

## RESEARCH ARTICLE

# Allele and haplotype frequencies of human leukocyte antigen-A, -B, -C, -DRB1, -DRB3/4/5, -DQA1, -DQB1, -DPA1, and -DPB1 by next generation sequencing-based typing in Koreans in South Korea

In-Cheol Baek<sup>1</sup>, Eun-Jeong Choi<sup>1</sup>, Dong-Hwan Shin<sup>1</sup>, Hyoung-Jae Kim<sup>1</sup>, Haeyoun Choi<sup>2</sup>, Tai-Gyu Kim<sup>1,2\*</sup>

**1** Hematopoietic Stem Cell Bank, College of Medicine, The Catholic University of Korea, Seoul, South Korea, **2** Department of Microbiology, College of Medicine, The Catholic University of Korea, Seoul, South Korea

\* [kimtg@catholic.ac.kr](mailto:kimtg@catholic.ac.kr)



## OPEN ACCESS

**Citation:** Baek I-C, Choi E-J, Shin D-H, Kim H-J, Choi H, Kim T-G (2021) Allele and haplotype frequencies of human leukocyte antigen-A, -B, -C, -DRB1, -DRB3/4/5, -DQA1, -DQB1, -DPA1, and -DPB1 by next generation sequencing-based typing in Koreans in South Korea. PLoS ONE 16(6): e0253619. <https://doi.org/10.1371/journal.pone.0253619>

**Editor:** Danilo G Augusto, University of California San Francisco, UNITED STATES

**Received:** January 30, 2021

**Accepted:** June 8, 2021

**Published:** June 21, 2021

**Copyright:** © 2021 Baek et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The results of this study were obtained from HLA genotype and NGS data for each individual. All relevant data are within the manuscript and its [Supporting information](#) files. Additionally, these data were submitted to NCBI's Sequence Read Archive (SRA) website (accession number: PRJNA721949).

**Funding:** This study was supported by a grant of the Korean Health Technology R&D Project,

## Abstract

Allele frequencies and haplotype frequencies of HLA-A, -B, -C, -DRB1, -DRB3/4/5, -DQA1, -DQB1, -DPA1, and -DPB1 have been rarely reported in South Koreans using unambiguous, phase-resolved next generation DNA sequencing. In this study, HLA typing of 11 loci in 173 healthy South Koreans were performed using next generation DNA sequencing with long-range PCR, TruSight<sup>®</sup> HLA v2 kit, Illumina MiSeqDx platform system, and Assign<sup>™</sup> for TruSight<sup>™</sup> HLA software. Haplotype frequencies were calculated using the PyPop software. Direct counting methods were used to investigate the association with DRB1 for samples with only one copy of a particular secondary DRB locus. We compared these allele types with the ambiguous allele combinations of the IPD-IMGT/HLA database. We identified 20, 40, 26, 31, 19, 16, 4, and 16 alleles of HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DQA1, HLA-DQB1, HLA-DPA1, and HLA-DPB1, respectively. The number of HLA-DRB3/4/5 alleles was 4, 5, and 3, respectively. The haplotype frequencies of most common haplotypes were as follows: A\*33:03:01-B\*44:03:01-C\*14:03-DRB1\*13:02:01-DQB1\*06:04:01-DPB1\*04:01:01 (2.89%), A\*33:03:01-B\*44:03:01-C\*14:03 (4.91%), DRB1\*08:03:02-DQA1\*01:03:01-DQB1\*06:01:01-DPA1\*02:02:02-DPB1\*05:01:01 (5.41%), DRB1\*04:05:01-DRB4\*01:03:01 (12.72%), DQA1\*01:03:01-DQB1\*06:01:01 (13.01%), and DPA1\*02:02:02-DPB1\*05:01:01 (30.83%). In samples with only one copy of a specific secondary DRB locus, we examined its association with DRB1. We, thus, resolved 10 allele ambiguities in HLA-B, -C (each exon 2 +3), -DRB1, -DQB1, -DQA1, and -DPB1 (each exon 2) of the IPD-IMGT/HLA database. Korean population was geographically close to Japanese and Han Chinese populations in the genetic distances by multidimensional scaling (MDS) plots. The information obtained by HLA typing of the 11 extended loci by next generation sequencing may be useful for more exact diagnostic tests on various transplantations and the genetic population relationship studies in South Koreans.

Ministry for Health & Welfare, Republic of Korea (HI14C3417) and the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2019R111A1A01064049).

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** HLA, Human leukocyte antigen; MDS, Multidimensional scaling; NGS, Next generation sequencing.

## Introduction

It is widely known that human leukocyte antigen (HLA) matching reduces morbidity and mortality in patients after hematopoietic stem cell transplantation (HSCT) [1]. Traditionally, Sanger sequencing has been the standard method for high-resolution HLA typing (PCR-SBT) [2], as the likelihood of accepting HLA mismatches for sensitized organ transplant candidates should be determined at high resolution, not at the antigen level [3]. However, SBT is unable to accurately phase heterozygous alleles and only provides limited sequencing information. Typically, SBT protocols cover only the exons 2, 3, and 4 for HLA class I genes and exons 2 and 3 for class II genes. This problem of ambiguity, the presence of two or more genotypes compatible with the same unphased sequence generated by SBT, is evidence for the complexity of the HLA region in the human genome, which contains more than 28,000 alleles identified in the IPD-IMGT/HLA database [4]. Time-consuming and costly additional testing is needed, and the vague list of delaying patient outcomes is still growing.

In recent years, efforts have been made to resolve this ambiguity and to reduce time and money, using next generation sequencing (NGS) techniques [5]. High-resolution and high-throughput HLA typing methods have been validated using amplicon-based next generation DNA sequencing [6, 7], greatly simplifying the workflow [8]. We reported on the distributions of HLA-A, -B (exon 2 and 3) and -DRB1 (exon 2) using amplicon-based NGS in a previous study [9]. Furthermore, HLA typing using NGS has subsequently been performed by long-range amplification, sequencing platforms, and analysis algorithms [10–12]. These efforts decreased the ambiguity of HLA alleles for all 11 HLA loci [3, 13–19].

The distribution of HLA alleles differed significantly between ethnic groups, and specific alleles and haplotypes are characteristic of each ethnic group [9, 14, 16, 17, 20–32]. Transplant and disease-related studies, such as organ or hematopoietic stem cell transplantation, require data on the distribution of HLA alleles and haplotypes in each ethnic group [3, 33–35]. More extended HLA typing region is needed to improve the success rate of unrelated hematopoietic stem cell transplantation [36]. In this study, we analyzed the allele and haplotype frequencies of 11 extended HLA loci using long-range PCR.

## Methods

### Sample preparation

DNA was collected from the blood of 173 genetically unrelated healthy Korean adults who were mainly consistent with students and staff from the Medical College of the Catholic University of Korea in Seoul in South Korea. The Korean people are originally derived from one ethnic group, Mongolian who migrated to the Korean peninsula about five thousand years ago, and preserved the unique physico/anthropological characteristics. South Korea is rapidly developed and urbanization is accelerating by fast-paced industrialization after the Korean War between 1950 and 1953. These phenomena make the huge population influx to Seoul from the different rural areas. In this way, Seoul became a metropolitan and features of special geographic origins have been diluted. Moreover, genetic homogeneity was revealed on the Korean peninsula level, except Jeju [37]. The population of the Seoul Capital Area (Seoul, Incheon, and Gyeonggi) amounted to 25.89 million persons in 2019, which accounted for 50.0% of the total population of South Korea. The 173 Korean adult from medical school, the subject of this study, has been used as a group representing Koreans because it includes diverse Koreans locally (<http://kostat.go.kr/portal/eng/pressReleases/8/7/index.board?bmode=download&bSeq=&aSeq=386088&ord=1>). Genomic DNA was freshly extracted from 4 mL of peripheral blood mixed with ethylenediaminetetraacetic acid (EDTA) using the TIANamp

Genomic DNA Extraction Kits (Tiangen Biotech Corporation, Beijing, China), according to the manufacturer's instructions. Extracted DNA was adjusted to a concentration of 50 ng/μL in Tris-ethylenediaminetetraacetic acid (TE) buffer [35], and DNA was quantified using a QuBit fluorometer (Life Technologies, Carlsbad, CA). After quantification, sample DNA was diluted to 10 ng/mL. All subjects provided informed consent to participate in genetic studies. Also, written informed consent was obtained from each participant. This research protocol was carried out in accordance with the Declaration of Helsinki with the approval of the Catholic University Institutional Review Board (IRB) (IRB number: MC13SISI0126).

## HLA gene amplification

Freshly extracted genomic DNA was used to amplify each HLA locus according to the manufacturer's instructions: TruSight<sup>®</sup> HLA v2 Sequencing Panel. PCR amplicons were confirmed using 2% agarose gel electrophoresis prior to preparing the NGS libraries. Twenty-four samples (192 HLA loci) were run in a single NGS experiment. The samples were of sufficient quality to ensure library preparation, data quality, and analysis, as well as the correct HLA typing by NGS.

## Genotyping of HLA alleles by Assign<sup>™</sup> for TruSight<sup>™</sup> HLA software

Data analysis was performed using Assign<sup>™</sup> for TruSight<sup>™</sup> HLA software (version 2.1.0.943, Illumina Inc., San Diego, CA). Sequencing data was interpreted on using the IPD-IMGT/HLA database 3.42.0 [4]. We compared the genotypes obtained with next-generation sequencing with the previous results acquired with Sanger sequencing, allowing the estimative of the NGS accuracy [35]. The Assign<sup>™</sup> for TruSight<sup>™</sup> HLA software was designed for the genotyping of HLA-A, -B, -C, -DRB1, -DRB3/4/5, -DQA1, -DQB1, -DPA1, and -DPB1 from the fastq sequencing reads provided by the Illumina MiSeqDx platform with the Illumina Pipeline software. The methods implemented within the Assign<sup>™</sup> for TruSight<sup>™</sup> HLA software utilizes a large number of reads per sequence.

## Statistical analysis

**Allele and haplotype frequencies.** The allele frequencies were determined using a direct counting method. Haplotypes were calculated using the iterative Expectation-Maximization (EM) algorithm [38, 39] implemented by the software PyPop-Win32-0.7.0 ([www.pypop.org](http://www.pypop.org)) for the HLA-A, -B, -C, -DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 (S1–S5 Tables) [40]. The list of files that make up the PyPop-Win32-0.7.0 software contains a minimal configuration file called “sample.ini”. The presence of [Emhaplofreq] section in the file enables haplotype estimation. In ‘lociToEstHaplo’ option in the section you can list the multi-locus haplotypes for which you wish the program to estimate (S6 Table).

Genotyping data were further investigated for predicted haplotypes using the PyPop-Win32-0.7.0 software. This analysis was not performed on all 11 loci of the HLA gene due to the limitation in the number of samples in this study. The association of DRB1 with DRB3/4/5 was analyzed by direct counting.

### Association of DRB1 with secondary DR loci (HLA-DRB3, -DRB4, and -DRB5).

Although not all samples present a secondary DR loci (HLA-DRB3, -DRB4, and -DRB5), direct counting methods were used to investigate association with DRB1 for samples with only one copy of a particular secondary DRB locus. The association of DRB1 with DRB3/4/5 loci were also investigated by comparison with previously reported DRB structures [41]. We referred to a two-field nomenclature found in volunteers from the US registry with European

backgrounds and a prior population study of the Netherlands consistent with the more recent high-resolution haplotype assignments [14, 22, 23, 25, 26, 34, 42].

**Hardy-Weinberg equilibrium.** Eight loci (HLA-A, -B, -C, -DRB1, -DQA1, -DQB1, -DPA1, and -DPB1) were tested for Hardy-Weinberg equilibrium using SNPStats (<https://www.snpstats.net/start.htm>). The haplotypes including HLA-DRB1 and one of the non-classic HLA-DRB loci (HLA-DRB3, -DRB4, or -DRB5) were considered one locus for the HW test. If the P-value was less than 0.05, the locus was considered unbalanced.

**Analysis of HLA alleles resolved on ambiguous allele combinations of the IPD-IMGT/HLA database (release version 3.42.0).** HLA alleles obtained from NGS by the TruSight<sup>®</sup> HLA v2 kit were compared with the IPD/IMGT/HLA ambiguous allele combinations (exon 2+3 in HLA-A, -B, and -C and exon 2 in HLA-DRB1/3/4/5, -DQA1, -DQB1, -DPA1, and -DPB1) (S7 Table) [4].

**Multidimensional scaling (MDS) analysis.** MDS analysis (for two dimensions, based on the Euclidian distance matrix computed using allele frequencies) were performed with ALSCAL procedure using SPSS 27 software package (SPSS Inc., Chicago, IL, USA). ALSCAL procedure options were set to 'interval' for measurement level, 'Symmetric' for data matrix shape, 'Dissimilarity' for type, and 'Leave Tied' for Approach to Ties. For HLA-A, -B, -DRB1, -DQA1, and -DQB1, we analyzed using HLA data reported by Johansson et al [43]. For HLA-C, -DPA1, and -DPB1, HLA data collected from a worldwide selection of populations in Allele frequencies.net database (<http://www.allelefreqencies.net>). The Japanese and Han Chinese populations were included because it was genetically close to this population in all analyzes, except for HLA-DPA1. Additionally, HLA data of deferent geographic regions was used in this study—HLA-C (13 populations): Japanese, Han Chinese, Australian, Southeast Asian, South Asian, West Asian, Oceanian, European, South American, North American, North African, and Sub-Saharan African; HLA-DPA1 (10 populations): Japanese, Southeast Asian, South Asian, Oceanian, European, Brazilian, South American, North American, and Sub-Saharan African; HLA-DPB1 (13 populations): Japanese, Han Chinese, Australian, Southeast Asian, South Asian, West Asian, Oceanian, European, South American, North American, North African, and Sub-Saharan African. The population groups were determined by 'Region' of HLA-Allele Frequency Search-Classical section in Allele frequencies.net. Additionally, we selected to each country for 'Country' and '2 field' for 'Level of resolution' (e.g. Han Chinese population was selected to 'North-East Asia' for 'Region', 'China' for 'Country', and '2 field' for 'Level of resolution'). HLA allele frequency of each population group was recalculated by following formula of Allele frequencies.net:

'Allele Frequency: Total number of copies of the allele in the population sample (Alleles / 2n) in decimal format.'

The genetic distances of 8 HLA loci were analyzed by the 2nd field allele frequencies of this study and the estimated allele frequencies of the population groups.

## Results

### Hardy-Weinberg equilibrium

Hardy-Weinberg equilibrium tests were performed on the eight HLA loci. The statistical P value of observed, expected homozygotes and heterozygotes are given in S8 Table. The results showed that the P values at the loci were all more than 0.05. There were no detectable deviations at each of the eight loci from the Hardy-Weinberg equilibrium. P values greater than 0.05 indicate that the population is consistent with the Hardy-Weinberg equilibrium [44].

### Allele frequencies of HLA-A, -B, -C, -DRB1, -DQA1, -DQB1, -DPA1, and -DPB1

Allele frequencies of HLA-A, -B, -C, -DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 are listed in Table 1 (>1%, Complete tables were showed in S9 and S10 Tables). For the HLA-A locus, the A2 group accounted for 35.5% of all the HLA-A alleles. In total, twenty distinct alleles were

**Table 1. Allele frequencies of HLA-A, -B, -C, -DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 in South Koreans (N = 173, >1%).**

HLA alleles	2n	%	HLA alleles	2n	%	HLA alleles	2n	%
A*02:01:01	79	22.83	C*03:02:02	19	5.49	DQA1*03:01:01	39	11.27
A*02:06:01	22	6.36	C*03:03:01	46	13.29	DQA1*03:02:01	19	5.49
A*02:07:01	18	5.20	C*03:04:01	32	9.25	DQA1*03:03:01	37	10.69
A*03:01:01	4	1.16	C*04:01:01	21	6.07	DQA1*05:01:01	11	3.18
A*11:01:01	45	13.01	C*05:01:01	4	1.16	DQA1*05:03:01	5	1.45
A*24:02:01	55	15.90	C*06:02:01	17	4.91	DQA1*05:05:01	16	4.62
A*26:01:01	14	4.05	C*07:02:01	39	11.27	DQA1*05:06:01	4	1.16
A*26:02:01	8	2.31	C*07:04:01	4	1.16	DQA1*05:08	7	2.02
A*30:01:01	10	2.89	C*07:06:01	8	2.31	DQA1*06:01:01	18	5.20
A*30:04:01	5	1.45	C*08:01:01	12	3.47			
A*31:01:02	12	3.47	C*08:03:01	6	1.73	DQB1*02:01:01	8	2.31
A*33:03:01	56	16.18	C*12:02:02	5	1.45	DQB1*02:02:01	26	7.51
			C*14:02:01	11	3.18	DQB1*03:01:01	57	16.47
B*07:02:01	17	4.91	C*14:03:01	19	5.49	DQB1*03:02:01	41	11.85
B*13:01:01	11	3.18	C*15:02:01	9	2.60	DQB1*03:03:02	21	6.07
B*13:02:01	17	4.91			DQB1*04:01:01	29	8.38	
B*15:01:01	38	10.98	DRB1*01:01:01	20	5.78	DQB1*04:02:01	6	1.73
B*15:07:01	4	1.16	DRB1*03:01:01	8	2.31	DQB1*05:01:01	23	6.65
B*15:11:01	5	1.45	DRB1*04:01:01	4	1.16	DQB1*05:02:01	12	3.47
B*15:18:01	4	1.16	DRB1*04:03:01	12	3.47	DQB1*05:03:01	18	5.20
B*27:05:02	11	3.18	DRB1*04:05:01	29	8.38	DQB1*06:01:01	46	13.29
B*35:01:01	18	5.20	DRB1*04:06:01	18	5.20	DQB1*06:02:01	22	6.36
B*38:02:01	5	1.45	DRB1*07:01:01	27	7.80	DQB1*06:03:01	4	1.16
B*40:01:02	11	3.18	DRB1*08:02:01	9	2.60	DQB1*06:04:01	21	6.07
B*40:02:01	15	4.34	DRB1*08:03:02	40	11.56	DQB1*06:09:01	11	3.18
B*40:06:01	7	2.02	DRB1*09:01:02	19	5.49			
B*44:02:01	4	1.16	DRB1*11:01:01	14	4.05	DPA1*01:03:01	152	43.93
B*44:03:01	20	5.78	DRB1*12:01:01	17	4.91	DPA1*02:01:01	54	15.61
B*44:03:02	9	2.60	DRB1*12:02:01	12	3.47	DPA1*02:02:02	139	40.17
B*46:01:01	21	6.07	DRB1*13:01:01	4	1.16			
B*48:01:01	13	3.76	DRB1*13:02:01	32	9.25	DPB1*02:01:02	87	25.14
B*51:01:01	17	4.91	DRB1*14:05:01	13	3.76	DPB1*02:02:01	18	5.20
B*51:02:01	4	1.16	DRB1*14:54:01	12	3.47	DPB1*03:01:01	16	4.62
B*52:01:01	5	1.45	DRB1*15:01:01	23	6.65	DPB1*04:01:01	21	6.07
B*54:01:01	27	7.80	DRB1*15:02:01	12	3.47	DPB1*04:02:01	30	8.67
B*55:02:01	6	1.73			DPB1*05:01:01	118	34.10	
B*58:01:01	20	5.78	DQA1*01:01:01	23	6.65	DPB1*09:01:01	11	3.18
B*59:01:01	10	2.89	DQA1*01:02:01	55	15.90	DPB1*13:01:01	20	5.78
			DQA1*01:03:01	49	14.16	DPB1*14:01:01	8	2.31
C*01:02:01	72	20.81	DQA1*01:04:01	27	7.80	DPB1*17:01:01	11	3.18
C*02:02:02	4	1.16	DQA1*02:01:01	27	7.80			

<https://doi.org/10.1371/journal.pone.0253619.t001>

identified for the HLA-A locus. Of these, A\*02:01:01 was the most common allele, followed by A\*33:03:01, A\*24:02:01, A\*11:01:01, A\*02:06:01, and A\*02:07:01 (>5%). The HLA-B locus showed the greatest diversity, with forty identified alleles. Allele frequencies of B\*15:01:01, B\*54:01:01, B\*46:01:01, B\*44:03:01, B\*58:01:01, and B\*35:01:01 were over 5%. In HLA-C, twenty six alleles were identified; C\*01:02:01 was the most common allele, followed by C\*03:03:01, C\*07:02:01, C\*03:04:01, C\*04:01:01, C\*03:02:02, and C\*14:03:01 (>5%). The allele frequencies of C\*07:01:02 and C\*07:06:01 were 0.29% and 2.31%, respectively. In the HLA-DRB1 locus, 31 alleles were identified. DRB1\*08:03:02 was the most common allele, followed by DRB1\*13:02:01, DRB1\*04:05:01, DRB1\*07:01:01, DRB1\*15:01:01, DRB1\*01:01:01, DRB1\*09:01:02, and DRB1\*04:06:01 (>5%). Interestingly, there were 19 alleles in HLA-DQA1 locus, making it very diverse. DQA1\*01:02:01 were the most common allele, followed by DQA1\*01:03:01, DQA1\*03:01:01, DQA1\*03:03:01, DQA1\*01:04:01, DQA1\*02:01:01, DQA1\*01:01:01, DQA1\*03:02:01, and DQA1\*06:01:01 (>5%). Sixteen alleles for the HLA-DQB1 were identified. The sum of allele frequencies of DQB1\*03:01:01 and DQB1\*03:03:02 were 22.54%. This included three DQ6 alleles (DQB1\*06:01:01, DQB1\*06:02:01, and DQB1\*06:04:01), DQB1\*03:02:01, DQB1\*04:01:01, DQB1\*02:02:01, DQB1\*05:01:01, and DQB1\*05:03:01 (>5%). The DQB1\*02:01:01 and DQB1\*02:02:01 were 2.31% and 7.51%, respectively. Four alleles for the HLA-DPA1 locus were identified, consisting of DPA1\*01:03:01, DPA1\*02:02:02, DPA1\*02:01:01, and DPA1\*01:04 in descending order frequency. In the HLA-DPB1 locus, 16 alleles were identified. DPB1\*05:01:01 and DPB1\*02:01:02 were the most common, followed by DPB1\*04:02:01, DPB1\*04:01:01, DPB1\*13:01:01, and DPB1\*02:02:01 (>5%). The number of HLA-DRB3/4/5 alleles was 4, 5, and 3, respectively. Allele frequencies of DRB3\*02:02:01, DRB3\*01:01:02, DRB3\*03:01:01, DRB3\*03:01:03, DRB4\*01:03:01, DRB4\*01:03:02, DRB4\*01:02:01, DRB4\*01:01:01, DRB5\*01:01:01, DRB5\*01:02:01, and DRB5\*02:02:01 were 32.12%, 16.36%, 16.36%, 6.67%, 51.52%, 8.48%, 2.42%, 0.61%, 0.61%, 13.94%, 6.67%, and 0.61%, respectively.

## Haplotype analysis

The 6-locus haplotypes of HLA-A, -B, -C, -DRB1, -DQB1, and -DPB1 are listed in [Table 2](#).

Haplotype frequencies of 14 HLA-A, -B, -C, -DRB1, -DQB1, and -DPB1 were >1%.

A\*33:03:01-B\*44:03:01-C\*14:03:01-DRB1\*13:02:01-DQB1\*06:04:01-DPB1\*04:01:01 (2.89%) was the most common haplotype.

The haplotype frequencies of HLA class I genes, HLA-A, -B, and -C, with values >1% are displayed in [S11 Table](#) except for overlapping parts of [Table 2](#). The remaining eight kinds were A\*11:01:01-B\*54:01:01-C\*01:02:01 (2.49%), A\*24:02:01-B\*59:01:01-C\*01:02:01 (2.31%), A\*02:01:01-B\*40:02:01-C\*03:04:01 (2.02%), A\*02:01:01-B\*13:01:01-C\*03:04:01 (1.45%), A\*02:01:01-B\*15:11:01-C\*03:03:01 (1.45%), A\*02:01:01-B\*27:05:02-C\*01:02:01 (1.44%), A\*11:01:01-B\*35:01:01-C\*03:03:01 (1.13%), and A\*02:06:01-B\*46:01:01-C\*01:02:01 (1.12%). The sum of the top ten haplotype frequencies was approximately 32.15%.

The frequencies of the 24 haplotypes of HLA class II genes, HLA-DRB1, -DQA1, -DQB1, -DPA1 and -DPB1, are shown in [S12 Table](#) (>1%). DRB1\*08:03:02-DQA1\*01:03:01-DQB1\*06:01:01-DPA1\*02:02:02-DPB1\*05:01:01 (5.41%) was the most common allele. The haplotype frequencies of 2-locus haplotypes of HLA class II genes are shown in [S13 Table](#) (>1%). 19 out of the 25 observed haplotypes including HLA-DRB1 and one of the non-classic HLA-DRB loci (HLA-DRB3, -DRB4, or -DRB5) presented frequencies over 1%. In samples with only one copy of a specific secondary DRB locus, we examined the association with DRB1. The DRB1\*04:05:01-DRB4\*01:03:01 haplotype had the highest frequency (22/173, 12.72%), followed by DRB1\*13:02:01-DRB3\*03:01:01 (16/173, 9.25%), DRB1\*04:06:01-DRB4\*01:03:01 (14/173, 8.09%), DRB1\*07:01:01-DRB4\*01:03:01 (11/173, 6.36%), DRB1\*15:01:01-DRB5\*01:01:01

**Table 2. Haplotype frequencies of HLA-A, -B, -C, -DRB1, -DQB1, and -DPB1 (>1%).**

HLA haplotypes	HF (%)
A*33:03:01-B*44:03:01-C*14:03-DRB1*13:02:01-DQB1*06:04:01-DPB1*04:01:01	2.89
A*02:07:01-B*46:01:01-C*01:02:01-DRB1*08:03:02-DQB1*06:01:01-DPB1*05:01:01	2.60
A*33:03:01-B*44:03:02-C*07:06-DRB1*07:01:01-DQB1*02:02:01-DPB1*13:01:01	1.73
A*30:01:01-B*13:02:01-C*06:02:01-DRB1*07:01:01-DQB1*02:02:01-DPB1*17:01:01	1.73
A*02:01:01-B*54:01:01-C*01:02:01-DRB1*04:05:01-DQB1*04:01:01-DPB1*05:01:01	1.73
A*24:02:01-B*07:02:01-C*07:02:01-DRB1*01:01:01-DQB1*05:01:01-DPB1*04:02:01	1.73
A*33:03:01-B*44:03:01-C*14:03-DRB1*13:02:01-DQB1*06:04:01-DPB1*02:01:02	1.45
A*02:01:01-B*15:01:01-C*03:03:01-DRB1*12:01:01-DQB1*03:01:01-DPB1*02:01:02	1.45
A*02:01:01-B*15:01:01-C*04:01:01-DRB1*04:06:01-DQB1*03:02:01-DPB1*05:01:01	1.45
A*24:02:01-B*07:02:01-C*07:02:01-DRB1*01:01:01-DQB1*05:01:01-DPB1*05:01:01	1.16
A*24:02:01-B*54:01:01-C*01:02:01-DRB1*04:05:01-DQB1*04:01:01-DPB1*05:01:01	1.16
A*11:01:01-B*15:01:01-C*04:01:01-DRB1*04:06:01-DQB1*03:02:01-DPB1*02:01:02	1.16
A*33:03:01-B*58:01:01-C*03:02:02-DRB1*13:02:01-DQB1*06:09:01-DPB1*05:01:01	1.16
A*24:02:01-B*52:01:01-C*12:02:02-DRB1*15:02:01-DQB1*06:01:01-DPB1*09:01:01	1.16

HF, haplotype frequency

<https://doi.org/10.1371/journal.pone.0253619.t002>

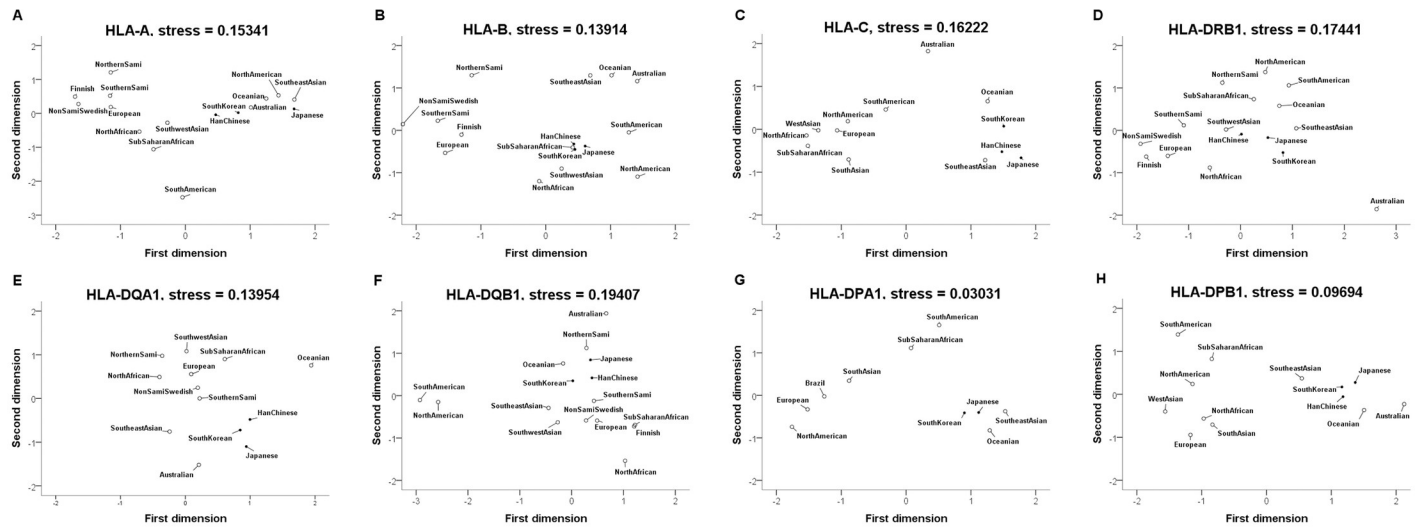
(11/173, 6.36%), DRB1\*09:01:02-DRB4\*01:03:02 (10/173, 5.78%), DRB1\*15:02:01-DRB5\*01:02 (10/173, 5.78%), DRB1\*12:01:01-DRB3\*02:02:01 (9/173, 5.20%), and DRB1\*14:54:01-DRB3\*02:02:01 (9/173, 5.20%) (>5%). Twenty DQA1~DQB1 and fifteen DPA1~DPB1 haplotypes presented frequencies over 1%, wherein DQA1\*01:03:01-DQB1\*06:01:01 (13.01%) and DPA1\*02:02:02-DPB1\*05:01:01 (30.83%) were the most common haplotypes, respectively.

### Expected PCR-SBT ambiguities (described in the IPD-IMGT/HLA database) solved by the NGS assay

Analysis of the data obtained from NGS by the TruSight<sup>®</sup> HLA v2 kit and Assign<sup>™</sup> for TruSight<sup>™</sup> HLA software resolved 10 allele ambiguities in HLA-B, -C (each exon 2+3), -DRB1, -DQB1, -DQA1, and -DPB1 (each exon 2) of the IPD-IMGT/HLA database (Release version 3.42.0) (S14 Table). The unresolved ambiguities observed using the PCR-sequence based typing method were in the IPD-IMGT/HLA database.

### Genetic distances by MDS plots

According to (1) the 2nd field allele frequencies of this study, (2) the estimated allele frequencies of the population groups, and (3) the allele frequencies reported by Johansson et al [43], allele frequencies of HLA-A, -B, -C, -DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 in 10–16 populations were collected to analyze Euclidian genetic distances by MDS plots (S15–S22 Tables). Since there is no data on HLA-DPA1 in some ethnicities, genetic distances were calculated only for 10 populations. The MDS analysis resulted in stress values of 0.15341, 0.13914, 0.16222, 0.17441, 0.13954, 0.19407, 0.03031, and 0.09694 for HLA-A, -B, -C, -DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 data, respectively. Interestingly, the stress values in HLA-DPA1 and -DPB1 were Excellent and Perfect by Kruskal's stress formula 1, respectively. In all analyzed genes, Korean, Japanese, and Han Chinese populations included in Northeast Asian populations were found to be close ethnicity, except for HLA-DPA1. In the analysis of HLA-DPB1, Northeast Asian was relatively close to Oceanian, Southeast Asian, and Australian populations compared to other ethnicities (Fig 1).



**Fig 1. Euclidian distances by two dimensional MDS plots for allele frequencies of 8 HLA loci.** (A) HLA-A, (B) HLA-B, (C) HLA-C, (D) HLA-DRB1, (E) HLA-DQA1, (F) HLA-DQB1, (G) HLA-DPA1, (H) HLA-DPB1.

<https://doi.org/10.1371/journal.pone.0253619.g001>

## Discussion

We first reported all the 11 loci HLA typing in South Koreans. In order to obtain high-resolution HLA types (3rd field allele types), 173 healthy Koreans were the HLA typed for all 11 HLA loci using next generation DNA sequencing (HLA-A, -B, -C, -DRB1, -DRB3/4/5, -DQA1, -DQB1, -DPA1 and -DPB1), of which were 100% concordant across HLA-A, -B, and -DRB1 (to 3rd field) [9]. No historical data was present for 3rd field of HLA-C, -DRB3/4/5, -DQA1, -DQB1, -DPA1, and -DPB1 (Table 1).

Compared with previous studies [9, 23, 24, 27, 30], the HLA-C, -DRB3/4/5, -DQA1, -DQB1, -DPA1, and -DPB1 alleles that were determined at high resolution in this study could be helpful for a clinical decision. The tolerance of a mismatch will depend on the patient's clinic and the discrepancy between unmatched alleles. We compared the allele types of 11 HLA loci in the Netherlands, UK, Hong Kong, Chinese, Philippines, and Vietnamese [14, 16, 17, 29, 31, 32]. A\*26:02 and B\*15:27 are already known alleles that are not uncommon in South Koreans and Japanese, and B\*55:07 is also found infrequently since it was reported as a variant of the B55 serotype in 1999 [45]. Some of the identified B\*54:19, C\*03:02:01, C\*08:06, C\*14:39, DRB1\*11:45, DRB1\*13:198, DRB1\*14:06:01, and DRB1\*14:142 alleles are included in the CWD list (CWD catalogue version 2.0, CIWD version 3.0.0) [46, 47], and some are rare ones. It is necessary to make a distinction between the two. Comparison of the Japanese populations (2nd field, 6 loci) [28] that were most close to the South Koreans, the alleles were B\*54:19, B\*55:07, C\*08:06, C\*14:39, DRB1\*11:45, DRB1\*13:198, and DRB1\*14:142. When creating commercial kits for South Koreans, manufacturers and developers must add these unique alleles to the test subject. When performing the HLA typing before transplantation for donors and patients, researchers will also need to include these unique alleles in the test so that they can be detected.

Compared with the Japanese population, in the 6 HLA loci [20], 7 kinds of the haplotypes with frequencies higher than 1% have been found in both populations, which were A\*02:07:01-B\*46:01:01-C\*01:02:01-DRB1\*08:03:02-DQB1\*06:01:01-DPB1\*05:01:01, A\*02:01:01-B\*15:01:01-C\*03:03:01-DRB1\*12:01:01-DQB1\*03:01:01-DPB1\*02:01:02, A\*33:03:01-B\*44:03:01-C\*14:03-DRB1\*13:02:01-DQB1\*06:04:01-DPB1\*04:01:01,



A\*24:02:01-B\*07:02:01-C\*07:02:01-DRB1\*01:01:01-DQB1\*05:01:01-DPB1\*05:01:01, A\*11:01:01-B\*15:01:01-C\*04:01:01-DRB1\*04:06:01-DQB1\*03:02:01-DPB1\*02:01:02, A\*24:02:01-B\*52:01:01-C\*12:02:02-DRB1\*15:02:01-DQB1\*06:01:01-DPB1\*09:01:01, and A\*24:02:01-B\*07:02:01-C\*07:02:01-DRB1\*01:01:01-DQB1\*05:01:01-DPB1\*04:02:01. The haplotype frequency of A\*33:03:01-B\*44:03:02-C\*07:06-DRB1\*07:01:01-DQB1\*02:02:01-DPB1\*13:01:01 (1.73%) in [Table 2](#) was lower than that of the top ten kinds of South Asian (2.17%) [16]. The top 10 haplotype frequencies of HLA-A, -B, and -C were the same at the 2nd field in the previous study ([S11 Table](#)) [24]. In the HLA class II, the haplotypes of the 3rd field of HLA-DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 (>1%) were not the same as the top ten most frequent haplotypes of different ethnic groups, European, South Asian, and African or Caribbean Black ([S12 Table](#)) [16]. Twenty-five DRB1-DRB3/4/5 haplotypes in [S13 Table](#) were investigated from the DRB1 and DRB3/4/5 genotyping data by comparison to the previously reported DRB structures [41, 42]. Only samples containing one copy of the secondary locus could also be evaluated against the previous study [14]. HLA genotyping data using the long-range PCR based NGS method are needed to understand the detailed DR haplotype structure and polymorphic generation. We compared the haplotype frequencies of 2-locus haplotypes of HLA class II genes in [S13 Table](#), HLA-DQA1-DQB1 and HLA-DPA1-DPB1 (>1%) with those in Fukuoka, Japan [21]. The haplotypes shared 8 for HLA-DQA1-DQB1 and 9 for HLA-DPA1-DPB1, respectively.

The PCR-sequence based typing method was observed as unresolved ambiguities in the IPD-IMGT/HLA database (Release version 3.42.0) ([S14 Table](#)). To increase the success rate of solid organ or hematopoietic stem cell transplantation, more extensive and high-resolution HLA typing is required. For the selection of solid organ donors in hypersensitized patients, a change in HLA type is needed, recognizing the need for high-resolution HLA types in traditional serologically defined HLA antigens [3]. More extended HLA typing region is needed to improve the success rate of unrelated hematopoietic stem cell transplantation [36].

In mitochondrial DNA study, MDS plot and unrooted phylogenetic tree were no significant differences [37]. The collected samples are composed from the same ethnic group in South Korea and can be said to have genetic homogeneity. Compared to other ethnic groups, we were able to account for genetic distances with MDS plots ([Fig 1](#)). It highlights the need to use multiple loci to study genetic population relationships such as genetic distance of HLA genes [43, 48]. The genetic distance of South Koreans was measured for more extended HLA loci. The genetic distances of the HLA-A, -B, -C, -DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 loci in the characteristics of the Korean samples support the theory that Koreans are primarily Northeast Asian origin. Our MDS analyses both genetic distance the South Koreans in close vicinity to Japanese and Han Chinese populations, whereas some analyses indicate a similarity to other Northeast Asian populations.

In conclusion, we analyzed the allele and haplotype frequencies of 11 entire and extensive HLA loci and the genetic distances by MDS plots compared to other ethnic groups. These data may be useful for more exact diagnostic tests of various transplantation and the genetic population relationship studies.

## Supporting information

**S1 Table. HLA-A, -B, and -DRB1 haplotype frequencies (>0.5%).**  
(DOCX)

**S2 Table. HLA-A, -B, -C, and -DRB1 haplotype frequencies (>0.5%).**  
(DOCX)

**S3 Table. HLA-A, -B, -C, -DRB1, and -DQB1 haplotype frequencies (>0.5%).**  
(DOCX)

**S4 Table. HLA-A, -B, -C, -DRB1, -DQA1, and -DQB1 haplotype frequencies (>0.5%).**  
(DOCX)

**S5 Table. HLA-DRB1, -DQB1, and -DPB1 haplotype frequencies (>0.5%).**  
(DOCX)

**S6 Table. The entire configuration file (.ini).**  
(DOCX)

**S7 Table. Alleles resolved on ambiguous allele combinations of the IPD-IMGT/HLA database (Release version 3.42.0) (n = 173).**  
(DOCX)

**S8 Table. The Hardy-Weinberg equilibrium of HLA-A, -B, -C, -DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 loci in South Korea population.**  
(DOCX)

**S9 Table. Allele frequencies of HLA-A, -B, and -C in South Koreans (N = 173).**  
(DOCX)

**S10 Table. Allele frequencies of HLA-DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 in South Koreans (N = 173).**  
(DOCX)

**S11 Table. Haplotype frequencies of HLA-A, -B, and -C except for overlapping parts of [Table 2](#) (>1%).**  
(DOCX)

**S12 Table. Haplotype frequencies of HLA-DRB1, -DQA1, -DQB1, -DPA1, and -DPB1 (>1%).**  
(DOCX)

**S13 Table. Haplotype frequencies of 2-locus haplotypes of HLA class II genes (>1%).**  
(DOCX)

**S14 Table. Expected PCR-SBT ambiguities (described in the IPD-IMGT/HLA database) solved by the NGS assay (n = 173).**  
(DOCX)

**S15 Table. HLA-A allele frequencies of 16 populations.**  
(DOCX)

**S16 Table. HLA-B allele frequencies of 16 populations.**  
(DOCX)

**S17 Table. HLA-C allele frequencies of 13 populations.**  
(DOCX)

**S18 Table. HLA-DRB1 allele frequencies of 16 populations.**  
(DOCX)

**S19 Table. HLA-DQA1 allele frequencies of 15 populations.**  
(DOCX)

**S20 Table. HLA-DQB1 allele frequencies of 16 populations.**  
(DOCX)

**S21 Table. HLA-DPA1 allele frequencies of 10 populations.**  
(DOCX)

**S22 Table. HLA-DPB1 allele frequencies of 14 populations.**  
(DOCX)

## Author Contributions

**Conceptualization:** In-Cheol Baek, Tai-Gyu Kim.

**Data curation:** In-Cheol Baek.

**Formal analysis:** In-Cheol Baek.

**Funding acquisition:** In-Cheol Baek, Tai-Gyu Kim.

**Investigation:** In-Cheol Baek, Eun-Jeong Choi, Dong-Hwan Shin, Hyoung-Jae Kim.

**Methodology:** In-Cheol Baek, Eun-Jeong Choi, Dong-Hwan Shin, Hyoung-Jae Kim.

**Project administration:** In-Cheol Baek, Hyoung-Jae Kim, Tai-Gyu Kim.

**Resources:** Dong-Hwan Shin, Hyoung-Jae Kim.

**Software:** In-Cheol Baek.

**Supervision:** Tai-Gyu Kim.

**Validation:** In-Cheol Baek, Dong-Hwan Shin, Hyoung-Jae Kim, Haeyoun Choi.

**Visualization:** In-Cheol Baek.

**Writing – original draft:** In-Cheol Baek.

**Writing – review & editing:** In-Cheol Baek.

## References

1. Lee SJ, Klein J, Haagenson M, Baxter-Lowe LA, Confer DL, Eapen M, et al. High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood*. 2007; 110(13):4576–83. Epub 2007/09/06. <https://doi.org/10.1182/blood-2007-06-097386> PMID: 17785583.
2. Cecka JM, Reed EF, Zachary AA. HLA high-resolution typing for sensitized patients: a solution in search of a problem? *American journal of transplantation: official journal of the American Society of Transplantation and the American Society of Transplant Surgeons*. 2015; 15(4):855–6. Epub 2015/03/18. <https://doi.org/10.1111/ajt.13169> PMID: 25778678.
3. Duquesnoy RJ, Kamoun M, Baxter-Lowe LA, Woodle ES, Bray RA, Claas FH, et al. Should HLA mismatch acceptability for sensitized transplant candidates be determined at the high-resolution rather than the antigen level? *American journal of transplantation: official journal of the American Society of Transplantation and the American Society of Transplant Surgeons*. 2015; 15(4):923–30. Epub 2015/03/18. <https://doi.org/10.1111/ajt.13167> PMID: 25778447.
4. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA Database. *Nucleic Acids Res*. 2020; 48(D1):D948–d55. Epub 2019/11/02. <https://doi.org/10.1093/nar/gkz950> PMID: 31667505.
5. Lan JH, Zhang Q. Clinical applications of next-generation sequencing in histocompatibility and transplantation. *Current opinion in organ transplantation*. 2015; 20(4):461–7. Epub 2015/06/25. <https://doi.org/10.1097/MOT.0000000000000217> PMID: 26107967.
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(5):393–403. Epub 2009/10/23. <https://doi.org/10.1111/j.1399-0039.2009.01345.x> PMID: 19845894.

7. Lange V, Bohme I, Hofmann J, Lang K, Sauter J, Schone B, et al. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC genomics*. 2014; 15:63. Epub 2014/01/28. <https://doi.org/10.1186/1471-2164-15-63> PMID: 24460756.
8. Moonsamy PV, Williams T, Bonella P, Holcomb CL, Høglund BN, Hillman G, et al. High throughput HLA genotyping using 454 sequencing and the Fluidigm Access Array System for simplified amplicon library preparation. *Tissue Antigens*. 2013; 81(3):141–9. Epub 2013/02/13. <https://doi.org/10.1111/tan.12071> PMID: 23398507.
9. Baek IC, Choi EJ, Shin DH, Kim HJ, Choi H, Kim TG. Distributions of HLA-A, -B, and -DRB1 alleles typed by amplicon-based next generation sequencing in Korean volunteer donors for unrelated hematopoietic stem cell transplantation. *Hla*. 2020. Epub 2020/11/13. <https://doi.org/10.1111/tan.14134> PMID: 33179442.
10. Wang C, Krishnakumar S, Wilhelmy J, Babrzadeh F, Stepanyan L, Su LF, et al. High-throughput, high-fidelity HLA genotyping with deep sequencing. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109(22):8676–81. Epub 2012/05/17. <https://doi.org/10.1073/pnas.1206614109> PMID: 22589303.
11. 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(4):305–16. Epub 2012/08/07. <https://doi.org/10.1111/j.1399-0039.2012.01941.x> PMID: 22861646.
12. 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. *Human immunology*. 2010; 71(10):1033–42. Epub 2010/07/07. <https://doi.org/10.1016/j.humimm.2010.06.016> PMID: 20603174.
13. Danzer M, Niklas N, Stabenheiner S, Hofer K, Pröll J, Stücker C, et al. Rapid, scalable and highly automated HLA genotyping using next-generation sequencing: a transition from research to diagnostics. *BMC genomics*. 2013; 14:221. Epub 2013/04/06. <https://doi.org/10.1186/1471-2164-14-221> PMID: 23557197.
14. Hou L, Enriquez E, Persaud M, Steiner N, Oudshoorn M, Hurley CK. Next generation sequencing characterizes HLA diversity in a registry population from the Netherlands. *Hla*. 2019; 93(6):474–83. Epub 2019/03/25. <https://doi.org/10.1111/tan.13535> PMID: 30907066.
15. Smith AG, Pereira S, Jaramillo A, Stoll ST, Khan FM, Berka N, et al. Comparison of sequence-specific oligonucleotide probe vs next generation sequencing for HLA-A, B, C, DRB1, DRB3/B4/B5, DQA1, DQB1, DPA1, and DPB1 typing: Toward single-pass high-resolution HLA typing in support of solid organ and hematopoietic cell transplant programs. *Hla*. 2019; 94(3):296–306. Epub 2019/06/27. <https://doi.org/10.1111/tan.13619> PMID: 31237117.
16. Guerra SG, Hamilton-Jones S, Brown CJ, Navarrete CV, Chong W. Next generation sequencing of 11 HLA loci characterises a diverse UK cord blood bank. *Hum Immunol*. 2020; 81(6):269–79. Epub 2020/04/20. <https://doi.org/10.1016/j.humimm.2020.04.001> PMID: 32305144.
17. Kwok J, Tang WH, Chu WK, Chan YS, Liu Z, Yang W, et al. High resolution allele genotyping and haplotype frequencies for NGS based HLA 11 loci of 5266 Hong Kong Chinese bone marrow donors. *Hum Immunol*. 2020; 81(10–11):577–9. Epub 2020/09/08. <https://doi.org/10.1016/j.humimm.2020.08.005> PMID: 32893027.
18. Liu C. A long road/read to rapid high-resolution HLA typing: The nanopore perspective. *Human immunology*. 2020. Epub 2020/05/11. <https://doi.org/10.1016/j.humimm.2020.04.009> PMID: 32386782.
19. Mosbrugger TL, Dinou A, Duke JL, Ferriola D, Mehler H, Pagkrati I, et al. Utilizing nanopore sequencing technology for the rapid and comprehensive characterization of eleven HLA loci; addressing the need for deceased donor expedited HLA typing. *Hum Immunol*. 2020; 81(8):413–22. Epub 2020/07/01. <https://doi.org/10.1016/j.humimm.2020.06.004> PMID: 32595056.
20. 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*. 2013; 82(1):82. Epub 2013/06/10. <https://doi.org/10.1111/tan.12142> PMID: 11169242.
21. Begovich AB, Moonsamy PV, Mack SJ, Barcellos LF, Steiner LL, Grams S, et al. Genetic variability and linkage disequilibrium within the HLA-DP region: analysis of 15 different populations. *Tissue Antigens*. 2001; 57(5):424–39. Epub 2001/09/15. <https://doi.org/10.1034/j.1399-0039.2001.057005424.x> PMID: 11556967.
22. Song EY, Park MH, Kang SJ, Park HJ, Kim BC, Tokunaga K, et al. HLA class II allele and haplotype frequencies in Koreans based on 107 families. *Tissue Antigens*. 2002; 59(6):475–86. Epub 2002/11/26. <https://doi.org/10.1034/j.1399-0039.2002.590604.x> PMID: 12445317.
23. Song EY, Park H, Roh EY, Park MH. HLA-DRB1 and -DRB3 allele frequencies and haplotypic associations in Koreans. *Hum Immunol*. 2004; 65(3):270–6. Epub 2004/03/26. <https://doi.org/10.1016/j.humimm.2003.12.005> PMID: 15041167.

24. Lee KW, Oh DH, Lee C, Yang SY. Allelic and haplotypic diversity of HLA-A, -B, -C, -DRB1, and -DQB1 genes in the Korean population. *Tissue Antigens*. 2005; 65(5):437–47. Epub 2005/04/28. <https://doi.org/10.1111/j.1399-0039.2005.00386.x> PMID: 15853898.
25. Gragert L, Madbouly A, Freeman J, Maiers M. Six-locus high resolution HLA haplotype frequencies derived from mixed-resolution DNA typing for the entire US donor registry. *Hum Immunol*. 2013; 74(10):1313–20. Epub 2013/06/29. <https://doi.org/10.1016/j.humimm.2013.06.025> PMID: 23806270.
26. Ozaki Y, Suzuki S, Shigenari A, Okudaira Y, Kikkawa E, Oka A, et al. HLA-DRB1, -DRB3, -DRB4 and -DRB5 genotyping at a super-high resolution level by long range PCR and high-throughput sequencing. *Tissue antigens*. 2014; 83(1):10–6. Epub 2013/12/21. <https://doi.org/10.1111/tan.12258> PMID: 24355003.
27. 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. *Annals of laboratory medicine*. 2015; 35(4):429–35. Epub 2015/07/02. <https://doi.org/10.3343/alm.2015.35.4.429> PMID: 26131415.
28. Ozaki Y, Suzuki S, Kashiwase K, Shigenari A, Okudaira Y, Ito S, et al. Cost-efficient multiplex PCR for routine genotyping of up to nine classical HLA loci in a single analytical run of multiple samples by next generation sequencing. *BMC genomics*. 2015; 16(1):318. Epub 2015/04/22. <https://doi.org/10.1186/s12864-015-1514-4> PMID: 25895492.
29. Zhou XY, Zhu FM, Li JP, Mao W, Zhang DM, Liu ML, et al. High-Resolution Analyses of Human Leukocyte Antigens Allele and Haplotype Frequencies Based on 169,995 Volunteers from the China Bone Marrow Donor Registry Program. *PloS one*. 2015; 10(9):e0139485. Epub 2015/10/01. <https://doi.org/10.1371/journal.pone.0139485> PMID: 26421847.
30. Park H, Lee YJ, Song EY, Park MH. HLA-A, HLA-B and HLA-DRB1 allele and haplotype frequencies of 10 918 Koreans from bone marrow donor registry in Korea. *International journal of immunogenetics*. 2016; 43(5):287–96. Epub 2016/08/12. <https://doi.org/10.1111/iji.12288> PMID: 27511726.
31. Geretz A, Cofer L, Ehrenberg PK, Currier JR, Yoon IK, Alera MTP, et al. Next-generation sequencing of 11 HLA loci in a large dengue vaccine cohort from the Philippines. *Hum Immunol*. 2020; 81(8):437–44. Epub 2020/07/14. <https://doi.org/10.1016/j.humimm.2020.06.010> PMID: 32654962.
32. Do MD, Le LGH, Nguyen VT, Dang TN, Nguyen NH, Vu HA, et al. High-Resolution HLA Typing of HLA-A, -B, -C, -DRB1, and -DQB1 in Kinh Vietnamese by Using Next-Generation Sequencing. *Frontiers in genetics*. 2020; 11:383. Epub 2020/05/20. <https://doi.org/10.3389/fgene.2020.00383> PMID: 32425978.
33. Ahn S, Choi HB, Kim TG. HLA and Disease Associations in Koreans. *Immune network*. 2011; 11(6):324–35. Epub 2012/02/22. <https://doi.org/10.4110/in.2011.11.6.324> PMID: 22346771.
34. Zhao LP, Alshiekh S, Zhao M, Carlsson A, Larsson HE, Forsander G, et al. Next-Generation Sequencing Reveals That HLA-DRB3, -DRB4, and -DRB5 May Be Associated With Islet Autoantibodies and Risk for Childhood Type 1 Diabetes. *Diabetes*. 2016; 65(3):710–8. Epub 2016/01/08. <https://doi.org/10.2337/db15-1115> PMID: 26740600.
35. Shin DH, Baek IC, Kim HJ, Choi EJ, Ahn M, Jung MH, et al. HLA alleles, especially amino-acid signatures of HLA-DPB1, might contribute to the molecular pathogenesis of early-onset autoimmune thyroid disease. *PLoS One*. 2019; 14(5):e0216941. Epub 2019/05/16. <https://doi.org/10.1371/journal.pone.0216941> PMID: 31091281.
36. Petersdorf EW, Anasetti C, Martin PJ, Gooley T, Radich J, Malkki M, et al. Limits of HLA mismatching in unrelated hematopoietic cell transplantation. *Blood*. 2004; 104(9):2976–80. Epub 2004/07/15. <https://doi.org/10.1182/blood-2004-04-1674> PMID: 15251989.
37. Hong SB, Kim KC, Kim W. Population and forensic genetic analyses of mitochondrial DNA control region variation from six major provinces in the Korean population. *Forensic science international Genetics*. 2015; 17:99–103. Epub 2015/04/23. <https://doi.org/10.1016/j.fsigen.2015.03.017> PMID: 25900647.
38. Dempster AP, Laird NM, Rubin DB. Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1977; 39(1):1–22. <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>.
39. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Molecular biology and evolution*. 1995; 12(5):921–7. Epub 1995/09/01. <https://doi.org/10.1093/oxfordjournals.molbev.a040269> PMID: 7476138.
40. Lancaster AK, Single RM, Solberg OD, Nelson MP, Thomson G. PyPop update—a software pipeline for large-scale multilocus population genomics. *Tissue Antigens*. 2007; 69 Suppl 1:192–7. Epub 2007/04/21. <https://doi.org/10.1111/j.1399-0039.2006.00769.x> PMID: 17445199.
41. Andersson G, Larhammar D, Widmark E, Serenius B, Peterson PA, Rask L. Class II genes of the human major histocompatibility complex. Organization and evolutionary relationship of the DR beta genes. *The Journal of biological chemistry*. 1987; 262(18):8748–58. Epub 1987/06/25. PMID: 3036826.

42. Sutton VR, Kienzie BK, Knowles RW. An altered splice site is found in the DRB4 gene that is not expressed in HLA-DR7, Dw11 individuals. *Immunogenetics*. 1989; 29(5):317–22. Epub 1989/01/01. <https://doi.org/10.1007/BF00352841> PMID: 2497069.
43. Johansson A, Ingman M, Mack SJ, Erlich H, Gyllensten U. Genetic origin of the Swedish Sami inferred from HLA class I and class II allele frequencies. *European journal of human genetics: EJHG*. 2008; 16(11):1341–9. Epub 2008/05/15. <https://doi.org/10.1038/ejhg.2008.88> PMID: 18478041.
44. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*. 1992; 48(2):361–72. Epub 1992/06/01. PMID: 1637966.
45. Lee KW, Yang SY. HLA-B22 diversity including a novel B54 variant (B\*5507) in the Korean population. *Human immunology*. 1999; 60(8):731–7. Epub 1999/08/10. [https://doi.org/10.1016/s0198-8859\(99\)00031-2](https://doi.org/10.1016/s0198-8859(99)00031-2) PMID: 10439319.
46. Mack SJ, Cano P, Hollenbach JA, He J, Hurley CK, Middleton D, et al. Common and well-documented HLA alleles: 2012 update to the CWD catalogue. *Tissue Antigens*. 2013; 81(4):194–203. Epub 2013/03/21. <https://doi.org/10.1111/tan.12093> PMID: 23510415.
47. Hurley CK, Kempenich J, Wadsworth K, Sauter J, Hofmann JA, Schefzyk D, et al. Common, intermediate and well-documented HLA alleles in world populations: CIWD version 3.0.0. *Hla*. 2020; 95(6):516–31. Epub 2020/01/24. <https://doi.org/10.1111/tan.13811> PMID: 31970929.
48. Kuranov AB, Vavilov MN, Abildinova G, Akilzhanova AR, Iskakova AN, Zhodybayeva EV, et al. Polymorphisms of HLA-DRB1, -DQA1 and -DQB1 in inhabitants of Astana, the capital city of Kazakhstan. *PLoS One*. 2014; 9(12):e115265. Epub 2014/12/23. <https://doi.org/10.1371/journal.pone.0115265> PMID: 25531278.