CA, USA) to analyze the NGS results of HLA-A, -B, -C, and -DRB1 in a Korean population and established updated allele and haplotype frequencies, which will be useful for more precise and extended applications of HLA types in clinical and research fields including disease-related HLA type analysis, drug-related adverse reaction analysis, immunologic interaction studies, and anthropological genetic studies. Additionally, we compared our NGS results with previous Sanger sequencing results to identify any discrepancies between the two methods.

MATERIALS AND METHODS

Samples collection

This retrospective study was performed at Asan Medical Center, Seoul, Korea, using the archival samples collected in 2003 from 128 genetically non-related, non-familial, healthy Korean adult volunteers >18 years old. Samples were obtained by collecting 10–20 mL of venous blood from each subject, and the concentration of the extracted DNA was adjusted to 40–100 ng/μL. These samples were previously analyzed for 6-digit HLA types using Sanger sequencing with Big Dye Terminator 3.1 (ABI Inc., Foster City, CA, USA) and an ABI 3730 DNA analyzer (ABI Inc.) [12]. Samples were stored in a -70°C deep freezer until our analysis. The Institutional Review Board (IRB) at Asan Medical Center approved this study (IRB No. S2018-2423-0001) and waived the need for informed consents from subjects.

NGS

The AllType NGS 11-Loci Amplification Kit (One Lambda) was used to amplify target DNA regions. The reagents were mixed according to the number of samples and pipetted into a PCR plate containing DNA. PCR was then performed for 11 loci by multiplex tagging rather than per locus. The primers for PCR were provided by the manufacturer, and the PCR cycling condition was as follows: 1 cycle at 94°C for 2 minutes, 22 cycles of 10 seconds at 98°C, 3 minutes at 69°C, 8 cycles of 98°C for 10

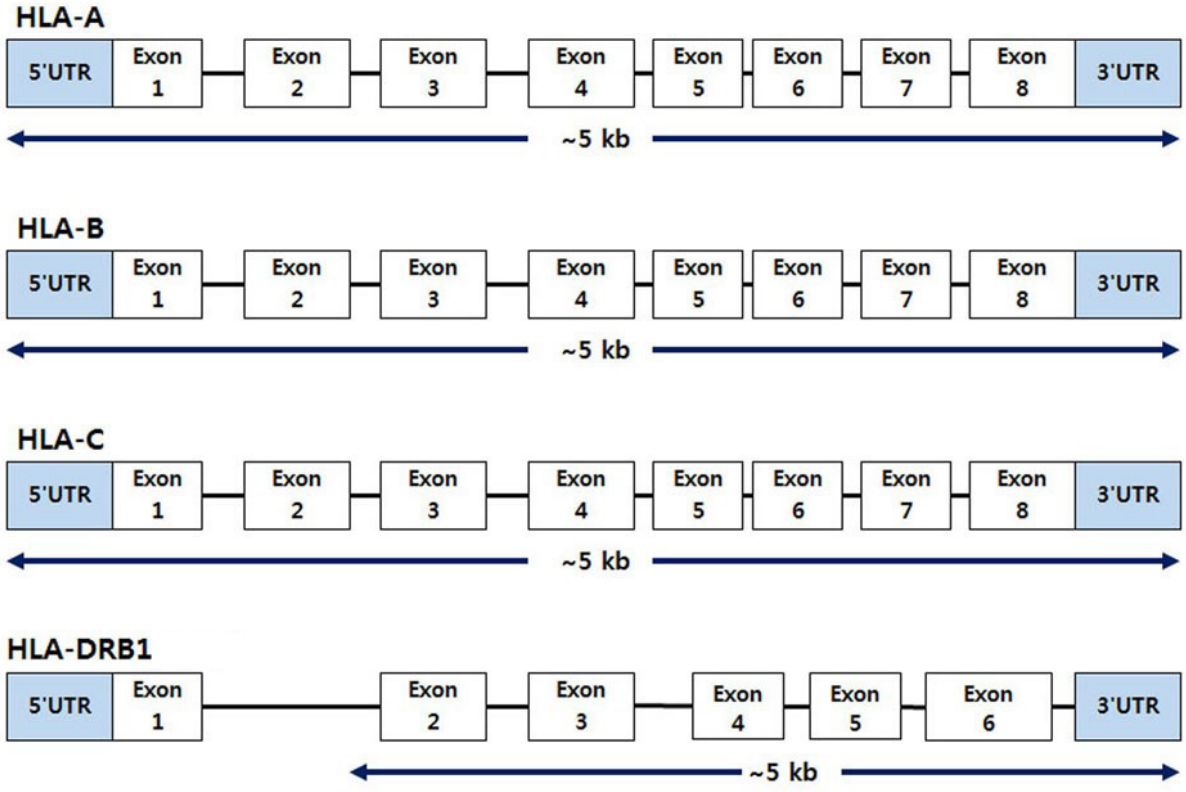

Fig. 1. Coverage of the HLA-A, -B, -C, and -DRB1 loci.
Abbreviations: HLA, human leukocyte antigen; UTR, untranslated region.

seconds, and 3 minutes at 60°C. The coverage for each HLA-A, -B, -C, and -DRB1 locus is shown in Fig. 1. DNA was purified by mixing Agencourt AMPure XP Beads (Beckman Coulter, Brea, CA, USA) with the amplicons. The purified amplicons were quantified using a Qubit 3.0 fluorometer and the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions and were diluted to ensure equal concentrations. The amplicons were fragmented using the Ion Shear Plus Reagent Kit (Thermo Fisher Scientific) and then ligated by mixing the Ion Xpress Barcode Adapters and Ion Plus Fragment Library Kit (Thermo Fisher Scientific) in each well. Next, size selection using Agencourt AMPure XP Beads (Beckman Coulter) was performed. The final products were obtained using PCR after adding the Ion Plus Fragment Library Kit to the amplicon plate. The primers for PCR were provided by the manufacturer, and the PCR cycling condition was as follows: 1 cycle at 96°C for 62 minutes, 8 cycles of 96°C for 16 seconds, followed by 15 seconds at 58°C and 1 minute at 70°C.

Using the NGS Calculator Excel file provided by One Lambda, DNA amount was calculated and pooled. Clonal amplification and sequencing were performed using Ion 520 & 530 ExT Kit–Chef and Ion 530 Chip Kit (Thermo Fisher Scientific). Reagents were accurately positioned according to the manufacturer's instructions; Ion Chef (Thermo Fisher Scientific) required 7 hours and Ion S5 XL (Thermo Fisher Scientific) required 6.5 hours. The obtained bam file was analyzed using Type Stream Visual (TSV) software (One Lambda). Analysis parameter settings are shown in Table 1, and the results were analyzed using the international ImMunoGeneTics information system (IMGT/HLA database) version 3.27 (the international ImMunoGeneTics information system, Montpellier, France).

Table 1. TSV analysis parameter configuration

Parameter setting	Coverage
Min read length	100
Max insertion	3
Max deletion	3
Max mismatch bases	5
Min base read depth	20
Min valid read	500
Noise cutoff value (%)	20
Min hetero allele balance	10
Max read for typing	300,000

Abbreviation: TSV, Type Stream Visual software.

HLA allele and haplotype frequency analysis

Allele frequency was calculated using the Maximum Likelihood Estimation method of the ALLELE procedure with SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). All loci were evaluated to determine whether they meet Hardy-Weinberg equilibrium using the same program. Haplotype frequency was calculated with the Expectation-Maximization Algorithm of the HAPLOTYPE procedure using the same program. Statistical significance level was $P<0.05$ (two-sided).

Additional exon analysis of discrepant NGS and 6-digit resolution Sanger sequencing results

The base position of samples showing discrepant results between NGS and previous Sanger sequencing HLA genotypes obtained by Jun, *et al.* [12] was confirmed using the BIOWITHUS SBT Analyzer (Biowithus Inc., Seoul, Korea). Next, we performed PCR using primer sets (Avita Plus Kit, Biowithus Inc., Seoul, Korea) harboring the necessary additional exon and Sanger sequencing according to previous methods to confirm the sequence difference in the base position of the additional exon [12].

RESULTS

HLA-A, -B, -C, and -DRB1 allele frequencies

NGS HLA genotyping revealed 24 types of HLA-A, 44 types of HLA-B, 25 types of HLA-C, and 28 types of HLA-DRB1 in our study group; allele frequencies are shown in Table 2. A few 8-digit alleles (2 HLA-A, 11 HLA-B, 4 HLA-C, and 2 HLA-DRB1), previously reported as 6-digit genotypes were observed for the first time.

HLA-A: previously typed 6-digit *HLA-A*24:02:01* was further typed as two 8-digit HLA types of *HLA-A*24:02:01:01* and *A*24:02:01:02L*.

HLA-B: four of the previously typed 6-digit HLA-B alleles were further typed as two 8-digit HLA types (*B*35:01:01* to *B*35:01:01:02, B*35:01:01:06; B*40:01:02* to *B*40:01:02:01, B*40:01:02:04; B*55:02:01* to *B*55:02:01:01, B*55:02:01:02; B*58:01:01* to *B*58:01:01:01, B*58:01:01:03*), and previously typed 6-digit *HLA-B*51:01:01* was further typed as three 8-digit HLA types of *HLA-B*51:01:01:01*, *B*51:01:01:03*, and *B*51:01:01:05*.

HLA-C: two of the previously typed 6-digit HLA-C alleles were further typed as two 8-digit HLA types (*C*03:02:02* to *C*03:02:02:01, C*03:02:02:03; C*07:02:01* to *C*07:02:01:01, C*07:02:01:03*).

HLA-DRB1: previously typed 6-digit *HLA-DRB1*07:01:01* was

https://doi.org/10.3343/alm.2021.41.3.310

Table 2. HLA-A, -B, -C, and -DRB1 allele frequencies in the Korean population (N=128)

*HLA-A**	N	Frequency (%)	*HLA-B**	N	Frequency (%)	*HLA-C**	N	Frequency (%)	*HLA-DRB1**	N	Frequency (%)
01:01:01:01	6	2.3	*07:02:01:01*	7	2.7	*01:02:01*	51	19.9	*01:01:01*	15	5.9
02:01:01:01	37	14.5	*07:05:01:01*	1	0.4	*01:03*	1	0.4	*03:01:01:01*	7	2.7
02:03:01	2	0.8	*08:01:01:02*	2	0.8	*02:02:02:01*	3	1.2	*04:03:01*	4	1.6
02:06:01:01	25	9.8	*13:01:01*	11	4.3	***03:02:02:01***	16	6.3	*04:04:01*	6	2.3
02:07:01	9	3.5	*13:02:01*	7	2.7	***03:02:02:03***	1	0.4	*04:05:01*	21	8.2
02:10	1	0.4	*14:01:01*	2	0.8	*03:03:01:01*	24	9.4	*04:06:01*	16	6.3
02:53N	1	0.4	*15:01:01:01*	26	10.2	*03:04:01:02*	29	11.3	*04:07:01*	1	0.4
03:01:01:01	3	1.2	*15:07:01*	2	0.8	*04:01:01:01*	23	9.0	*04:10:01*	2	0.8
11:01:01:01	34	13.3	*15:11:01*	1	0.4	*05:01:01:02*	3	1.2	*04:10:03*	2	0.8
11:02:01	4	1.6	*15:18:01:02*	1	0.4	*06:02:01:01*	8	3.1	***07:01:01:01***	13	5.1
11 : 20	1	0.4	*15:38:01*	2	0.8	***07:02:01:01***	14	5.5	***07:01:01:02***	2	0.8
24:02:01:01	50	19.5	*27:04:01*	1	0.4	***07:02:01:03***	7	2.7	*08:02:01*	5	2.0
24:02:01:02L	1	0.4	*27:05:02*	11	4.3	*07:04:01:01*	1	0.4	*08:03:02*	24	9.4
24:08	1	0.4	***35:01:01:02***	17	6.6	*07:06*	8	3.1	*09:01:02*	26	10.2
24 : 20	1	0.4	***35:01:01:06***	2	0.8	*07:18*	1	0.4	*11:01:01:01*	14	5.5
26:01:01:01	9	3.5	*37:01:01*	1	0.4	*08:01:01*	14	5.5	*12:01:01:03*	9	3.5
26:02:01	9	3.5	*38:02:01*	2	0.8	*08:02:01:02*	2	0.8	*12:02:01*	13	5.1
26:03:01	1	0.4	*39:01:01:03*	3	1.2	*08:03:01*	2	0.8	*12:10*	3	1.2
26 : 10	1	0.4	***40:01:02:01***	2	0.8	*08:22*	2	0.8	*13:01:01:01*	7	2.7
29:01:01:01	2	0.8	***40:01:02:04***	6	2.3	*12:02:02:01*	9	3.5	*13:02:01*	19	7.4
30:01:01	5	2.0	*40:02:01:01*	6	2.3	*12:03:01:01*	1	0.4	*14:03:01*	2	0.8
30:04:01	2	0.8	*40:03*	2	0.8	*14:02:01:01*	20	7.8	*14:05:01*	8	3.1
31:01:02:01	11	4.3	*40:06:01:01*	7	2.7	*14:03*	14	5.5	*14:07:01*	4	1.6
33:03:01	40	15.6	*44:02:01:01*	3	1.2	*15:02:01:03*	1	0.4	*14:12:01*	1	0.4
			44:03:01:01	14	5.5	*15:05:02*	1	0.4	*14:54:01*	4	1.6
			44:03:02	8	3.1				*15:01:01:03*	18	7.0
			46:01:01	14	5.5				*15:02:01:02*	7	2.7
			48:01:01	10	3.9				*16:02:01:02*	3	1.2
			51:01:01:01	17	6.6						
			51:01:01:03	5	2.0						
			51:01:01:05	1	0.4						
			51:02:01:02	1	0.4						
			52:01:01:02	8	3.1						
			54:01:01	18	7.0						
			55:01:01	1	0.4						
			55:02:01:01	2	0.8						
			55:02:01:02	5	2.0						
			56:01:01:04	1	0.4						
			57:01:01	1	0.4						
			58:01:01:01	1	0.4						
			58:01:01:03	13	5.1						
			59:01:01:01	5	2.0						
			67:01:01	4	1.6						
			67:01:02	2	0.8						

Bold letters: previously typed as 6-digit HLA alleles.
Abbreviation: HLA, human leukocyte antigen.

further typed as two 8-digit HLA types of *HLA-DRB1*07:01:01:01* and *DRB1*07:01:01:02*.

HLA haplotype frequencies

HLA haplotypes with frequencies >1% are shown in Tables 3 and 4. There were 25 HLA-A-B-C haplotypes with >1% frequency, and the following were the most frequent haplotypes: *HLA-A*33:03:01-B*44:03:01:01-C*14:03, HLA-A*11:01:01:01-B*15:01:01:01-C*04:01:01:01*, and *HLA-A*33:03:01-B*58:01:01:03-C*03:02:02:01*. There were 20 HLA-A-B-C-DRB1 haplotypes with >1% frequency, and the following were the most frequent haplotypes: *HLA-A*33:03:01-B*44:03:01:01-C*14:03-DRB1*13:02:01, HLA-A*11:01:01:01-B*15:01:01:01-C*04:01:01:01-DRB1*04:06:01*, and *HLA-A*02:01:01:01-B*13:01:01-C*03:04:01:02-DRB1*12:02:01*.

Table 3. *HLA-A-B-C* haplotype frequencies in the Korean population (N=128)

Haplotype	Frequency (%)	Standard error	95% confidence limit
33:03:01-44:03:01:01-14:03	4.8	0.013	0.022–0.075
11:01:01:01-15:01:01:01-04:01:01:01	4.7	0.013	0.021–0.072
33:03:01-58:01:01:03-03:02:02:01	4.2	0.012	0.017–0.066
24:02:01:01-52:01:01:02-12:02:02:01	3.1	0.011	0.010–0.053
24:02:01:01-51:01:01:01-14:02:01:01	2.9	0.011	0.009–0.050
33:03:01-44:03:02-07:06	2.7	0.010	0.007–0.047
02:07:01-46:01:01-01:02:01	2.3	0.009	0.005–0.042
24:02:01:01-54:01:01-01:02:01	2.1	0.009	0.004–0.039
02:06:01:01-51:01:01:01-14:02:01:01	2.1	0.009	0.004–0.039
02:01:01:01-27:05:02-01:02:01	2.0	0.009	0.003–0.037
24:02:01:01-35:01:01:02-03:03:01:01	2.0	0.009	0.003–0.037
30:01:01-13:02:01-06:02:01:01	2.0	0.009	0.003–0.037
02:01:01:01-13:01:01-03:04:01:02	1.9	0.009	0.002–0.036
02:06:01:01-54:01:01-01:02:01	1.8	0.008	0.001–0.034
24:02:01:01-07:02:01:01-07:02:01:03	1.6	0.008	0.000–0.031
33:03:01-51:01:01:03-03:02:02:01	1.6	0.008	0.000–0.031
31:01:02:01-35:01:01:02-03:03:01:01	1.5	0.008	0.000–0.031
02:01:01:01-15:01:01:01-04:01:01:01	1.2	0.007	0.000–0.025
02:01:01:01-40:06:01:01-08:01:01	1.2	0.007	0.000–0.025
02:01:01:01-48:01:01-08:01:01	1.2	0.007	0.000–0.025
02:01:01:01-54:01:01-01:02:01	1.2	0.007	0.000–0.025
11:01:01:01-67:01:01-07:02:01:01	1.2	0.007	0.000–0.025
24:02:01:01-46:01:01-01:02:01	1.2	0.007	0.000–0.025
24:02:01:01-59:01:01:01-01:02:01	1.2	0.007	0.000–0.025
24:02:01:01-40:02:01:01-03:04:01:02	1.1	0.007	0.000–0.024

Abbreviation: HLA, human leukocyte antigen.

Discrepancies between NGS and Sanger sequencing

A comparative analysis of the NGS and Sanger sequencing results revealed three discrepant HLA alleles (Table 5). HLA-A and -B alleles showed 100% concordance between the two methods, while there were two discrepant alleles for HLA-C and one discrepant allele for HLA-DRB1. Following additional SBT exon analyses of these samples, the SBT results were all corrected to the NGS results, yielding 100% concordance.

DISCUSSION

We used the One Lambda AllType 11-loci multiplex kit to perform HLA NGS and obtained 8-digit HLA typing results in the Korean population.

The AllType NGS Amplification kit uses multiplex individual tagging of 11 HLA loci. Instead of detecting an independent locus at the PCR stage, the kit uses independent index combinations to detect all 11 loci from each sample. Thus, AllType has the following advantages: during the PCR stage only one step is needed, it requires less DNA and fewer pipetting events than other commercial kits, amplicon pooling is not needed, and the library preparation step is 8 hrs shorter, including the PCR stage.

Eight of the HLA alleles showed frequencies of ≥10% in the Korean population. In previous studies examining HLA allele frequencies in a Korean population, In, *et al.* [14] reported five HLA alleles with frequencies of ≥10% [*A*24:02* (22.7%), *A*02:01* (17.6%), *A*33:03* (16.1%), *C*01:02* (17.8%), *C*03:03* (11.8%)] by SBT, while Chung, *et al.* [15] reported seven HLA alleles [*A*24:02* (22.9%), *A*02:01* (16.5%), *A*33:03* (15.4%), *B*51:01* (10.2%), *C*01:02* (17.4%), *C*03:03* (10.9%), *DRB1*09:01* (10.4%)] by high-resolution DNA typing. Our results were mostly consistent with previous results in 4-digit resolution comparisons [14, 15]. Interestingly, *B*51:01* (10.2%), reported by Chung, *et al.* [15], was divided into three 8-digit alleles, and *C*03:03:01:01* was nearly 10% (9.4%) in our study.

A comparison of some of the most frequent HLA alleles in Koreans with those in individuals from other countries revealed high frequencies of *HLA-A*24:02:01:01* (19.5%) in the Japanese (37.9%) and Hong Kong Chinese (14.7%) populations, whereas the frequencies were lower in US Caucasian (7.5%), British Caucasian (6.9%), and Saudi (8.5%) populations [16-20]. *HLA-B*15:01:01:01* (10.2%) was not as frequent in other populations (Hong Kong Chinese 2.9%, US Caucasian 6.0%, British Caucasian 5.9%, and Saudi 0.3%) except in the Japa-

https://doi.org/10.3343/alm.2021.41.3.310

Table 4. *HLA-A-B-C-DRB1* haplotype frequencies in the Korean population (N=128)

Haplotype	Frequency (%)	Standard error	95% confidence limit
33:03:01-44:03:01:01-14:03-13:02:01	3.9	0.012	0.015–0.063
11:01:01:01-15:01:01:01-04:01:01:01-04:06:01	3.1	0.011	0.010–0.053
02:01:01:01-13:01:01-03:04:01:02-12:02:01	2.0	0.009	0.003–0.037
02:01:01:01-27:05:02-01:02:01-01:01:01	2.0	0.009	0.003–0.037
30:01:01-13:02:01-06:02:01:01-07:01:01:01	2.0	0.009	0.003–0.037
33:03:01-44:03:02-07:06-07:01:01:01	2.0	0.009	0.003–0.037
24:02:01:01-07:02:01:01-07:02:01:03-01:01:01	1.6	0.008	0.000—0.031
33:03:01-51:01:01:03-03:02:02:01-13:01:01:01	1.6	0.008	0.000–0.031
33:03:01-58:01:01:03-03:02:02:01-03:01:01:01	1.6	0.008	0.000–0.031
33:03:01-58:01:01:03-03:02:02:01-13:02:01	1.6	0.008	0.000–0.031
02:06:01:01-54:01:01-01:02:01-04:05:01	1.6	0.008	0.000–0.031
24:02:01:01-52:01:01:02-12:02:02:01-15:02:01:02	1.6	0.008	0.000–0.031
02:01:01:01-15:01:01:01-04:01:01:01-04:06:01	1.2	0.007	0.000–0.025
02:07:01-46:01:01-01:02:01-08:03:02	1.2	0.007	0.000–0.025
11:01:01:01-15:01:01:01-04:01:01:01-11:01:01:01	1.2	0.007	0.000–0.025
24:02:01:01-46:01:01-01:02:01-08:03:02	1.2	0.007	0.000–0.025
24:02:01:01-51:01:01:01-14:02:01:01-04:05:01	1.2	0.007	0.000–0.025
24:02:01:01-51:01:01:01-14:02:01:01-09:01:02	1.2	0.007	0.000–0.025
24:02:01:01-54:01:01-01:02:01-04:05:01	1.2	0.007	0.000–0.025
26:02:01-40:06:01:01-08:01:01-09:01:02	1.2	0.007	0.000–0.025

Abbreviation: HLA, human leukocyte antigen.

Table 5. HLA allele discrepancies between NGS and 6-digit-resolution Sanger sequencing methods

Locus	NGS	SBT (before discrepancy analysis)	SBT (after discrepancy analysis)	Performed exon analysis
A	No discrepancy	No discrepancy		
B	No discrepancy	No discrepancy		
C	*C:15:05:02*	*C*15:05:01*	*C*15:05:02*	SBT Exon 1
	C:03:02:02:01	*C*03:02:01*	*C*03:02:02*	SBT Exon 5
DRB1	*DRB1*04:10:03*	*DRB1*04:10:01*	*DRB1*04:10:03*	SBT Exon 4

Abbreviations: HLA, Human leukocyte antigen; NGS, next-generation sequencing; SBT, sequence-based typing.

nese population (8.7%). *HLA-C*01:02:01* (19.9%) was also observed in significantly high frequencies in the Japanese (14.8%) and Hong Kong Chinese (19.0%) populations, whereas the frequencies were lower in the other populations (US Caucasian 2.1%, British Caucasian 4.0%, Saudi 1.0%). *HLA-DRB1*09:01:02* (10.2%) was also significantly high in the Japanese (12.4%) and Hong Kong Chinese (15.9%) populations, but was low in the other populations (US Caucasian 1.2%, British Caucasian 1.6%, and Saudi 0.3%).

Numerous intron variants have been predicted to exist; however, identifying them was not regarded as cost-effective relative to their clinical importance [21]. However, HLA typing solely of exons can provide incomplete information for two reasons. First, regulatory elements in promoter genes or introns need to be identified to determine the difference in HLA expression, as this is related to disease phenotype [22, 23]. Several studies have reported that promoter or intron identification resolved the problems of null to low expression of certain HLA alleles [24-27]. Second, without complete gene sequencing, it would be difficult to identify disease causes and drug interactions arising from the high linkage disequilibrium in the HLA region. Thus, the clinical use of NGS for HLA typing is expected to increase in the era of precision medicine.

A limitation of this study is that we were unable to include 8-digit HLA allele frequencies of the DQ or DP loci. We observed numerous ambiguities at the 8-digit resolution in these loci, which could not be resolved by this study alone. As the IMGT database is constantly being updated, further analysis would be needed to report 8-digit DQ and DP loci without ambiguities. This is the

first study on high-resolution 8-digit HLA typing using the One Lambda AllType NGS HLA Typing kit in the Korean population. We identified updated frequencies of HLA alleles and haplotypes by analyzing not only the exons but also the whole locus, including introns, 3′ untranslated regions (UTR), and 5′UTRs of HLA-A, -B, -C, and by resolving ambiguities for HLA-DRB1. Our data can be used as additional information in identifying cases where the same 4-digit or 6-digit HLA types show different characteristics, especially in studies involving disease-related HLA type analysis, drug-related adverse reaction analysis, immunologic interaction studies, and anthropological genetic studies in the Korean population.

ACKNOWLEDGEMENTS

None.

AUTHOR CONTRIBUTIONS

Choe W and Oh H-B planned the study; Choe W implemented the study; Chae J-D, Yang JJ, and Hwang S-H participated in the experiments and analysis; Choi S-E helped with statistical analysis; Choe W and Oh H-B wrote this manuscript.

CONFLICTS OF INTEREST

No potential conflicts of interest relevant to this study are reported.

RESEARCH FUNDING

This study was supported by EMBRI Grants 2018 SN0002 from the Eulji University, Korea.

ORCID

Wonho Choe	https://orcid.org/0000-0002-9685-4291
Jeong-Don Chae	https://orcid.org/0000-0002-2654-8299
John Jeongseok Yang	https://orcid.org/0000-0002-7739-1009
Sang-Hyun Hwang	https://orcid.org/0000-0003-3201-5728
Sung-Eun Choi	https://orcid.org/0000-0001-6707-6817
Heung-Bum Oh	https://orcid.org/0000-0001-6728-4640

REFERENCES

1. Borghans JAM, Keşmir C, *et al.* MHC diversity in individuals and populations. In: Flower D and Timmis J, eds. In silico immunology. Boston, MA: Springer, 2007:177-95.
2. Cao K, Chopek M, Fernández-Viña MA. High and intermediate resolution DNA typing systems for class I HLA-A, B, C genes by hybridization with sequence-specific oligonucleotide probes (SSOP). Rev Immunogenet 1999;1:177-208.
3. Olerup O and Zetterquist H. HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. Tissue Antigens 1992;39:225-35.
4. Saiki RK, Walsh PS, Levenson CH, Erlich HA. Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. Proc Natl Acad Sci U S A 1989;86:6230-4.
5. Lind C, Ferriola D, Mackiewicz K, Heron S, Rogers M, Slavich L, et al. Next-generation sequencing: the solution for high-resolution, unambiguous human leukocyte antigen typing. Hum Immunol 2010;71:1033-42.
6. Bentley G, Higuchi R, Hoglund B, Goodridge D, Sayer D, Trachtenberg EA, et al. High-resolution, high-throughput HLA genotyping by next-generation sequencing. Tissue Antigens 2009;74:393-403.
7. Erlich RL, Jia X, Anderson S, Banks E, Gao X, Carrington M, et al. Next-generation sequencing for HLA typing of class I loci. BMC Genomics 2011;12:42.
8. Holcomb CL, Höglund B, Anderson MW, Blake LA, Böhme I, Egholm M, et al. A multi-site study using high-resolution HLA genotyping by next generation sequencing. Tissue Antigens 2011;77:206-17.
9. Shiina T, Suzuki S, Ozaki Y, Taira H, Kikkawa E, Shigenari A, et al. Super high resolution for single molecule-sequence-based typing of classical HLA loci at the 8-digit level using next generation sequencers. Tissue Antigens 2012;80:305-16.
10. Hosomichi K, Jinam TA, Mitsunaga S, Nakaoka H, Inoue I. Phase-defined complete sequencing of the HLA genes by next-generation sequencing. BMC Genomics 2013;14:355.
11. Ehrenberg PK, Geretz A, Baldwin KM, Apps R, Polonis VR, Robb ML, et al. High-throughput multiplex HLA genotyping by next-generation sequencing using multi-locus individual tagging. BMC Genomics 2014;15:864.
12. Jun JH, Hwang K, Kim SK, Oh HB, Cho MC, Lee KJ. Estimation of the 6-digit level allele and haplotype frequencies of HLA-A, –B, and -C in Koreans using ambiguity-solving DNA typing. Tissue Antigens 2014;84:277-84.
13. IPD-IMGT/HLA Statistics. https://www.ebi.ac.uk/ipd/imgt/hla/stats.html (updated on Nov 2019).
14. In JW, Roh EY, Oh S, Shin S, Park KU, Song EY. Allele and haplotype frequencies of Human Leukocyte Antigen-A, -B, -C, -DRB1, and -DQB1 from sequence-based DNA typing data in Koreans. Ann Lab Med 2015;35:429-35.
15. Chung HY, Yoon JA, Han BY, Song EY, Park MH. Allelic and haplotypic diversity of HLA-A, -B, -C, and -DRB1 genes in Koreans defined by high-resolution DNA typing. Korean J Lab Med 2010;30:685-96.
16. Saito S, Ota S, Yamada E, Inoko H, Ota M. Allele frequencies and haplotypic associations defined by allelic DNA typing at HLA class I and class II loci in the Japanese population. Tissue Antigens 2000;56:522-9.
17. Kwok J, Guo M, Yang W, Lee CK, Ho J, Tang WH, et al. HLA-A, -B, -C, and -DRB1 genotyping and haplotype frequencies for a Hong Kong Chinese population of 7595 individuals. Hum Immunol 2016;77:1111-2.
18. Skibola CF, Akers NK, Conde L, Ladner M, Hawbecker SK, Cohen F, et al. Multi-locus HLA class I and II allele and haplotype associations with follicular lymphoma. Tissue Antigens 2012;79:279-86.
19. Alfirevic A, Gonzalez-Galarza F, Bell C, Martinsson K, Platt V, Bretland G,

et al. In silico analysis of HLA associations with drug-induced liver injury: use of a HLA-genotyped DNA archive from healthy volunteers. Genome Med 2012;4:51.
20. Hajeer AH, Al Balwi MA, Aytül Uyar F, Alhaidan Y, Alabdulrahman A, Al Abdulkareem I, et al. HLA-A, -B, -C, -DRB1 and -DQB1 allele and haplotype frequencies in Saudis using next generation sequencing technique. Tissue Antigens 2013;82:252-8.
21. Qu H and Fang X. A brief review on the Human Encyclopedia of DNA Elements (ENCODE) project. Genomics Proteomics Bioinformatics 2013; 11:135-41.
22. Cocco E, Meloni A, Murru MR, Corongiu D, Tranquilli S, Fadda E, et al. Vitamin D responsive elements within the HLA-DRB1 promoter region in Sardinian multiple sclerosis associated alleles. PLoS One 2012;7:e41678.
23. Thomas R, Apps R, Qi Y, Gao X, Male V, O'hUigin C, et al. HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. Nat Genet 2009;41:1290-4.
24. Dubois V, Tiercy JM, Labonne MP, Dormoy A, Gebuhrer L. A new HLA-B44 allele (B*44020102S) with a splicing mutation leading to a complete deletion of exon 5. Tissue Antigens 2004;63:173-80.
25. Elsner HA, Bernard G, Eiz-Vesper B, de Matteis M, Bernard A, Blasczyk R. Non-expression of HLA-A*2901102 N is caused by a nucleotide exchange in the mRNA splicing site at the beginning of intron 4. Tissue Antigens 2002;59:139-41.
26. Laforet M, Froelich N, Parissiadis A, Bausinger H, Pfeiffer B, Tongio MM. An intronic mutation responsible for a low level of expression of an HLA-A*24 allele. Tissue Antigens 1997;50:340-6.
27. Tamouza R, El Kassar N, Schaeffer V, Carbonnelle E, Tatari Z, Marzais F, et al. A novel HLA-B*39 allele (HLA-B*3916) due to a rare mutation causing cryptic splice site activation. Hum Immunol 2000;61:467-73.