UNIVERSITY OF MEDICINE AND PHARMACY
"CAROL DAVILA", BUCHAREST
DOCTORAL SCHOOL
MEDICINE FIELD

DOCTORAL THESIS
SUMMARY OF THE DOCTORAL THESIS

Doctoral supervisor:
PROF. UNIV. DR. CONSTANTINESCU ILEANA
Doctoral student:
CARAGEA ANDREEA MIRELA

2025

UNIVERSITY OF MEDICINE AND PHARMACY

"CAROL DAVILA", BUCHAREST

DOCTORAL SCHOOL

LABORATORY MEDICINE

FREQUENCY OF HLA ALLELES AND HAPLOTYPES IN THE ROMANIA POPULATION

SUMMARY OF THE DOCTORAL THESIS

Doctoral supervisor:

PROF. UNIV. DR. CONSTANTINESCU ILEANA

Doctoral student:

CARAGEA ANDREEA MIRELA

2025

Contents

Introduction

The human leukocyte antigen (HLA) represents the major histocompatibility complex (MHC) in humans, located on the short arm of chromosome 6 (region 6p21), and comprises over 200 genes, many of which are highly polymorphic.

The HLA system is the most polymorphic human genetic system, with thousands of allelic variants identified (e.g., over 5000 alleles at the HLA-B locus). HLA genes are divided into two main classes with distinct functions: class I molecules (encoded by genes such as HLA-A, -B, -C) are expressed on the surface of all nucleated cells and present endogenous self or viral peptides to CD8^+ cytotoxic T lymphocytes, while class II molecules (HLA-DR, -DQ, -DP, expressed on specialized antigen-presenting cells) display exogenous peptides to CD4^+ helper T lymphocytes.

The fundamental role of these molecules is to allow the immune system to discriminate between self and non-self, triggering immune responses against pathogens or foreign cells. Because of their high variability, HLA genes are directly involved in transplant compatibility: HLA differences between a donor and a recipient can lead to recognition of the allograft as "non-self" and the initiation of immune rejection of the transplant.

Therefore, pre-transplant HLA compatibility testing is essential in solid organ and hematopoietic stem cell transplantation (HSCT) – HLA mismatch can cause serious phenomena such as graft-versus-host disease in bone marrow transplantation or acute rejection in organ transplantation. In addition, certain HLA allelomorphs predispose to autoimmune diseases: for example, over 90% of European patients with ankylosing spondylitis possess the HLA-B*27 allele, highlighting the close association between HLA and genetic susceptibility to this disease. Similar mechanisms link HLA to other immune pathologies (type 1 diabetes, rheumatoid arthritis, celiac disease, etc.), thus HLA represents a key factor in both transplantation immunology and the immunopathology of autoimmune diseases.

An important aspect of HLA genetics is the inheritance of HLA haplotypes. HLA genes on the same chromosome 6 are transmitted in a hereditary block, forming haplotypes specific to each parent. Each individual possesses two haplotype sets (maternal and paternal), and siblings have a 25% probability of receiving the same HLA haplotype from both parents. Consequently, only about a quarter of siblings are HLA-identical – a crucial aspect in choosing a family donor for transplantation.

Since ~70% of patients requiring a transplant do not have an HLA-identical related donor (sibling), it has become necessary to search for compatible unrelated donors through volunteer donor registries. In Romania, the National Registry of Voluntary Hematopoietic Stem Cell Donors (RNDVCSH) has developed over the last two decades to support patients without a family donor; by the end of 2019, approximately 66,000 volunteer donors were registered, of which only a fraction had been HLA-tested at high resolution. The number of hematopoietic stem cell transplants from unrelated donors has increased exponentially (an increase of >400% during 2004–2014 globally), and in 2020, 252 such transplants were performed in Romania.

The diversity of HLA alleles in the population poses major challenges in the rapid identification of a compatible donor: the more heterogeneous a population is in terms of HLA, the lower the chances of a perfect match, requiring the most extensive databases and knowledge of the immunogenetic profile of each population.

I. GENERAL PART

1. The major histocompatibility complex (MHC) and its role in the immune system

The genes of the major histocompatibility complex in humans, also known as the HLA system genes, have the highest degree of polymorphism in the human genome [1-7].

The major histocompatibility complex (MHC) is a group of genes located on chromosome 6 (Fig. 1.3.1), at cytogenetic location 6p21.3 [7, 10-15]. The term major histocompatibility complex refers to the primary biological function of this highly polymorphic system, namely that the molecules encoded by this complex are involved in organ and cell transplantation [7, 11, 12, 16, 17, 18].

HLA genes encode molecules that act as receptors for the body's own peptides (self) or peptides foreign to the body (nonself), the initiation of the immune response is done by their presentation to cytotoxic T lymphocytes or to helper T lymphocytes [7, 11, 19-24] .

The MHC contains a very large number of genes involved in different stages of the immune response, in addition to the HLA genes [24,25].

CMH is divided into three classes, namely :

- Class I genes
- Class II genes
- Class III genes

2. Solid organ and stem cell transplantation involving HLA genes

The determination of HLA variants and haplotypes is called HLA typing or tissue typing and is performed in clinical laboratories to help determine histocompatibility between the donor and recipient of solid organs or stem cells [26]. HLA typing is also used for anthropological studies, in order to make inferences regarding population relationships and history [20-26].

Transplantation is the "gold standard " in treatment of numerous diseases that are, in their evolutionary stages, in a terminal stage [19,26].

We have two types of transplant, namely cell transplant and solid organ transplant.

Hematopoietic stem cell transplantation allows the replacement of abnormal hematopoietic cells with normal ones [19,26]. Indications for stem cell transplantation include : myelodysplastic syndromes, bone marrow aplasias, thalassemias, hemoglobinopathies, hematological malignancies (acute or chronic leukemias, lymphomas) and solid tumors [18,19].

Stem cells can come from HLA-identical or haploidentical related donors or from HLA-identical unrelated donors [26].

3. HLA and associated diseases

Some autoimmune diseases are associated with HLA genes, but the mechanism has not been fully elucidated to date. Population studies have shown that the genetic predisposition to human autoimmune diseases is linked to HLA genes, especially class II genes [1,7, 11, 12] . Due to the polymorphism of HLA genes and the heterogeneity of autoimmune diseases, the mechanisms explaining this association are not yet fully understood.

II. PERSONAL CONTRIBUTIONS

4. Working hypothesis and general objectives

The purpose of the doctoral thesis

Given the importance of knowing the distribution of HLA alleles and haplotypes at the population level (both for optimizing clinical transplantation procedures and for anthropological and genetic association studies), numerous international studies have mapped HLA frequencies in various ethnic groups around the world.

These studies have highlighted marked differences between populations, allowing the groups to be related on phylogenetic and historical criteria (for example, the identification of ancestral haplotypes characteristic of Western Europe, and others prevalent in the Mediterranean area or Southeast Asia). In Romania, few previous works have investigated HLA distribution, and those that have used lower resolution methods (serological typing or partial genotyping) have not provided a complete picture of all HLA loci.

The immunogenetic profile of the Romanian population at high resolution has not been exhaustively characterized, leaving the frequency of rare alleles or the haplotype configuration in this ethnic group unknown. In this context, the general aim of this doctoral thesis was to perform a detailed, state-of-the-art mapping of the frequency of HLA alleles and haplotypes in the Romanian population, using the most modern genomic typing techniques.

To achieve this goal, several specific objectives were formulated:

(1) determining the distribution of HLA alleles at high resolution for the main class I (HLA-A, -B, -C) and class II (HLA-DRB1, -DQB1, -DQA1, -DPB1, -DPA1) loci among the autochthonous population and comparing these distributions with those reported for other peoples, in order to highlight the phylogenetic relationships of Romanians with other ethnic groups;

(2) identification of rare alleles and haplotypes present in the investigated population, an important aspect to estimate the probability of finding a compatible donor for Romanian patients (the rarity of some alleles reduces the chance of compatibility and requires the expansion of registries);

(3) providing a reference data set on HLA frequencies, also useful for exploring the association of certain HLA constellations with population-specific autoimmune pathologies (e.g., genetic predisposition to diseases such as type 1 diabetes or celiac disease, still insufficiently studied in Romania).

5. Materials and methods:

To achieve the objectives, a cross-sectional observational study was conducted on a sample of volunteer stem cell donors within the RNDVCSH.

The study group included approximately 420 healthy adult individuals, recruited as potential unrelated donors for transplantation and considered representative of the general population in Romania.

Biological samples (peripheral blood) were collected in EDTA vacutainers and transported to the Fundeni Clinical Institute (Immunogenetics and Virology Laboratory) for processing, between 2020–2021.

Genomic DNA was extracted from each sample using the commercial QIAamp DNA Blood Mini kit (Qiagen, Germany) according to the standard protocol based on silicate columns. The extraction method used allows the rapid obtaining of high purity and integrity DNA, suitable for PCR amplifications and sequencing, without the need for additional leukocyte isolation procedures. The concentration and quality of the extracted DNA were assessed spectrophotometrically, ensuring that all samples met the requirements for further analysis.

Subsequently, DNA samples were subjected to HLA genotyping using Next Generation Sequencing (NGS) technology. The Immucor MiaFora NGS Flex kit and the Illumina MiniSeq sequencing platform were used, which allow simultaneous high-resolution typing of multiple HLA genes. Extensive genomic regions containing the HLA genes of interest (including critical exons for allele determination) were amplified using a long-range PCR protocol, followed by fragmentation and preparation of indexed sequencing libraries. Through NGS, each sample was sequenced in massively parallel fashion, generating millions of DNA fragments that covered the target loci.

This approach offers the major advantage of resolving allelic ambiguities and accurately identifying rare variants, due to the high sequencing capacity and paired reads that determine allelic phase.

Compared to traditional methods (such as Sanger sequencing or PCR-SSP), the introduction of NGS in HLA typing has led to a substantial increase in the number of identifiable alleles and a reduction in the time required to obtain a complete result. In the present study, NGS allowed the acquisition of HLA alleles at a resolution level of 4–6 digits (equivalent to high-resolution allelic specificity), covering virtually all clinically relevant variants.

The raw sequencing data were analyzed using the dedicated MiaFora NGS Analysis software (vers. 2.0), which aligns the sequences to the reference genome and assigns the corresponding alleles to each typed locus, based on international HLA databases. The genotypes of each individual for each HLA gene investigated were thus determined. Subsequently, allelic frequencies (the number of occurrences of each allele, relative to the total alleles for the respective locus) and haplotype frequencies (the combined incidence of certain alleles on the same chromosome, relative to the total haplotypes possible to observe) were calculated.

For comparative analysis, the results obtained in the Romanian cohort were compared with published data from other populations – both geographically close European populations and wider continental environments – using statistical tests of difference in proportions and methods of analysis of genetic distances between populations (such as the calculation of the genetic divergence index based on allele frequencies). The validity of the data was supported by the fact that the observed distributions do not

differ significantly from those previously reported in the literature for European populations of comparable ethnic origin.

6. Results

Frequency of HLA alleles (class I) :

In total, in the ~420 genomes analyzed, 109 distinct HLA alleles were identified at high resolution, distributed as follows: 31 different alleles at the HLA-A locus, 49 at the HLA-B locus, and 29 at the HLA-C locus.

This reflects the high diversity of the HLA system even in a relatively small sample size. Among the class I alleles genotyped, those at the HLA-A locus showed the following frequency hierarchy: HLA-A*02:01:01 – the most common allele, with a relative frequency of approximately 26.1% among the individuals tested, followed by HLA-A*01:01:01 (~12.5%), HLA-A*24:02:01 (~11.7%) and HLA-A*03:01:01 (~9.7%). Other HLA-A alleles with notable prevalence were A*11:01:01 and A*32:01:01, each present at ~8.6%.

These first 6 HLA-A alleles account for ~80% of all identified A alleles, indicating the existence of a core of common alleles in the population. At the same time, relatively rare HLA-A alleles were also detected: approximately 15 of the HLA-A alleles (out of a total of 31) had individual frequencies below 1%, but accounted for only ~5.7% of the observed alleles. This distribution highlights a rich polymorphism, with a few very common alleles and many minority alleles present at low proportions.

At the HLA-B locus, the most frequent allele in the Romanian population studied was HLA-B*18:01:01, representing ~11.3% of alleles at this locus. The next most frequent alleles were HLA-B*51:01:01 (~10.8%) and HLA-B*08:01:01 (~7.8%). It is therefore found that no HLA-B allelomorph exceeds ~12% of the total – unlike some island or restricted populations where HLA-B alleles such as *07 or *51 can reach much higher frequencies.

The distribution of HLA-B alleles in Romania seems aligned with southeastern European trends (for example, the B18 allele is common in the Balkan basin), although the B35 allele (reported as frequent in older studies) is not found among the top three in frequency in the high-resolution sequencing data. In contrast, the position of the B*08 allele, part of the ancestral European haplotype 8.1 (A1-B8-DR3), which has a moderate frequency in the Romanian population, is noteworthy, suggesting a non-homogeneous presence (possibly due to genetic input from Western Europe through historical migrations). At the HLA-C locus, the three most common alleles identified were HLA-C*07:01:01 (approximately 17.4%), HLA-C*04:01:01 (~13.5%) and HLA-C*12:03:01 (~10.7%).

The C*07 allele, combining both the *07:01 and 0*7:02 subtypes, has an even higher presence if the allelic group is considered as a whole – a fact also found in other European populations, where alleles from the C07 family are often dominant. Also, the C*04 allele (including the common subtype C*04:01) is found with notable frequencies in Europe and appears among the first in Romania (~13-14%). The

third allele in frequency, C*12:03, is more specific to Eastern Europe and Western Asia, which could suggest oriental genetic influences in the autochthonous gene pool.

As with the other loci, in addition to these main alleles, numerous HLA-C alleles with low individual frequencies (below 5%) were detected, which once again confirms the high degree of polymorphism of the HLA system in Romania.

Compared to other populations, it can be stated that the HLA class I allelic profile of Romania follows the general European pattern, in which HLA-A02:01 is usually the most frequent allele (both in Western and Central-Eastern Europe). In almost all European countries, A02 has the highest share – a fact confirmed by the present data. Also, the A01, B07, B51, C*07 alleles appear constantly among the dominant variants in various European populations, being also found in the Romanian set (B*07 and B*51 are very widespread in southern Europe, and their significant presence in Romania was expected; similarly, C*07 is widespread throughout Europe). However, each population presents specific nuances: for example, in Romania the B*18 allele stood out as the first in frequency, while in other close populations (Bulgaria, Serbia) B*35 or B*51 sometimes occupy the first place. These differences emphasize the importance of local HLA frequency studies.

Frequency of HLA alleles (class II) :

For class II genes, the analysis focused mainly on the HLA-DRB1, -DQB1, -DQA1 loci (encoding the HLA-DQ heterodimer) and the HLA-DPB1, -DPA1 loci (encoding the HLA-DP heterodimer).

The allelic distribution at the HLA-DRB1 locus (present on all antigen-presenting cells) shows that the most frequent allelic groups in the Romanian population are HLA-DRB1*11 and HLA-DRB1*13, each representing approximately 10-11% of the total. These are followed in prevalence by the alleles from the DRB1*04 (~9.5%), DRB1*07 (~9.3%) and DRB1*01 (~8.4%).

Together, these five DRB1 specificities account for over 45% of the detected class II alleles, with the remaining specificities having lower frequencies (for example, alleles in the DRB115 groups – which also includes the DRB1*16 subtype, DRB1*14, DRB1*03, DRB1*08, etc. each having less than 6% of the total). The HLA-DRB1 profile in Romania proves to be similar to European averages: at continental level, the most common DRB1 alleles reported are, in descending order, DRB1*01:01, DRB1*11:01, DRB1*13:01, DRB1*04:01 and DRB1*15/16:01, which closely corresponds to the set of peak alleles found in the present study (with the mention that in the Romanian population the DRB1*16 allele – equivalent to part of DRB1*15 in the old nomenclature – has a relatively lower frequency than DRB1*11 or *13). The results obtained confirm previously published statistics for HLA-DRB1 in comparable European populations, the differences being minor and probably explainable by sample size or regional particularities.

At the HLA-DQB1 locus, which encodes the β chain of the HLA-DQ molecule, the allele distribution was relatively uniform, with no single allele clearly dominating. None of the HLA-DQB1 alleles individually exceeds ~15% frequency in the studied cohort. We can deduce, from the association with

the correlative DQA1 alleles, that the most widespread DQ specificities in the Romanian population are those corresponding to the HLA-DQ2 and HLA-DQ7 heterodimers.

For example, the DQ2.5 heterodimer (formed by the DQA1*05:01 and DQB1*02:01 alleles) is well represented, and this is also the HLA complex associated with celiac disease in over 90% of cases. The presence of the DQB1*02:01 and DQA1*05 alleles in the general population of Romania (even if at moderate frequencies) explains why a significant proportion of individuals may be genetically predisposed to gluten enteropathy, an autoimmune disease whose incidence depends on the distribution of these alleles in the population.

Also, the HLA-DQ8 heterodimer (DQA1*03 together with DQB1*03:02) – associated with a lower risk of celiac disease and a predisposition to type 1 diabetes – is present in the population, although its component alleles have frequencies below 5% each.

Overall, considerable allelic diversity is also noted at the DQB1 locus, suggesting that the population does not have a simplified DQ profile (as occurs in closely related or isolated groups, where often 1-2 DQB1 alleles may be predominant). Regarding the HLA-DQA1 locus (the α chain of the DQ molecule), the distribution in Romania indicates a single relatively common allele, namely HLA-DQA1*05:05:01:01, which reached ~11.4% of the total – being also the only DQA1 allele with a frequency above the 10% threshold.

The following DQA1 alleles had frequencies below this threshold (second in frequency ~9.4%, third ~6-7%, etc.), which shows that the variability is distributed among several alleles of secondary importance. The DQA1*05:05 allele (subtype of the DQA1*05 group) is notable because it often associates in haplotype with DQB1*03:01 (together forming the DQ7 molecules) and with DRB1*11:04 – a combination characteristic of Southeastern Europe.

Moreover, this constellation (DR11-DQ7) explains the close frequencies of the observed DRB1*11, DQA1*05:05 and DQB1*03:01 alleles. For the HLA-DP complex, formed by the α (DPA1) and β (DPB1) chains, the results reveal an interesting situation: at the HLA-DPA1 locus, a single allele dominates at the population level, namely HLA-DPA1*01:03:01, which represents approximately 80% of the identified DPA1 alleles. This overwhelming predominance of the DPA1*01:03 allele is a phenomenon known in other populations as well – the DP heterodimer often having a relatively monomorphic alpha chain, in contrast to the greater variability of the beta chain.

Thus, HLA-DP variability is particularly evident at the HLA-DPB1 locus, where several alleles with notable frequencies were detected.

The three most frequent DPB1 alleles in the Romanian sample were: HLA-DPB1*04:01:01 (approx. 33.3%), HLA-DPB1*02:01:02 (~20.7%) and HLA-DPB1*04:02:01 (~15.7%).

Together, these three alleles total around 70% of the DPB1 alleles present, indicating moderate polymorphism at this locus – a few common alleles and the rest (numerous) of low frequency.

None of the HLA-DPB1 alleles individually reach a frequency comparable to the dominant DPA1*01:03 allele, but their combined distribution is sufficiently balanced to confer heterogeneity to

DP haplotypes. This is of practical relevance: while most individuals will share the same DPA1 allele (which simplifies matching on the DP alpha chain), differences on the DPB1 beta chain can generate incompatibilities – hence knowledge of the DPB1 allelic repertoire in the population is important when searching for a perfectly matched DP donor.

Frequency of HLA haplotypes :

In addition to the distribution of each individual allele, the thesis also investigated the frequency of HLA haplotypes – specific combinations of alleles inherited together on the same chromosome. Haplotype analysis provides an integrative perspective, revealing how often certain common allelic associations occur in the population (e.g., HLA-A with HLA-B with HLA-DRB1).

The main result in the case of extended haplotypes (including alleles from HLA-A, -B, -C, -DRB1) is the absence of any pronounced "modal" haplotype in the Romanian population, the frequency distribution being extremely dispersed. No complete HLA haplotype exceeds a few percent frequency – the maximum threshold observed being below ~5% for the most common haplotype. Approximately 98-99% of the identified HLA haplotypes had frequencies below 1%, indicating an exceptionally high haplotypic diversity.

In other words, almost every individual in the population has a different haplotype combination, the chances of two unrelated individuals sharing an identical HLA haplotype being very small (except for common pan-European haplotypes with a frequency of a few percent). However, certain relatively frequent (without being dominant) haplotypes have also been highlighted in Romania that fit into the known picture of Europe.

A notable example is the so-called "Balkan" haplotype: HLA-A*02:01 – B*18:01 – C*07:01/02 – DRB1*11:04 – DQB1*03:01, which was found in a modest proportion of individuals, but still stands out as one of the most frequently observed (around 3-4% of chromosomal haplotypes). This haplotype is known for its high frequency in populations from southeastern Europe (Balkan Peninsula) and probably reflects a common regional genetic pool.

Its components include the predominant alleles in the aforementioned class I distribution (A*02 and B*18, C*07) and DRB1*11, all of which are among the most frequent in Romania – so it is not surprising that they appear together as a relatively widespread haplotype.

Another haplotype of interest – this time associated with Western Europe – is the ancestral haplotype "8.1": HLA-A01:01 – B*08:01 – C*07:01 – DRB1*03:01 – DQB1*02:01 (known for its association with the HLA-DR3/DQ2 suite).

In the general Romanian population, this haplotype occurs, but with a low frequency (below 2%), lower than in northwestern Europe. Interestingly, in ethnic subgroups of Germanic origin in Romania – such as the Transylvanian Saxons or among Romanian emigrants settled in Germany – the A1-B8-DR3 haplotype is much more common (~6%), in parallel with a minimal presence of the Balkan haplotype A2-B18-DR11 in those groups.

This contrast reflects the influence of ancestral origin on haplotype frequencies: communities of Germanic origin retain a western HLA profile (rich in haplotype 8.1), while majority Romanians have an eastern/southeastern European profile (rich in haplotype B18-DR11).

7. Discussions

The frequencies of HLA alleles and haplotypes obtained in this study provide a series of essential information with practical and scientific applications. In practical terms, for the bone marrow transplantation program, the data suggest that the absence of high-frequency HLA haplotypes in the Romanian population translates into an increased difficulty in finding fully compatible donors for each patient.

Each patient of Romanian origin has, on average, relatively rare HLA combinations, which makes it unlikely that another unrelated individual in the population will have exactly the same HLA profile. Consequently, the chance of 100% HLA compatibility in the absence of a sibling is lower, compared to populations where there are dominant haplotypes (where certain donors are compatible with more than one patient).

This highlights the importance of expanding the volunteer donor base so that as many rare haplotypes as possible are represented in the registry, increasing the probability of finding at least one compatible donor for each patient. At the same time, the results argue for the need for ethnic diversification of registered donors: HLA matches are much more likely within the same ethnic or regional group, so the inclusion of national minorities and people with mixed ethnic origins in the RNDVCSH will increase the chances of successful searches for compatible donors.

In addition, given the mobility of the population (emigration of some Romanians and immigration of new groups), the constant updating of the HLA profile becomes relevant for international cooperation in transplantation: Romania can benefit from sharing data and donors globally, integrating into the international donor network, so that a Romanian patient can find a compatible donor in another country and vice versa.

From a technical point of view, the study also demonstrates the added value of high-resolution HLA typing via NGS for donor registries. The NGS method allowed for precise identification of alleles and haplotypes, eliminating phase ambiguities and highlighting rare alleles, information that was previously missing.

This detailed data can be used to improve virtual donor-patient matching algorithms – for example, by more accurately calculating the probability of finding a compatible donor for a given patient, taking into account the frequency of the patient's haplotype in the population. Knowledge of local frequencies thus allows an estimate of the chances of matching: very rare haplotypes require extensive searches (perhaps internationally), while relatively common haplotypes have a higher chance of success in the domestic registry.

NGS typing also highlighted the presence of novel or very rarely documented allelomorphs in the population, contributing to the update of global HLA databases and underlining the importance of continued genetic surveillance.

Last but not least, these results can also guide biomedical research: for example, correlating the frequency of some risk HLA alleles with the incidence of autoimmune diseases in Romania. If we know that the HLA-B27 allele has ~4-5% frequency in the population (according to the study data, including subtypes B27:05 and B27:02), we can understand why the prevalence of ankylosing spondylitis remains similar to other European countries; likewise, the frequency of the DQ2 heterodimer (DQA105:01-DQB1*02:01) in the population explains why celiac disease has a significant predisposing pool, even if phenotypically the disease is underdiagnosed in our country.

8. Conclusions

The present study makes a substantial contribution to the knowledge of the immunogenetic profile of the Romanian population, being the first exhaustive research at high resolution on the HLA distribution in this population group.

The results obtained clearly demonstrate that the distribution of HLA alleles and haplotypes in Romania is very broad and balanced, without the existence of any alleles or haplotypes that overwhelmingly dominate.

This picture of diversity is consistent with the patterns observed in other large European populations, suggesting that the Romanian population fits into a European genetic model in which variability is the rule, not the exception. The characterization of the native HLA profile was thus achieved and it was shown, through international comparisons, that Romania is in a common genetic cluster with Western Europe, while also having regional particularities (such as the high frequency of certain Balkan alleles). Rare alleles and haplotypes present in the population were identified, essential information for estimating the chances of compatibility in transplant programs.

Also, by highlighting the distribution of haplotypes such as DQ2 and DQ8, the paper provides a starting point for future studies on the correlation of HLA genotype with predispositions to autoimmune diseases in Romania – a field still little explored in our country. From an applications perspective, the conclusions support the need for complete reporting of HLA data for each national population, as a basis for optimizing hematopoietic stem cell donor registries. In the case of Romania, the results of this thesis can be immediately used by RNDVCSH and transplant centers, both to guide new donor recruitment strategies (targeting underrepresented ethnic subgroups), and to refine graft allocation procedures based on virtual compatibility criteria.

The large volume of knowledge generated – including the detailed frequencies of each HLA allele and the list of haplotypes with their probability of occurrence – also constitutes a benchmark for any future research in the field of population immunogenetics in the Romanian space.

As demographic dynamics and population admixture continue, HLA distributions are expected to evolve, increasing the observed diversity. In this regard, the continued study of HLA frequency in the population is essential not only for transplant medicine, but also for understanding the genetics of complex immune-mediated diseases.

In conclusion, the present research provides an updated and detailed overview of the HLA landscape of the analyzed population in Romania and underlines its importance in a clinical and scientific context. Finally, we emphasize that the integration of this knowledge into medical practices (especially in transplantation and in the monitoring of HLA-associated diseases) will lead to tangible benefits for the healthcare system.

Bibliography

1. G Rodey. HLA beyond tears: Introduction to human histocompatibility. DeNovo Publishing House, 2000, 2nd edition.
2. SY Choo. The HLA system: genetics, immunology, clinical testing, and clinical implications. Yonsei Medical Journal, 2007; 48(1):11-23.
3. EP Kransdorf, MJ Pando, L Gragert, B Kaplan. HLA Population Genetics in Solid Organ Transplantation. Transplantation, 2017 ; 101(9).
4. G Tumer, B Simpson, TK Roberts. Genetics, Human Major Histocompatibility Complex (MHC). StatPearls Publishing, 2021.
5. I Constantinescu, V Boscaiu, P Cianga, AA Dinu, E Gai, M Melinte, A moise. The frequency of HLA alleles in the Romanian population. Immunogenetics, 2016; 68(3).
6. http://www.allelefrequencies.net/default.asp
7. I Constantinescu. Transplant Immunology. UMF Carol Davila University Publishing House, 2007.
8. M Carp-Cărare, D Timofte. Immunology and immunopathology .
9. A Olinescu. Immunology. Didactic and Pedagogical Publishing House, 1995, Bucharest
10. CA Janeway, P Travers, M Walport, et al. Immunobiology: The Immune System in Health and Disease. The components of the immune system. 5th edition. Garland Science, 2001
11. Mihăilescu G. Immunology and immunochemistry. University of Bucharest Publishing House, 2001.
12. Rich R, Fleisher A, Shearer W, Schroeder H, Frew A, Weyand C. Clinical Immunology. Third edition. Elsevier, 2008.
13. Image 1 Westover J. Immune dysfunction in autism spectrum disorder. DOI:10.5772/22318. (Chromosome 6 picture). Fig 1.1
14. B Lewin. Genesis VII. Oxford University Press, 2000.
15. Shiina T, Inoko H, Kulski JK. An update of the HLA genomic region, locus information and disease association. Tissue Antigens 2004 ; 64(6):631-649. DOI:10.1111/j.1399-0039.2004.00327.x

16. Belich, MP, Madrigal, JA, Hildebrand, WH, et al. (1992) Unusual HLA-B alleles in two tribes of Brazilian Indians. Nature 357, 326–329.
17. Klein, K. (1986) Natural History of the Major Histocompatibility Complex. John Wiley and Sons, Inc., Toronto, Canada.
18. Cao, K., Chopek, M., and Fernandez-Vina, MA (1999) High and intermediate resolution DNA typing systems for class I HLA-A, B, C genes by hybridization with sequence-specific oligonucleotide probes (SSOP). rev. Immunogenic. 1, 177–208.
19. Balazs, I., Beekman, J., Neuweiler, J., Liu, H., Watson, E., and Ray, B. (2001) Molecular typing of HLA-A, -B, and DRB using a high throughput micro array format. Hum. Immunol. 62, 850–857.
20. Consolandi, C., Frosini, A., Pera, C., et al. (2004) Polymorphism analysis within the HLA-A locus by universal oligonucleotide array. Hum. Moved. 24, 428–434.
21. Li, AX, Seul, M., Cicciarelli, J., Yang, JC, and Iwaki, Y. (2004) Multiplexed analysis of polymorphisms in the HLA gene complex using bead array chips. Tissue Antigens 63, 518–528.
22. Olerup, O. and Zetterquist, H. (1992) 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 39, 225–235.
23. Bunce, M., O'Neill, CM, Barnardo, MC, et al. (1995) Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes using sequence-specific primers (PCR-SSP). Tissue Antigens 46, 355–367.
24. Newton, CR, Graham, A., Heptinstall, LE, et al. (1989) Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). Nucleic Acids Res. 17, 2503–2516.
25. van der Vlies, SA, Voorter, CE, and van den Berg-Loonen, EM (1998) A reliable and efficient high resolution typing method for HLA-C using sequence-based typing. Tissue Antigens 52, 558–568.
26. Turner, S., Ellexson, ME, Hickman, HD, et al. (1998) Sequence-based typing provides a new look at HLA-C diversity. J. Immunol. 161, 1406–1413.

List of publications

Original ISI listed articles

1. Caragea AM, Ursu RI, Maruntelu I, Tizu M, Constantinescu AE, Tălăngescu A, Constantinescu I. High Resolution HLA-A, HLA-B, and HLA-C Allele Frequencies in Romanian Hematopoietic Stem Cell Donors. Int J Mol Sci. 2024 Aug 14;25(16):8837. doi: 10.3390/ijms25168837, Q1, IF 4.9. Website: https://www.mdpi.com/1422-0067/25/16/8837 .
2. Caragea AM, Ursu RI, Bohîltea LC, Iordache P, Constantinescu AE, Constantinescu I. Characterization of HLA-A/HLA-B/HLA-C/HLA-DRB1 Haplotypes in Romanian Stem Cell Donors Through High-Resolution Next-Generation Sequencing. *International J. Mol. Sci.* 2025, *26* , 5250. https://doi.org/10.3390/ijms26115250 , Q1, IF 4.9. Website: https://www.mdpi.com/1422-0067/26/11/5250
3. Caragea AM, Ursu RI, Vișan LD, Mărunțelu I, Iordache P, Constantinescu AE, Tizu M, Tălăngescu A, Constantinescu I. High-resolution HLA-DRB1 allele frequencies in a Romanian cohort of stem cell donors. BJMG. 2024, 27(1):1-8. Doi: 10.2478/bjmg-2024-0009, Q4, IF 0.5.

Original BDI quoted articles

1. Caragea AM, Ursu RI, Iacob MM, Surugiu MS, Iordache P, Matei A, Constantinescu AE, Vișan LD, Constantinescu I. High Resolution HLA-DPA1 AND HLA-DPB1 allele frequencies in a Romanian cohort. Romanian Archives of Microbiology and Immunology, 2024, 83(2):87-93, DOI: 10.54044/RAMI.2024.02.02

Book chapters in international publishing houses

1. Caragea AM, Bohiltea L, Constantinescu A, Constantinescu I, Ursu RI. NGS and Immunogenetics: Sequencing the HLA Genes. Book chapter from DNA Sequencing - History, Present and Future Genetics, edited by Abdurakhmonov IY. IntechOpen; 2025. Available from: http://dx.doi.org/10.5772/intechopen.1008527 .
2. Caragea AM, Bohiltea L, Constantinescu A, Constantinescu I, Ursu RI. Perspective Chapter: Decoding the Complexity of HLA Genes – The Heart of Modern Immunogenetics. Advances and Trends in Population Genetics Studies [Working Title]. IntechOpen; 2025. Available from: http://dx.doi.org/10.5772/intechopen.1008543

Awards obtained

1. "Grigore Ghyka" Clinical Immunology Award , for the paper "Frequency of HLA haplotypes in the Romanian population - preliminary data " . Awarded on the occasion of the 49th Conference of the Romanian Society of Immunology, 2020.