DOI 10.18699/SBB-2022-212

Alleles and haplotypes diversity of HLA-A, -B, -C, -DRB1, and -DQB1 genes in Russia

Ananeva A.[1]*, Ryzhkova T.V.[2], Gaifullina R.[3], Gusev O.[1], Shagimardanova E.[1]

[1] *Extreme Biology Laboratory, Regulatory Genomics Research Center, Institute of Fundamental Medicine and Biology, Kazan Federal University, Kazan, Russia*

[2] *Regulatory Genomics Research Center, Institute of Fundamental Medicine and Biology, Kazan Federal University, Kazan, Russia*

[3] *National Bone Marrow Donor Registry, Moscow, Russia*

[4] *Volga-Region Bone Marrow Donor Registry, Kazan, Russia*

** nastya.ananeva@gmail.com*

Key words: HLA, novel allele, diversity, haplotype

Motivation and Aim: HLA is well known as a highly polymorphic genetic system. HLA genes are in the MHC (major histocompatibility complex) and the frequency of alleles and haplotypes varies widely among human populations. The aim of our study was to characterize allele frequency distribution of the HLA-A, HLA-B, HLA-C, HLA-DQB1, HLA-DRB1 loci of the Russian population from different regions of Russia.

Methods and Algorithms: Blood samples from 5492 adult potential donors of hematopoietic stem cells recruited by National bone marrow donor registry named after Vasya Perevoshchikov were analyzed. We implemented an NGS-based workflow for the full length typing of HLA-A, -B, -C, -DRB1 and -DQB1 using Holotype HLA™ (Omixon, Hungary). Sequencing was performed using the MiSeq reagent kit v.2 through MiSeq instrument (Illumina, USA). Analysis of resulting FastQ data files was carried by HLA Twin™ software (Omixon, Hungary). HLA allele and haplotype frequencies were estimated via maximum-likelihood analysis from genotypic data through an expectation-maximization (EM) algorithm for unknown gametic phase implemented in the Arlequin program version 3.5.2.2.

Results: A total of 72 HLA-A, 122 HLA-B, 67 HLA-C, 34 HLA-DQB1, 89 HLA-DRB1 alleles were identified. The most frequent among the HLA-A were A*02:01:01 (26,7 %), A*03:01:01 (13,5 %), A*01:01:01 (11,2 %) and A*24:02:01 (11,2 %). Among the HLA-B locus, the most frequent variants were B*07:02:01 (11,3 %), B*13:02:01 (7,1 %), B*35:01:01 (6,4 %), B*18:01:01 (6,4 %), B*08:01:01 (6.3 %). Among the HLA-C locus C*07:02:01 (13,1 %), C*04:01:01 (13 %), C*06:02:01 (12,8 %), C*07:01:01 (11 %), C*12:03:01 (9,3 %). The most frequent allele of HLA-DQB1 was defined as DQB1*03:01:01 (19.1 %), other variants were DQB1*02:02:01 (14,7 %), DQB1*05:01:01 (14,1 %), DQB1*06:02:01 (10,9 %) and finally among the DRB1 locus - DRB1*07:01:01 (14,8 %), DRB1*01:01:01 (12 %) and DRB1*15:01:01 (11,7 %). There were a trend for higher frequency of HLA haplotype A*01:01:01~B*08:01:01~C*07:01:01~DQB1*02:01:01~DRB1*03:01:01, determined for 2,38 % and haplotype A*02:01:01~B*13:02:01~C*06:02:01~DQB1*02:01:01~ DRB1*07:01:01 determined for 1.05 % of donors.

Among 5,492 potential donors we identified 18 rare alleles and 17 very rare alleles. Also, we identified 16 novel alleles: 6 HLA-A, 7 HLA-B, 8 HLA-C, 4 HLA-DQB1, 3 HLA-DRB1. The names A*02:974, C*06:327, A*68:01:58, C*04:441:01:02, A*02:01:204,

C*12:349, A*03:444, B*39:189, B*07:458, B*44:348, A*03:390:02, B*27:05:57, B*08:295, C*05:269, A*02:1042, C*01:230 has been officially assigned by the WHO Nomenclature Committee. We also identified 12 potentially novel alleles among this group, which were submitted to the WHO Nomenclature Committee, but official names have not been assigned yet. This follows the agreed policy that, subject to the conditions stated in the most recent Nomenclature Report (Marsh et al. 2010), names will be assigned to new sequences as they are identified. Lists of such new names will be published in the following WHO Nomenclature Report.

Conclusion: The most frequent alleles and haplotype of Russian population are typical for Europeans populations. In the cohort analyzed 0.64 % of potential donors can be assumed as bearing rare or very rare HLA alleles. 28 novel and potentially novel alleles have been described among 5492 potential donors. Analysis of the HLA haplotypes frequencies in the population is necessary to assess the probability of detecting a matching donor. Also, such studies can be used for population genetic and HLA-disease association fields.

Acknowledgements: We would like to thank the Charity fund "National Bone Marrow Donors Registry named after Vasya Perevoshchikov" for support of the HLA typing and for the donor recruiting. This study was the part of the Kazan Federal University Strategic Academic Leadership Program (PRIORITY-2030).

References

1. Marsh S.G.E., Albert E.D., Bodmer W.F., Bontrop R.E., Dupont B., Erlich H.A., Fernández-Viña M., Geraghty D.E., Holdsworth R., Hurley C.K., Lau M., Lee K.W., Mach B., Maiers M., Mayr W.R., Müller C.R., Parham P., Petersdorf E.W., Sasazuki T., Strominger J.L., Svejgaard A., Terasaki P.I., Tiercy J.M., Trowsdale J. Nomenclature for factors of the HLA system, 2010. *Tissue Antigens*. 2010;75(4):291-455.