COMMON HLA ALLELES AND HAPLOTYPES IN THE AUSTRIAN POPULATION BASED ON 1370 UMBILICAL CORD BLOOD DONORS

MASTER THESIS

Stefanie Nerstheimer

Univ.-Prof. Dr. Peter Macheroux

Department of Biochemistry

University of Technology

Graz

Graz, December 19th, 2012

ABSTRACT

In humans the MHC (Major Histocompatibility Complex) is called HLA (Human Histocompatibility Complex) after its discovery on a human leucocyte. The HLA proteins play an important role not only in the immune response but also in transplantations. By presenting antigens to T-lymphocytes the HLA molecules help the immune system to discriminate between "self" and "non-self" and therefore are the main immunological barrier for transplantations. Additionally, the HLA gene region is the most polymorphic in the human genome. Due to migration and evolution there are immense differences between various populations concerning the frequency of HLA-alleles and HLA-haplotypes. As a consequence, it is extremely important to match donor and recipient as well as possible and also to provide extensive HLA-data for various populations. Only with this strategy it is possible to establish stem cell banks and donor registries in order to find a possible match as quickly as possible when a donor is needed. It is also a way to actively recruit donors from ethnic minorities or identify highly compatible donors. The aim of this master thesis is to present extensive HLA-data for the Austrian population for the first time. Previously published studies including Austrian data were rare, only included small population samples and the data sets were not complete. In this study 1370 unrelated Austrian umbilical cord blood samples were typed for HLA A, HLA B, and HLA DRB1. Frequent alleles (low and high resolution) and haplotypes (low resolution) were listed. HLA C allele frequencies were investigated in a subgroup of 503 samples. Low resolution HLA-allele frequencies were compared to previously published data of Austrian bone marrow and blood donors - no significant differences were found, confirming the data presented in this master theses. Additionally HLA-haplotypes were compared to those presented in a large German population study - the results also did not differ. This surprising equality concerning frequent HLA haplotypes between Germans and Austrians led to the conclusion that Austrian stem cell donors can be considered for German recipients and vice versa. Furthermore it is important to focus more specific on minority donors since no uncommon haplotype was detected in this sample. To conclude, an extensive report of HLA-data is needed for every population in order to optimize stem cell donor databases. This master thesis shows such data for the very first time and also it was demonstrated that there are no significant differences between the Austrian and the German population regarding HLA haplotype distribution.

ZUSAMMENFASSUNG

Der menschliche MHC (Major Histocompatibility Complex) wird nach seiner Entdeckung auf menschlichen Leukozyten HLA - Complex (Human Histocompatibility Complex) genannt. HLA Proteine spielen nicht nur im Zuge der Immunantwort, sondern auch bei Transplantationen eine wichtige Rolle. Indem sie Antigene für T-Lymphozyten präsentieren, unterstützen HLA Moleküle das Immunsystem bei der Unterscheidung zwischen "Fremdem" und "Eigenem" und fungieren so auch als immunologische Barriere für Transplantationen. Die HLA Genregion besitzt außerdem die größte Polymorphie im gesamten menschlichen Genom. Aufgrund von Evolution und Migration gibt es in Bezug auf die häufigsten Allele und Haplotypen auch signifikante Unterschiede zwischen den einzelnen Bevölkerungsgruppen. Aus all diesen Fakten folgt, dass die HLA-Übereinstimmung zwischen einem Spender und einem Empfänger im Zuge einer Transplantation so gut wie möglich sein sollte, und dass eine möglichst detaillierte Darstellung der häufigsten Allele und Haplotypen für jede Bevölkerungsgruppe notwendig ist. Nur dadurch ist gewährleistet, dass Stammzellbanken und Spenderregister aufgebaut werden können um einen möglichen Spender so schnell wie möglich identifizieren zu können. Diese Informationen tragen auch dazu bei, Spender aus ethnischen Minderheiten gezielt rekrutieren zu können. Das Ziel dieser Masterarbeit ist es, zum ersten Mal eine große Menge an HLA Daten für die österreichische Bevölkerung zu präsentieren und zu analysieren. Zuvor publizierte österreichische HLA Daten basierten auf einer wesentlich geringeren Anzahl von Proben und waren teilweise unvollständig. In dieser Studie wurden 1370 nicht-verwandte österreichische Nabelschnurblutspenden für HLA-A, HLA-B, und HLA-DRB1 typisiert. Häufige Allele (niedrig- und hochaufgelöst) sowie häufige Haplotypen (niedrig aufgelöst) wurden identifiziert. HLA-C wurde bei 503 Spenden getestet. Die niedrig aufgelösten HLA-Allele Frequenzen wurden mit bisher publizieren österreichischen Daten verglichen. Es wurde keine signifikanten Abweichungen entdeckt, was die Daten in dieser Masterarbeit bestätigt. Zusätzlich wurden die HLA-Haploytpen mit Daten einer Publikation aus Deutschland verglichen - hier wurden ebenfalls keine signifikanten Unterschiede festgestellt. Diese überraschen Gleichheit zwischen der österreichischen und der deutschen Bevölkerung in Bezug auf die HLA-Typen führt zu dem Schluss, dass österreichische Stammzellspender für deutsche Patienten berücksichtig werden können und umgekehrt. Die Daten haben weiters gezeigt, dass der Focus der Spenderrekrutierung mehr auf die Mobilisierung von Spendern aus ethnischen Minderheiten gelegt werden muss, da in den hier präsentierten Daten keine seltenen HLA-Haploytpen entdeckt werden konnten.

INDEX

ABSTRACT ..... **- 1 -**

ZUSAMMENFASSUNG ..... **- 2 -**

INDEX ..... **- 3 -**

LIST OF FIGURES ..... **- 5 -**

LIST OF TABLES ..... **- 6 -**

LIST OF ABBREVIATIONS ..... **- 7 -**

THEORETICAL PART ..... **- 8 -**

1. INTRODUCTION ..... - 8 -
2. THE MAJOR HISTOCOMPATIBILITY COMPLEX (MHC) AND ITS ROLE IN THE IMMUNE SYSTEM ..... - 9 -
   - *2.1. The adaptive immune system* ..... *- 9 -*
   - *2.2. T-lymphocytes* ..... *- 10 -*
   - *2.3. The Major Histocompatibility Complex (MHC)* ..... *- 10 -*
3. MHC PROTEINS ..... - 12 -
   - *3.1. MHC class I proteins* ..... *- 12 -*
   - *3.2. MHC class II proteins* ..... *- 13 -*
4. HLA GENES OF THE MHC ..... - 14 -
   - *4.1. HLA class I genes* ..... *- 15 -*
   - *4.2. HLA class II genes* ..... *- 16 -*
   - *4.3. Nomenclature* ..... *- 17 -*
     - 4.3.1. Alternatively expressed alleles ..... - 19 -
     - 4.3.2. P- and G-groups ..... - 19 -
5. POLYMORPHISM AND POPULATIONS ..... - 20 -
6. HLA AND TRANSPLANTATION ..... - 20 -

PRACTICAL PART ..... **- 21 -**

1. INTRODUCTION ..... - 21 -
2. MATERIALS AND METHODS ..... - 22 -
   - *2.1. Materials* ..... *- 22 -*
     - 2.1.1. Population sample ..... - 22 -
     - 2.1.2. Chemicals ..... - 22 -
     - 2.1.3. Equipment and analysis software ..... - 23 -
   - *2.2. Methods* ..... *- 23 -*
     - 2.2.1. DNA purification and determination of DNA concentration ..... - 23 -
       - 2.2.1.1. Maxwell® 16 Purification System ..... - 23 -
       - 2.2.1.2. NanoDrop-ND-2000 Spectrophotometer ..... - 24 -
     - 2.2.2. DNA sequencing and HLA typing ..... - 24 -
       - 2.2.2.1. Gene-specific PCR-amplification ..... - 25 -
       - 2.2.2.2. PCR product purification ..... - 25 -

2.2.2.3. Agarose gel electrophoresis ..... - 25 -
2.2.2.4. Exon-specific fluorescent PCR ..... - 26 -
2.2.2.5. Sequencing reaction clean up ..... - 27 -
2.2.2.6. Sequencing ..... - 27 -
2.2.3. Data processing ..... - 27 -
2.2.3.1. Analysis of electropherograms ..... - 27 -
2.2.3.2. Resolving ambiguities ..... - 28 -
2.2.3.3. Resolution and nomenclature of obtained data ..... - 29 -
2.2.4. Statistical analysis ..... - 29 -
3. RESULTS AND DISCUSSION ..... - 31 -
3.1. Hardy Weinberg Equilibrium ..... - 31 -
3.2. Null-alleles ..... - 32 -
3.3. Frequent HLA alleles – low resolution data ..... - 34 -
3.4. Frequent HLA alleles – high resolution data ..... - 37 -
3.5. Frequent HLA haplotypes – low resolution data ..... - 39 -
4. CONCLUSION ..... - 41 -
REFERENCES ..... - 42 -
APPENDIX ..... - 46 -
ACKNOWLEDGEMENTS ..... - 48 -
STATUTORY DECLARATION ..... - 49 -

LIST OF FIGURES

Figure 1: Structure of MHC class I and class II proteins ..... - 13 -
Figure 2: Map of the MHC showing the HLA genes ..... - 14 -
Figure 3: Gene structure of HLA class I genes ..... - 15 -
Figure 4: The major genes within the HLA class II region ..... - 16 -
Figure 5: Gene structure of HLA class II genes ..... - 17 -
Figure 6: Picture of the Maxwell® 16 Purification Cartridge filled with reagents ..... - 24 -
Figure 7: Example of an agarose gel photo ..... - 26 -
Figure 8: Example of an electropherogram ..... - 28 -
Figure 9: Example of an HARP resolving an ambiguity ..... - 29 -
Figure 10: Comparison of the most frequent alleles of HLA-A, HLA-B, and HLA-DRB1 reported in the Austria population ..... - 36 -
Figure 11: Comparison of the most frequent three loci haplotypes ..... - 40 -
Figure 12: Location of collection sites ..... - 46 -

LIST OF TABLES

Table 1: HLA-Nomenclature (Marsh, et al., 2010) ..... - 18 -
Table 2: List of used chemicals ..... - 22 -
Table 3: Details of the exon-specific primers ..... - 23 -
Table 4: List of used equipment and analysis software ..... - 23 -
Table 5: Mastermix for gene-specific PCR ..... - 25 -
Table 6: PCR protocol for the gene-specific amplification of the coding regions on chromosome 6 ..... - 25 -
Table 7: PCR protocol for the exon-specific fluorescent PCR ..... - 26 -
Table 8: Null-alleles detected in all 1370 samples ..... - 32 -
Table 9: Frequent HLA A, HLA B, HLA C, and HLA DRB1 alleles ..... - 34 -
Table 10: Frequent three loci haplotypes (n=1370) ..... - 39 -
Table 11: Frequent four loci haplotypes (n=503) ..... - 39 -
Table 12: Complete list of HLA alleles (low resolution) ..... - 46 -
Table 13: Frequent HLA alleles (high resolution) ..... - 47 -

LIST OF ABBREVIATIONS

Abbreviation	Meaning
APC	Antigen Presenting Cells
BM	Bone Marrow
CD	Cluster of Differentiation
DNA	Deoxyribonucleic Acid
EM algorithm	Expectation-Maximization algorithm
ER	Endoplasmic Reticulum
GA	Genotype Ambiguity
GSSP	Group-Specific Sequencing Primer
HA	Heterozygous Ambiguity
HARP	Heterozygous Ambiguity Resolution Primer
HWE	Hardy-Weinberg Equilibrium
Ig	Immunoglobulin
LD	Linkage Disequilibrium
MHC	Major Histocompatibility Complex
MIC genes	MHC class I Chain Related Genes
NK Cells	Natural Killer Cells
PCR	Polymerase Chain Reaction
PMP	Paramagnetic Particles
RNA	Ribonucleic Acid
SBT	Sequence-Based Typing
SC	Stem Cell
T_C	Cytotoxic T-Cell
TCR	T-Cell-Receptor
T_H	T-Helper-Cell
UCB	Umbilical Cord Blood

THEORETICAL PART

1. INTRODUCTION

HLA (Human Histocompatibility Complex) genes code for cell-surface molecules that help the immune system to discriminate between "self" and "nonself" by presenting antigen-derived peptides to T-lymphocytes (Klein, et al., 2000).

Due to this function, these genes also play a major role in organ and hematopoietic stem cell transplantation by being the main immunological barrier for to transplantation. The graft rejection is mediated by T-lymphocytes, which react against the foreign version of the HLA-molecules on the cell surface of the donor tissue. Since HLA genes are the most polymorphic known in the human genome, it is extremely important to match donor and recipient as good as possible in the case of transplantation (Petersdorf, et al., 2001).

Extensive differences are found in the distribution of HLA alleles and haplotypes among various populations. Consequently it is important to not only know the common HLA allele and haplotypes in a population for donor selection purposes, but also to establish stem cell banks and donor registries (Clayton, et al., 1997).

2. THE MAJOR HISTOCOMPATIBILITY COMPLEX (MHC) AND ITS ROLE IN THE IMMUNE SYSTEM

In order to function properly and protect the human body from diseases, the immune system has to be able to discriminate between "self" and "nonself". Though here are various functional elements that interact with each other, immunity can be broadly categorized into innate and adaptive immunity. Both immune responses work closely together, thus a strict separation of molecules only involved in the innate or the adaptive immune response is not possible (Roitt, et al., 1995).

Innate immunity is the first immediate response to a specific molecule (antigen) on the surface of a pathogen and is subsequently necessary for the delayed adaptive immune response. The characteristics of the adaptive immunity are immunological memory and immunological specificity (Crotty, et al., 2004).

The immune response is mainly mediated by leukocytes. There exist various different types and they are all derived from multipotent stem cells, the hematopoietic stem cells. These stem cells give rise to all blood cell types and can be found in the bone marrow, the umbilical cord blood (Cairo, et al., 1997) (Lai, et al., 2008) and the peripheral blood after pre-treatment with cytokines that trigger the release of hematopoietic stem cells from the bone marrow (Körbling, et al., 2001).

2.1. The adaptive immune system

The adaptive immune system is mediated by B- and T-lymphocytes and activated by the innate immune system after exposure to a specific antigen or via immunization. Antigen presenting cells (APC) such as Dendritic Cells, macrophages, and B-lymphocytes present foreign antigens with the help of Major Histocompatibility Complex (MHC) molecules on their cell surface (Roitt, et al., 1995).

Phagocytes and Lymphocytes interact in various ways:

- Lymphocytes only respond to antigens when the innate immune response was activated first (Medzhitov, et al., 1997)
- Phagocytes present antigens with the help of MHC-molecules on their surface and active the T-lymphocytes (Loureiro , et al., 2006)
- Activated T-lymphocytes release factors that stimulate phagocytosis (Zhao, et al., 2008)

- B-lymphocytes release antibodies (soluble receptor molecules that detected the antigen) that bind to the pathogen and thus "labeling" it for phagocytosis (Zhao, et al., 2008)

2.2. T-lymphocytes

T-lymphocyte progenitors are derived from lymphoid progenitors and mature in the thymus. The two major groups of T-lymphocytes are T-helper cells (T_H) that stimulate the response of other cells, and cytotoxic T-lymphocytes (T_C), that directly kill infected cells (Lu, et al., 2009). The difference is due to the positive and negative selection during maturation and differentiation (Spits, 2002):

Every T-lymphocyte has an antigen specific T-cell receptor (TCR) to recognize the MHC-presented peptides. Before entering the thymus for maturation, T-lymphocytes lack the TCR and are "double negative" (CD4-CD8-). That means that they neither express the co-receptor for the MHC class I (CD8) nor the co-receptor for MHC class II (CD4) molecules. After expressing various other cell surface molecules and undergoing gene rearrangement of TCR-chains, they are "double positive" (CD4+CD8+) before maturing into single positive T-lymphocytes (either CD4+CD8- or CD4-CD8+).

As a result T_H (CD4+) that recognize antigens presented by MHC class II molecules and T_C (CD8+) that recognize AGs presented by MHC class I molecules leave the thymus.

A third group of T-lymphocytes, the regulatory (suppressor) T-lymphocytes, suppress the activity of other lymphocytes such as T_H and T_C in order to prevent self-tissue damage (Beissert, et al., 2006).

2.3. The Major Histocompatibility Complex (MHC)

MHC proteins were first discovered in the 1930s during the course of transplantation reactions: Tumors could be transplanted between mice which shared a group of alleles, which were named Antigen II. Mice with different Antigen II alleles rejected the transplanted tumors. Soon after, in a series of mouse crossing experiments, it was shown that the acceptance of a graft is controlled by a small number of loci that were named Major Histocompatibility loci (Thorsby, 2009).

The first human MHC proteins were reported by Dausset in the 1950s: Antibodies in human sera from multitransfused patients or multiparous women reacted with leukocytes

from many but not all individuals who were tested. Dausset concluded that antibodies in these sera detected alloantigens on human leukocytes (Dausset, 1958).

B-lymphocytes carry out the antibody response by recognizing intact antigens. T-lymphocytes on the other hand can either initiate adaptive immune responses (T_H) or help to eliminate pathogens that are inside host cells and would otherwise be missed. Furthermore, T-lymphocytes act at the site of infection directly with a "target cells" by either killing it - if it is infected - or by signaling and recruiting this cell for help with the immune response.

T-lymphocytes need APC that present the foreign pathogenic antigen on surface molecules (MHC) in order to recognize it. The antigen has been partly degraded inside the APC before it is presented on the cell surface.

There are two distinct classes of MHC-molecules:

- MHC class I proteins present foreign antigens mainly derived from the cytoplasm to cytotoxic T cells
- MHC class II proteins present mainly endocytosed plasma membrane and extracellular foreign antigens to helper T cells

All somatic cells express MHC class I proteins, although the level of expression varies depending on the tissue. On contrast, MHC class II proteins are only expressed by APC.

Due to the fact that these APC not only have MHC class II, but also class I surface proteins, T cells have another way of recognizing the target cell. Each class of T cells also expresses a co-receptor that recognizes a distinct, invariant part of the appropriate MHC protein: CD4 for helper T cells and CD8 for cytotoxic T cells (see 2.2).

Both MHC proteins - as shown in figure 2 are transmembrane heterodimers with an extracellular N-terminal antigen-binding site.

3. MHC PROTEINS

3.1. MHC class I proteins

Class I MHC proteins consist of an α-chain, that has five domains: two peptide-binding domains (α_1 and α_2), one immunoglobulin-like domain (α_3), the transmembrane region, and the cytoplasmic tail (see Figure 1). These domains are encoded by separate exons within the MHC gene cluster (see chapter 4). The protein also consists of an invariant β_2-microglobulin, which is not encoded within the MHC and does not span the membrane (Bodmer, 1987).

The α1 and α2 domains form the cleft responsible for binding of self and nonself proteins. The α3 domain is necessary in forming molecular bonds with β_2-microglobulin, which is essential for the folding of the molecule on the cell surface.

The MHC class I peptide binding site consist of a groove, that can bind peptides about 8-10 amino acids long. The terminal amino group and the terminal carboxyl group of the degraded antigen are bound to invariant amino acids at both ends of the groove. These invariant binding sites recognize features of the peptide backbone that are common to all peptides. The side chains of the antigen are either bound to polymorphic amino acids or pointed out to be recognized by TCRs. The polymorphic amino acids in the groove ensure that each allelic form binds and presents a distinct characteristic set of peptides thus guaranteeing that a broad range of antigens can be presented (Natarajan, et al., 1999).

MHC class I proteins present two kinds of proteins: Either defective proteins in the cytosol or proteins that origin from viral mRNA. Both kinds are degraded by barrel-shaped structures, the proteasomes, and then transported into the endoplasmic reticulum (ER). There the degraded proteins are loaded onto MHC class I proteins and transported to the cell surface for presenting (Ljunggren, et al., 1996).

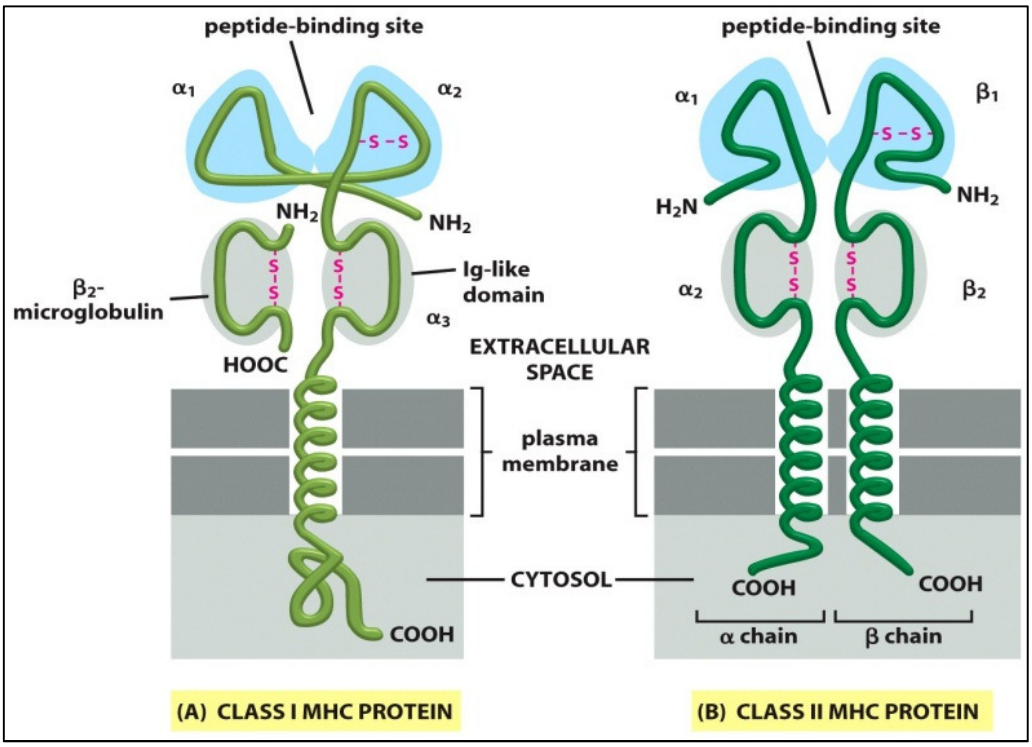


Figure 1: Structure of MHC class I and class II proteins
MHC class I molecules (A) consist of five α-chain domains and one β_2-microglobulin, while MHC class II molecules (B) consist of an α- and a β-chain (Alberts, et al., 2008).

3.2. MHC class II proteins

The class II MHC proteins consist of an α- and a β-chain, encoded within the MHC gene cluster. As seen in figure 2, each chain has four domains: the peptide-binding domain (α_1 + β_1), the immunoglobulin-like domain (α_2 + β_2), the transmembrane region, and the cytoplasmic tail (Bodmer, 1987).

The peptide-binding groove of the MHC class II peptides can present antigens that have been degraded to 12-20 amino acids. Another difference to MHC class I molecules is that the presented peptide is held by interactions with invariant amino acids along the length of the groove and can therefore present a more heterozygous set of peptides.

The antigen processing pathway also differs from class I: Extracellular proteins (self or foreign) are taken up by endocytosis or phagocytosis and loaded into endosomes. Those endosomes fuse with lysosomes that are loaded with MHC class II molecules (previously synthesized in the ER). Enzymes from the lysosome degrade the protein so the peptides can be loaded onto the class II molecules. The MHC-protein-complexes are then exported via exocytosis to the cell surface to present the antigen (Creswell, 1994).

4. HLA GENES OF THE MHC

In humans MHC is called Human Leukocytes Antigen (HLA) after its discovery on a human leukocyte by Jean Dausset in 1958 (Dausset, 1958). The genes of the MHC are located on the short arm of chromosome 6 in region 6p21.3 (see Figure 2). In addition to the HLA genes, the MHC contains over 200 genes (non-HLA genes) – some of which also play a role in the immune system. For many of these genes the function is yet to be known. (Klein, et al., 2000).

As already mentioned in chapter 2.3 the HLA genes can be divided into two classes (I and II) and are structurally and functionally different. In addition to that, there also exists a region III, that contains genes that code for soluble factors involved in the immune response (e.g. cytokines) (Horton, et al., 2004).

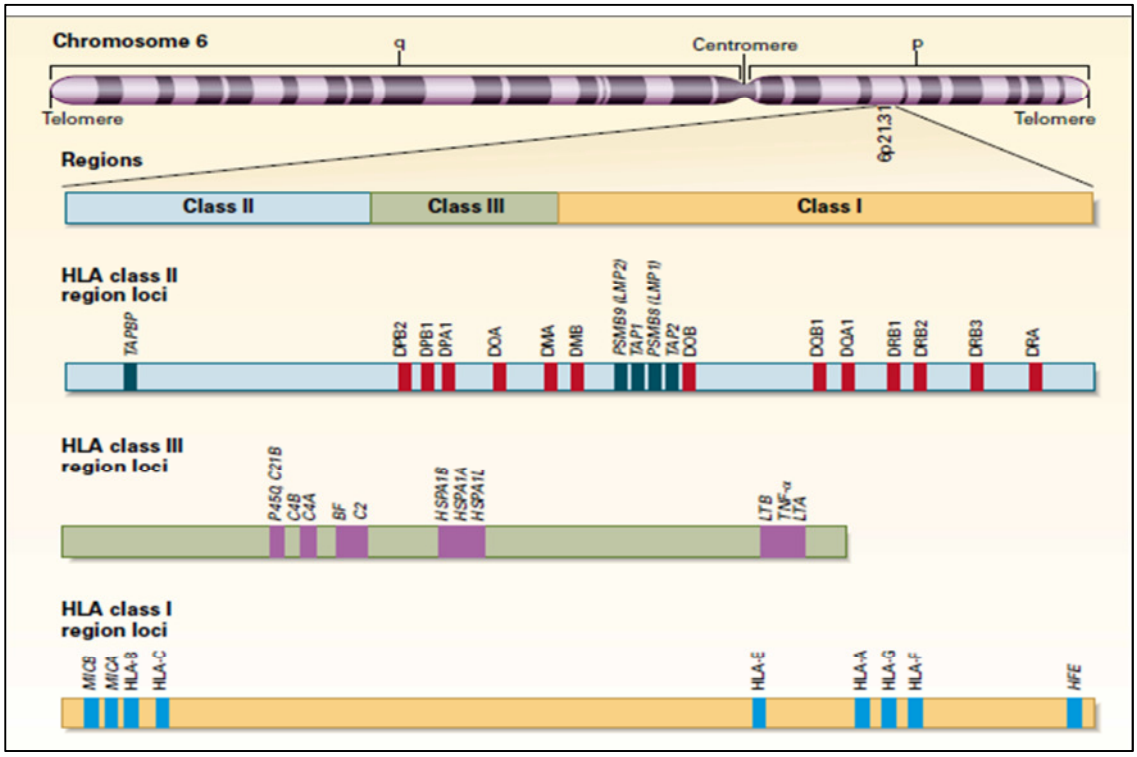


Figure 2: Map of the MHC showing the HLA genes
On top the broad structure of chromosome 6 is shown – the HLA genes are located in the region 6p21.3. Beneath a simplified map of the class 1, 2 and 3 region of the MHC is shown (Robinson, et al., 2011).

4.1. HLA class I genes

HLA class I genes are located on the region I in the MHC gene cluster on chromosome 6. They code for HLA cell-surface molecules that present endogenous proteins to cytotoxic T-lymphocytes triggering an immune response (Klein, et al., 2000).

The most important genes for the immune response – and therefore also called "classical HLA-genes" - in the HLA class I region are HLA A, HLA B, and HLA C. They are the most polymorphic in class I and expressed on all nucleated cells.

The classical HLA class I gene consists of 6 exons (see Figure 3): The first exon initiates transcription. Exon 2 and 3 encode the α1- and α2-chains which together form the antigen binding site, where the presented peptide binds. Exon 4 encodes the α3-chain which is responsible for the folding of the cell-surface molecule. The remaining two exons code for the transmembrane portion and the cytoplasmic tail. This explains why exon 2, 3, and 4 are the most important for HLA-typing, because they encode parts that are responsible of the make-up of the molecule, consequently the immune response and are the most polymorphic (Horton, et al., 2004).

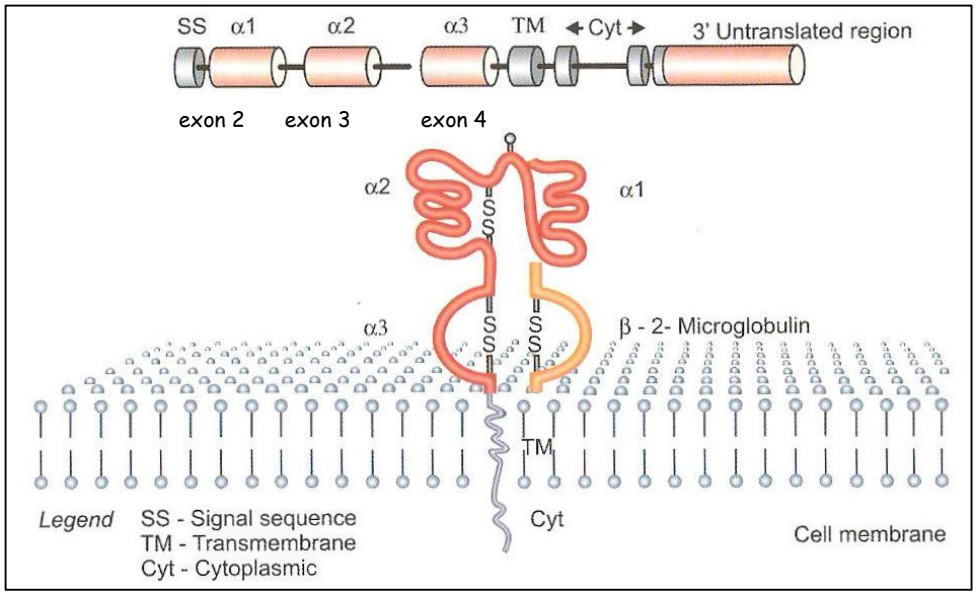


Figure 3: Gene structure of HLA class I genes
The figure shows the gene-structure relationship for classical HLA class I proteins. The exons are shown as cylindrical segments and the introns are shown as black lines. The exons are labeled to indicate the corresponding domain of the class I molecule for which they code. Extracellular domains of the protein are shown in red, the β2-microglobulin is shown in orange (Mehra, et al., 2010).

Non classical class I HLA genes are HLA E, F, G and the nonexpressed HLA H, J, K, and L. They are less polymorphic, have limited expression and play various roles in the innate and adaptive immune system.

In addition, there are also a series of genes with class I like structure in the MHC, which become more and more important for HLA-matching for transplantation:

- MIC genes (MHC Class I Chain Related Genes; (Collins, 2004):
  There are seven MIC genes, but only two (MIC-A and MIC-B) code for functional products. MIC-A and MIC-B are upregulated under cellular stress and play a role during the innate immune response.
- HFE genes (Human Hemochromatosis Protein; (Feder, et al., 1996):
  They have a non-immunological function and regulate iron absorption.

Both non classical class I HLA genes and genes with class I like structure are not included in the HLA analysis in this master thesis.

4.2. HLA class II genes

HLA class II genes are located on the region II in the MHC gene cluster on chromosome 6. They code for HLA cell-surface molecules that present exogenous proteins to helper T-lymphocytes triggering an immune response (Klein, et al., 2000).

The HLA class II region consists of three "classical" subregions: DR, DQ, and DP. These genes share a basic common structure, but have diverse sequences. Besides that there are also non-classical HLA genes like HLA DM/DO, which play a role in the antigen presentation.

The DR subregion consists of 10 (DRA, DRB1-9), the DQ subregion consists of five (DQA1-2, DQB1-3), and the DP subregion also consists of five (DPA1-2, DPB1-3) genes. In every subregion there are nonexpressed genes due to gene duplication and mutations during the evolution of the MHC.

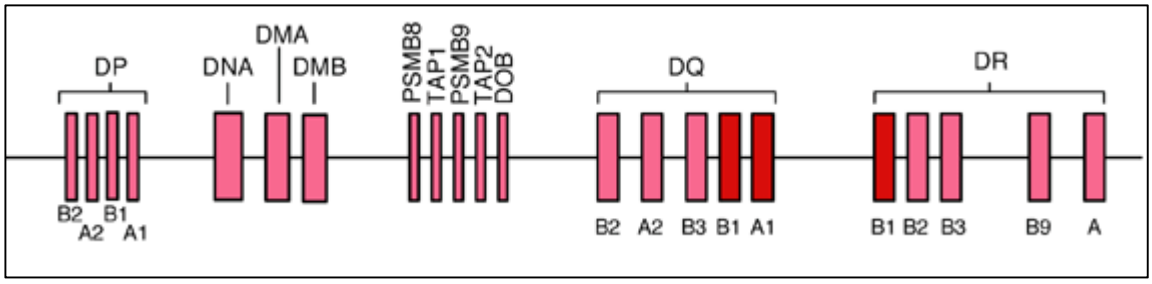


Figure 4: The major genes within the HLA class II region
The figure shows that the class II region is divided into HLA DP, DN, DM, DO, DQ, and DR. For DP only 4 out of 5 genes (DPB2, DPA2, DPB1, and DPA1) are shown. For DQ all existing 5 genes are shown and for DR also only 5 (DRA, DB1, DRB2, DRB3, and DRB9) out of ten genes are displayed (Brand, et al., 2005).

All class II genes have five exons: the first codes for the transcription initiation, the second and third for two extracellular domains, the forth for the transmembrane part of the molecule and the fifth for the cytoplasmic tail. Since class II molecules are heterodimers, there is an α- and a β-chain with corresponding domains (see Figure 5). Similar to HLA class I genes, not all exons are of interest for HLA-typing: usually only exon 2 of both chains, which encodes for the peptide-binding site - and therefore the most polymorphic - is sequenced (Horton, et al., 2004).

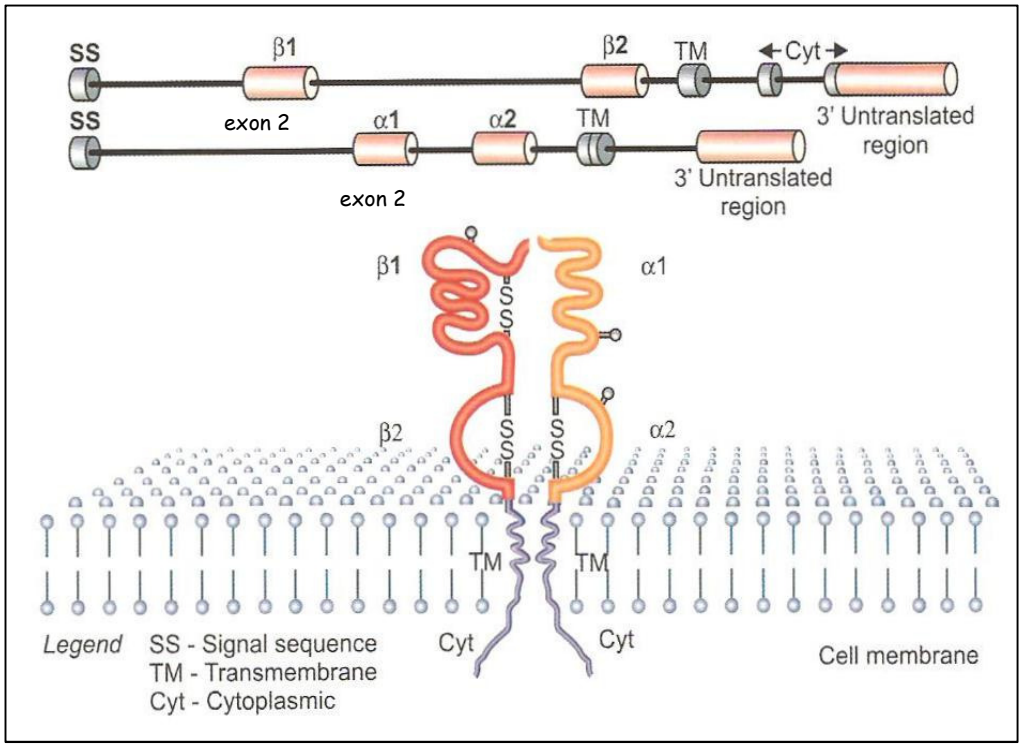


Figure 5: Gene structure of HLA class II genes
The figure shows the gene-structure relationship for classical HLA class II proteins. As in figure 4 the exons are shown as cylindrical segments and the introns are shown as black lines. The exons are labeled to indicate the corresponding domain of the molecule for which they. The α-chain is shown in red, the β-chain displayed in orange (Mehra, et al., 2010).

One of the most important features of HLA class II is that there is a variation of the number of genes which comprise observed haplotypes, which adds to the polymorphism of this region. The only heterodimer present on all haplotypes is the DRA/DRB1 molecule which is often referred to as "DRB1" (Guardiola , et al., 1996).

4.3. Nomenclature

The first HLA typing methods were based on a serological lymphocytotoxicity assay, where found antigens where given numbers as they were discovered. The prefix to those

numbers was the locus the antigen belonged to (e.g.: B5). After the HLA-typing methods were further refined and the full extent of the polymorphism of this gene regions was discovered, the nomenclature changed completely (Thorsby, 2009):

To illustrate the nomenclature it is explained step by step in Table 1. The left column shows an example of the nomenclature of an allele, where in the right column the example is explained in the corresponding row. The explained part of the allele-nomenclature is highlighted in green.

Table 1: HLA-Nomenclature (Marsh, et al., 2010)

Example	Explanation
HLA-A	Every HLA allele is given a prefix which refers to the gene locus (e.g. locus A).
HLA-A	The hyphen between "HLA" and the locus is used if an allele is described.
HLA-DRB1	The gene-loci of the class II genes consist of three letters: the first (D) indicates class II, the second (M, O, P, Q, or R) the subregion and the third (A or B) the chain (α or β) (Klein, et al., 2000).
HLA-A*	The locus is followed by an asterisk to show that the HLA-typing was not done by serological methods but by sequencing.
HLA-A*02	Position1: The first two digits indicate the serological group to which the sequence belongs, i.e. a group of alleles which encode this specific antigen
HLA-A*02:	A colon between numbers indicates a new position in the nomenclature that has a different meaning than the one before, i.e. with every new position the allele is identified more thoroughly and therefore "higher resolved".
HLA-A*02:101	Position 2: These digits show the specific HLA protein of the previously identified allele group. Since those numbers have been also designated as they have been discovered and there are currently more than 99 different alleles for certain serological groups, this position can consist of two or three digits.
HLA-A*02:101:01	Position 3: If there is a non-synonymous difference between alleles (difference in the coding region), they are given a different number at this position.
HLA-A*02:101:01:02	Position 4: Differences outside the coding region or in intron sequences are indicated in this position
HLA-A*02:101:01:02N	Alternatively expressed allele are given a suffix (see chapter 4.3.1)

Depending on the number of positions defined in the allele-nomenclature, allele-data can be either presented at a "low-" or a "high-resolution" level.
Low-resolution or "2-digit resolution" means, that only the serological allele group is specified (e.g.: HLA-A*02).
High-resolution or "4-digit resolution" or "resolution at a subtype/allelic level" is defined by also specifying the next position (e.g.: HLA-A*02:101). In order to higher resolve an allele, reported ambiguities must be excluded (see chapter 2.2.3.2 Resolving ambiguities).

4.3.1. Alternatively expressed alleles

There are various suffixes that are designated in order to indicate alternatively expressed alleles (Robinson, et al., 2011):

- N: null-alleles are not expressed at all
- L: the allele encodes a protein with significantly reduced or "low" cell surface expression compared to normal levels
- Q: the allele has a mutation that has a significant effect on the expression, but it is not yet confirmed and therefore "questionable"
- S: the allele is expressed as a soluble "secreted" molecule and not present on the cell surface
- C: the allele codes for a molecule that is present in the "cytoplasm" but not on the cell surface
- A: the allele has an "aberrant" expression where there is doubt as to whether the protein is expressed ort not

4.3.2. P- and G-groups

In order to define HLA alleles which are not resolved to the level of a single allele (position three or four) there is the P and G suffix (Robinson, et al., 2011).
P refers to "peptides/proteins" and means that all alleles which have a nucleotide sequence that encodes for the same peptide binding domain are summarized with this suffix following the allele of the group with the lowest allele number, i.e. these alleles have the same protein sequence in the antigen binding site (e.g.: the group of HLA-A*01 alleles that share the same peptide sequence as HLA-A*01:01:01:01 are summarized as HLA-A*01:01P).

G groups summarizes groups of alleles with not only have the same peptide sequence in the antigen binding domain, but also identical nucleotide sequences encoding this part of the molecule (e.g.: HLA-A*24:02:01G).

5. POLYMORPHISM AND POPULATIONS

The human HLA class I and class II genes are the most polymorphic observed in the whole genome. This is important because they need to present pathogen derived peptides to initiate an immune response. Pathogens are constantly undergoing mutations and the human immune system (and therefore the HLA genes) needs to adapt in order for the human race to survive the changing environment.

This constant ability of the HLA genes to mutate according to the changing circumstances is the reason why there are immense differences between various populations concerning the frequency of HLA-alleles. It can be explained by historical migration but also different selection patterns due to pathogen variations a population is exposed to.

6. HLA AND TRANSPLANTATION

HLA antigens are the most important transplantation antigens, both for solid organ transplantation as well as hematopoietic stem cell transplantation. During their development, the molecules of the immune system learn to discriminate between self and nonself. This happens due to the self-HLA antigens expressed on the cell surface. In case of a transplantation, where the cells of the graft are marked with nonself-HLA-molecules, a severe rejection reaction follows. Therefore a transplantation between a fully HLA-matched donor and recipient in the important loci (HLA A, B, and C for class I and HLA DRB1 for class II) is necessary for the success of this treatment. In this case the immune system recognizes the transplanted tissue as “self” (Petersdorf, et al., 2001).

PRACTICAL PART

1. INTRODUCTION

HLA molecules play an important role not only in the immune response but also in transplantations. Since this region of the human genome is the most polymorphic and show immense differences between populations, it is important to obtain HLA-data for each population (Jin, et al., 2003).

The aim of this master thesis is to present extensive HLA-data for the Austrian population, since in the past, only a few studies have described HLA data for this country (Schipper, et al., 1997) (Gonzales-Galarza, et al., 2011).

1370 unrelated Austrian umbilical cord blood samples were typed for HLA A, HLA B, and HLA DRB1. Frequent alleles and haplotypes were listed. HLA C allele frequencies were investigated in a subgroup of 503 samples. For a subpopulation of 100 donors all ambiguities were resolved on the subtype level. The data was compared to previously published data of Austrian bone marrow and blood donors (Schipper, et al., 1997) (Gonzales-Galarza, et al., 2011).

The Austrian population was about 8.40 million in 2011. Although the population is relatively homogenous, about a sixth of the Austrian population has been born abroad and/or holds foreign citizenships. The majority of migrants originate from Germany. The second largest group was born in Serbia, Kosovo or Montenegro, followed by people of Turkish origin and migrants from Bosnia and Herzegovina (Statistik Austria, 2011).

On the basis of these facts it was also examined if the results regarding most frequent haplotypes in the Austrian population are comparable to the results presented in a large German population study (Müller, et al., 2003).

2. MATERIALS AND METHODS

2.1. Materials

2.1.1. Population sample

HLA A, HLA B, and HLA DRB1 PCR-SBT was performed on 1370 unrelated randomly selected Austrian cord blood donors. The cord blood was collected during a period from 2009 to 2010 at 49 different private and public hospitals in Austria. A complete list and a figure of the various collection sites are provided in the appendix.

HLA C typing was performed on a subgroup of 503 of these donors. No information of ethnic group or country of birth was obtained from any of the donors. The umbilical cord blood samples were stored at 4 °C and the DNA was purified within 14 days of collection at the Laboratory of Genetic Analysis, Vivocell Biosolutions, Graz, Austria.

2.1.2. Chemicals

All used chemicals are listed below:

Table 2: List of used chemicals

Chemicals	Source
Maxwell 16® Purification Kit	Promega
Allele SEQR HLA-A, -B, -C, and -DRB1 Kombi Kit	Abbott
10x TBE-Buffer	Roth/Lactan
ddH_2O	Roth/Lactan
LE Agarose	Biozym
Gel Red DNA Stain 10,000x	Biotium
Quantitas DNA Marker (300 bp-2 kb)	Biozym
Gel Electrophoresis Buffer	Roth/Lactan
Sephadex G-50	Sigma-Aldrich

The used primers for the PCR-amplifications are included in the Allele SEQR Kombi Kits. No data concerning the primer sequence is provided by the manufacturer. All kits include locus-specific primers for HLA A, B, C, and DRB1 as well as exon-specific primers in both forward and reverse direction.

The exon-specific primers are located in the introns for complete sequence coverage. Exceptions are the exon-specific primers for DRB1: the forward primer is located in the intron, the reverse primer is located in the exon.

Details for the locus-specific primers are presented in Table 3. For HLA A and HLA DRB1 a primer for single-tube locus specific PCR is provided, for HLA B and HLA C primers for a single-tube duplex PCR are included in the kit.

Table 3: Details of the exon-specific primers

Locus	Range	Size PCR-product
HLA A	exon 1 to intron 4	2 kb
HLA B (single tube duplex PCR)	exon 1 to intron 3 intron 3 to exon 7	1,5 kb 1,2 kb
HLA C (single tube duplex PCR)	exon 1 to intron 3 intron 3 to exon 7	1,4 kb 1,2 kb
HLA DRB1	exon 2	300 bp

2.1.3. Equipment and analysis software

All equipment and analysis software used is listed in Table 4:

Table 4: List of used equipment and analysis software

Equipment	Manufacturer
Maxwell 16® System	Promega
NanoDrop-ND-2000	Thermo Fisher Scientific
Thermocycler DNA Engine Dyad	Bio Rad
Centrifuge 5810R	Eppendorf
Electrophoresis Chamber	Single Cell
Image Master VDS-CL	Amersham Biosciences
Gel Doc™ 2000	Bio Rad
TDS Quantity One Software	Bio Rad
3500xL Dx Genetic Analyzer	Applied Biosystems
Assign-SBT™ Sequencing Software	Conexio Genomics
ARLQUIN Software	Institute of Ecology and Evolution, University of Bern

2.2. Methods

2.2.1. DNA purification and determination of DNA concentration

2.2.1.1. Maxwell® 16 Purification System

The DNA of the umbilical cord blood samples was isolated with the Maxwell® 16 Purification System. The system uses cartridges with prefilled reagents in wells. First 300 µL of a whole blood sample were mixed with a lysis puffer, a chaotropic agent to disrupt the structure and denaturize the DNA, and a detergent. Then the nucleic acids were bound to magnetized silica particles (Paramagnetic Particles PMPs) on a plunger to selectively capture and release the genomic DNA through the following steps. Finally impurities like other cellular components were washed away during several wash-steps and the DNA was suspended in 300 µl elution puffer for further processing.

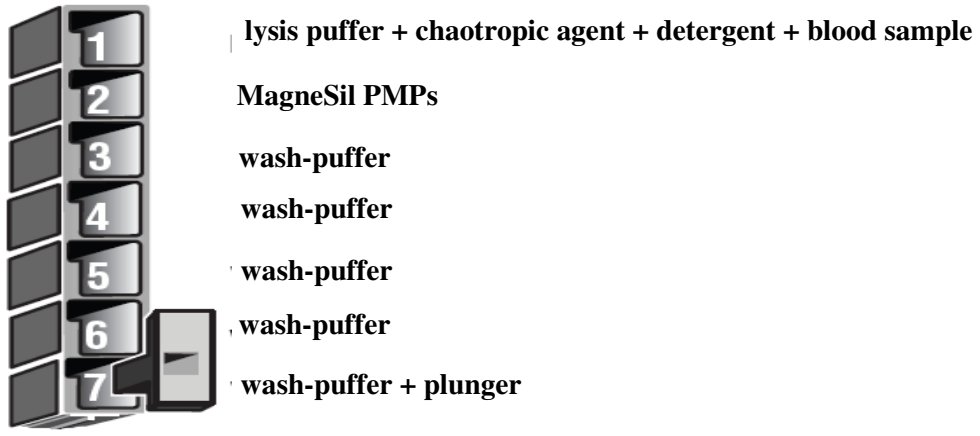


Figure 6: Picture of the Maxwell® 16 Purification Cartridge filled with reagents.
Well 1 is filled with a lysis puffer, a chaotropic agent and a detergent. The blood is pipetted into well 1. Well 2 is filled with MagneSil particles, which are taken up by the plunger. The plunger is first located in well 7 and then taken up and moved to well 2. There it takes up the MagneSil particles with which the DNA is transported from well 1 to well 3 to 7 (wash steps). The isolated DNA is finally moved to well 7 where is suspended in elution puffer.

2.2.1.2. NanoDrop-ND-2000 Spectrophotometer

1 µL of the purified DNA suspended in elution puffer was pipetted onto a measurement pedestal, which has a fiber optic cable embedded. A second fiber optic cable is brought into contact with the sample causing the liquid to bridge the gap between the two fibers. A spectrometer analyzes the light (xenon flash lamp) going through the sample with a CCD array.

For the following amplification and sequencing steps a concentration of 40-80 ng of genomic DNA was needed. The concentration was determined with the NanoDrop-ND-2000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, Delaware, USA). The concentration was measured at a wavelength of 260 nm.

Additionally the ratio of absorbance at 260 nm and 280 nm is measured. It shows the purity of the DNA sample. For our experiments we accepted a ratio between 1.8 and 2.0, a lower ratio indicates impurities and the sample was not used for further analysis.

2.2.2. DNA sequencing and HLA typing

The DNA sequencing step was performed with the AlleleSEQR Sequencing Kombi Kit (Abbott, Illinois, USA) according to manufacturer's protocol:

2.2.2.1. Gene-specific PCR-amplification

Gene-specific PCR-amplification of the coding regions on the chromosome 6 was performed for HLA A, HLA B, HLA C (class I) and HLA DRB1 (class I). The mastermix for each locus was prepared for 100 reactions respectively, according to Table 5. Every necessary reagent was already mixed and provided in the kit in a single-tube, only Taq-I-polymerase was added. For the PCR mix isolated DNA was added to the mastermix according to Table 5.

Table 5: Mastermix for gene-specific PCR

Locus	mastermix for 100 reactions		PCR mix for 1 reaction	
	single-tube reagents	**polymerase**	**mastermix**	**isolated DNA**
HLA A	1600 µL	30 µL	16 µL	4 µL
HLA B	1600 µL	30 µL	16 µL	4 µL
HLA C	1600 µL	30 µL	16 µL	4 µL
HLA DRB1	800 µL	10 µL	8 µL	2 µL

PCR was performed according to the following protocol:

Table 6: PCR protocol for the gene-specific amplification of the coding regions on chromosome 6

temp [°C]	time	no of cycles
95	10 min	1
96	20 s	36
60	30 s	
72	3 min	
4	∞	

2.2.2.2. PCR product purification

For PCR product clean-up Exo-SAP-IT provided in the sequencing kit was used. It contains exonuclease and alkaline phosphatase. 3 µL of the ExoSAP were added to each PCR product. The samples were incubated at 37 °C for 30 min followed by 30 min at 80 °C. After cooling the samples down to 4 °C and mixing them with 20 µL nuclease-free water, they were either stored at -20 °C or used immediately for further processing.

2.2.2.3. Agarose gel electrophoresis

An agarose gel electrophoresis was performed to control the PCR-product quality:
500 mg agarose was added to 50 mL 1x TBE buffer (1 % agarose gel), melted and 5 µL GelRed 10,000x was added. 5 µL of each DNA amplicon sample and 5 µL standard (DNA marker) were loaded onto the gel and analyzed. For the electrophoresis 700 mL 1x TBE running buffer was used and 90 V were applied. The gels were analyzed with the

Image Master VDS-CL (Amersham Biosciences, Freiburg, Germany) and the corresponding software.

The agarose gel electrophoresis was not used for quantifying the DNA - the sole existing of a DNA band in each amplicon slot was proof for an effective gene-specific amplification of the coding regions.

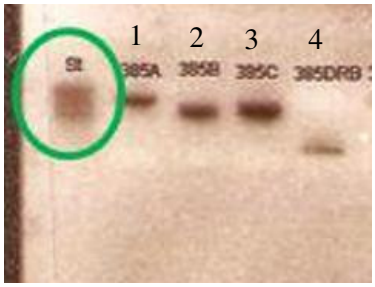


Figure 7: Example of an agarose gel photo
The figure shows a highly enlarged part of an agarose gel photo.
The standard band is on the very left side (highlighted in green) - the corresponding bands of the class I amplicons (HLA A, HLA B, and HLA C) as well as the band of the class II amplicon (HLA DRB1) are easily visible. The HLA A (marked as 1), HLA B (marked as 2), and HLA C (marked as 3) bands are visible at approximately 2kb, the HLA DRB1 band (marked as 4) is visible at approximately 300bp. The reason for the shifted HLA DRB1 bands is that less DNA was used for amplification (4 µL of class I and 2 µL of class II).

2.2.2.4. Exon-specific fluorescent PCR

In the next step the amplicons underwent an exon-specific fluorescent amplification using BigDye® Terminator chemistry: For class I exon 2, 3, and 4 and for class II exon 2 were amplified with a forward and a reversed primer, respectively. This prevents peak shifts that can occur with gene-specific amplification at heterozygous positions (Blasczyk, 2003).

The four ddNTPs, which terminate the chain after being integrated, are labeled with different fluorescent dyes, which emit light at distinct wavelengths.

Before performing the PCR, the wells in the sequencing plates were preloaded with 8 µL of the corresponding sequencing mix (provided for each forward and reversed primer of each exon in the sequencing kit). Then 2 µL of the previously prepared amplicons were added. PCR was performed according to the protocol shown in table 4:

Table 7: PCR protocol for the exon-specific fluorescent PCR

temp [°C]	time	no of cycles
96	20 s	25
50	30 s	
60	2 min	
4	∞	

2.2.2.5. Sequencing reaction clean up

The PCR products of the sequencing reaction were cleaned up via gel filtration using Sephadex G-50 (Sigma Aldrich, St. Louis, Missouri, USA) according to standard operating protocols.

2.2.2.6. Sequencing

The PCR products were mixed with 10 µL ddH_2O and then sequenced with the 3500xl Dx Genetic Analyzer (Applied Biosystems, Life Technologies, Carlsbad, California, California, USA) via Sanger capillary electrophoresis resulting electropherograms:
First, the capillaries are filled with the sample. Then a high electric field is created between an anode buffer and the negative voltage applied to the load header of the capillary array. The negatively charged DNA is pulled by the electric field through a separation polymer.
To ensure optimal separation and maintain denaturation of the sample, the capillaries are thermally controlled in an oven and in the detection cell. Smaller fragments migrate faster than larger fragments and therefore reach the detector first.
In the detector the fluorescent dyes are excited by a laser beam and emit a larger wavelength light in all directions. The emitted light is captured in the instruments optics and imaged onto a CCD array.

2.2.3. Data processing

2.2.3.1. Analysis of electropherograms

The electropherograms obtained from the sequencing step were analyzed using Assign SBT (Conexio Genomics, Fremantle, WA, Australia) software.
The resulting sequences of the performed exon-specific PCR with both the forward and the reversed primer were imported for each locus (HLA A, HLA B, HLA C, and HLA DRB1) separately.
Various steps are done with the assistance of the software:

- base calling (analysis of the nucleotide sequence)
- calculation of consensus sequence and consensus BCS (base call scoring - i.e. applying an algorithm including the obtained data to assign a "confidence score" to each consensus base)
- reverse complementing of appropriate sequences
- alignment with complementary strand

- splicing of intron sequence
- connection of sequences from multiple exons
- allele assignment by matching the consensus sequence within a library

Although most of the sequence analysis was done by the program, a confirmation of the base calls (especially with homozygous and heterozygous sequences) and a manual sequence editing (low quality peaks, mismatched positions against selected alleles from the library) was needed to gain correct HLA alleles.

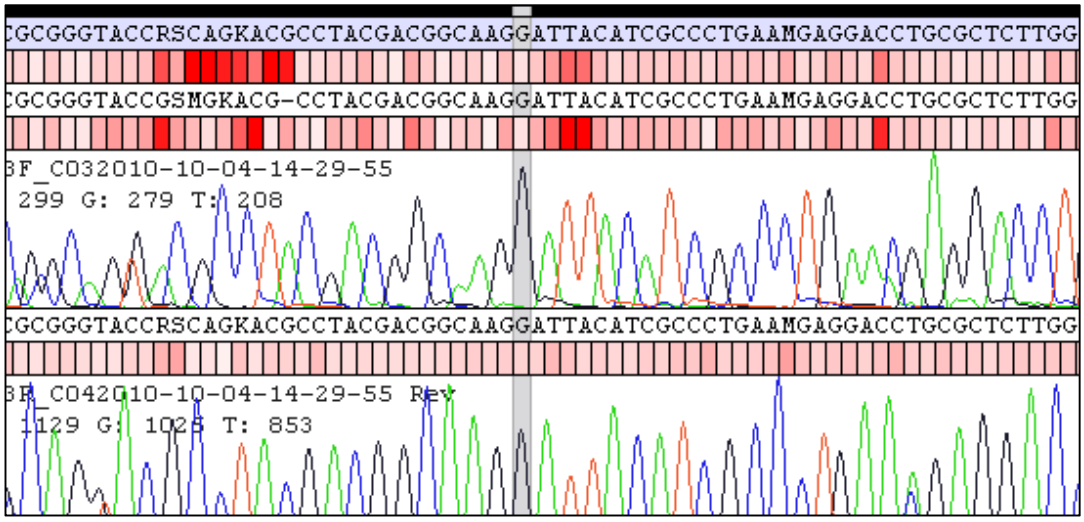


Figure 8: Example of an electropherogram
Figure 8 shows an example of an electropherogram. The peaks at the bottom show the sequence of the reverse primer, on top in white are the bases assigned to the peaks. The peaks above that show the forward primer and the assigned bases. The base-sequence on the very top (highlighted in purple) is the consensus sequence. The base highlighted in grey is marked for confirmation.

2.2.3.2. Resolving ambiguities

For 100 donors all ambiguities were resolved on a subtype level, i.e. 4-digit resolution. Generally speaking, two types of ambiguities can occur:

- genotype ambiguities (GA), where alleles differ outside the region sequenced
- heterozygous ambiguities (HA), where two or more pairs of allele combinations result in an identical heterozygous sequence

GA can only be resolved by sequencing additional regions (e.g. exon 7) – this was not routinely done. All samples, which could not be resolved at an allele group level (two-digits) without sequencing other exons than exon 2, 3, and 4 for class I and exon 2 for class II were excluded from the data pool. The exceptions are all HLA C alleles resolved at an allelic level, where this was necessary.

To resolve HA group-specific sequencing primers (GSSP) - also called heterozygous ambiguity resolution primers (HARP) - provided in the corresponding Allele SEQR HLA kits, were used. The appropriate HARPs to use were suggested by the software after the first analyzing step - all sequencing steps from exon-specific PCR on were performed with them. The HARPS are directed to motifs that discriminate the alleles involved in HA and result in a hemizygous sequence.

The so obtained additional electropherogram was loaded to the already existing data in order to gain a higher resolution.

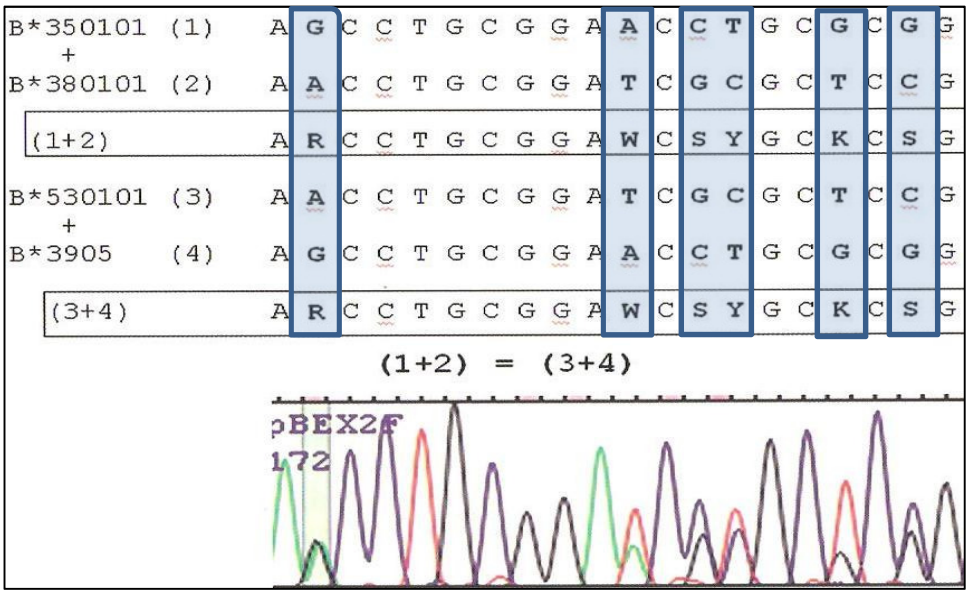


Figure 9: Example of an HARP resolving an ambiguity
Figure 9 shows an example of an HARP resolving an ambiguity: The difference between the allele combination at the bottom and the one on top are highlighted in blue. Both combinations result in the same consensus sequence. The appropriate hemizygous HARP for this part of the sequence excludes one allele combination by ruling it out.

2.2.3.3. Resolution and nomenclature of obtained data

1370 samples were resolved at an allele group level (two-digit form), i.e. describing the type which often corresponds to the serological antigen carried by an allotype (Marsh, et al., 2010).

100 samples were resolved at the allelic level (subtype level, 4-digits), i.e. describing a specific HLA protein, described by P-groups (see chapter 4.3 HLA nomenclature).

2.2.4. Statistical analysis

All statistical analysis was performed with the help of Dr. Gudrun Schappacher-Tilp, Department for Mathematics and Scientific Computing, Karl-Franzens-University Graz.

For all test the ARLQUIN software (Institute of Ecology and Evolution, University of Bern, Switzerland; http://cmpg.unibe.ch/software/arlequin35) was used.

Maximum-likelihood estimates for HLA A, HLA B, HLA C, and HLA DRB1 allele frequencies were calculated. Three-locus (HLA A, HLA B, HLA DRB1) and four-locus (HLA A, HLA B, HLA C, HLA DRB1) haplotype frequencies were estimated using an expectation-maximization (EM) algorithm. Only haplotypes with a minimum frequency of 10^{-5} are presented.

Hardy-Weinberg exact tests were performed for each of the four HLA loci and linkage disequilibrium analysis was performed for pairs of loci (Guo, et al., 1992).

3. RESULTS AND DISCUSSION

3.1. Hardy Weinberg Equilibrium

Hardy-Weinberg exact tests, that predict the expected genotype frequencies, were performed on each of the four loci. The observed and the expected frequencies are compared – if there is no significant deviation, the population is said to be in Hardy-Weinberg Equilibrium (HWE) (Guo, et al., 1992).

No significant deviations from HWE in loci HLA B, HLA C and HLA DRB1 were detected, i.e. the null hypothesis (population is in HWE) was not rejected.

However, the analysis revealed a significant deviation at locus HLA A. These results are in accordance to a study performed by Maiers et al (Maiers, et al., 2006) who also found significant values at locus HLA-A in a population of European Americans. Since a deviation from HWE was only observed in one locus and due to the sample size only a limited influence on haplotype frequency estimations based on an expectation-maximization algorithm is expected (Fallin, et al., 2000).

3.2. Null-alleles

Null alleles describe alleles for which no HLA molecule is expressed. Since HLA-typing reports are usually given with more than one allele combination possible (ambiguity), it is important to identify and – if possible - exclude them from the data before entering HLA-alleles in stem cell registries. This is especially important in the case of hematopoietic stem cell transplantation. In the worst case, the allele combination including a null allele cannot be excluded and the patient is transplanted with this stem cell sample assuming an HLA-match. If the null-allele combination is actually the one encoded in the genes of the graft, the recipients T-lymphocytes do not recognize the graft's HLA-molecules, it is misidentified as "non-self" and it can lead to a rejection of the graft (Smith, et al., 2005).
A complete list of the current null-alleles identified can be found at the IMGT/HLA database (Robinson, et al., 2011).
In the presented data nine different null-alleles were detected in approximately 30 % of all samples in the HLA A, B, and C locus. No null-alleles were detected in HLA DRB1. The description for all mutations that cause the occurring null-alleles also found at the IMGT/HLA database. According to this data the respective intron sequences and in a few cases additional exons were analyzed. This lead to the exclusion of seven null-allele combinations in every sample they occurred, one could not have been excluded in any sample, and one was excluded in a few samples:

Table 8: Null-alleles detected in all 1370 samples

Locus	Null-alleles	Comment
HLA A	HLA-A*01:01:01:02N	excluded through analysis of intron sequences in every occurring sample
	HLA-A*03:01:01:02N	excluded through analysis of intron sequences in every occurring sample
	HLA-A*29:01:01:02N	excluded through analysis of intron sequences in every occurring sample
	HLA-A*68:11N	excluded through analysis exon 1 in every occurring sample
HLA B	HLA-B*15:01:01:02N	excluded through analysis of intron sequences in every occurring sample
	HLA-B*18:17N	excluded through analysis exon 1 in every occurring sample
	HLA-B*44:19N	never excluded
HLA C	HLA-C*03:20N	excluded through analysis exon 1 in every occurring sample
	HLA-C*04:09N	excluded in samples were exon 7 was analyzed

For statistical reasons the null alleles were not excluded from the data presented in this master thesis, mainly because they don't change the frequency of allele or haplotypes. Nevertheless, these results play a significant role in the further treatment of the data especially in regard to uploading them to stem cell registry database. In this case all samples, where null-alleles cannot be excluded, are not used for hematopoietic stem cell transplantation. In total 290 of the 1370 samples (21.17 %) were eliminated before entering the database. No comparable data was reported in other Austrian population studies.

3.3. Frequent HLA alleles – low resolution data

Only 17 of 965 HLA A, 29 of 1543 HLA B, 13 of 626 HLA C, and 13 of 762 HLA DRB1 alleles were found among World Health Organization (WHO)-recognized HLA A, HLA B, HLA C, and HLA DRB1 alleles, respectively (Marsh, et al., 2010). This can be partly explained by the limited number of samples. It is known that relatively few common alleles are expressed by a great majority of people in individual ethnic groups (Middleton, et al., 2000). Therefore, rare alleles could be missing in this data.

HLA most frequent alleles are presented in Table 9. All frequencies are provided in the appendix.

The most frequent HLA A alleles were HLA-A*02, -A*03, and -A*01, the most frequent HLA B alleles were HLA-B*44, -B*07, and -B*35. The most frequent alleles of HLA C and HLA DRB1 loci were HLA-C*07 and -C*04, and HLA-DRB1*11, -DRB1*13, and -DR1*15, respectively. Frequencies of HLA C are based on a subsample of 503 donors.

Table 9: Frequent HLA A, HLA B, HLA C, and HLA DRB1 alleles

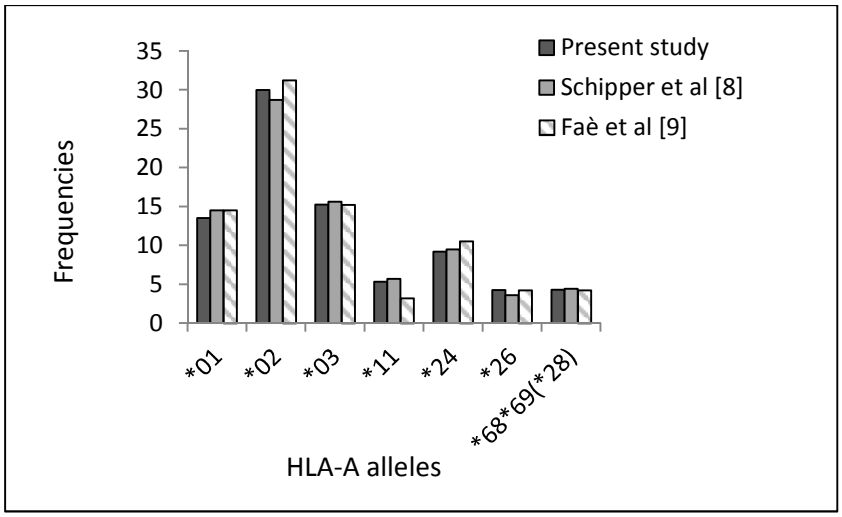
HLA A		**HLA B**	
allele	*frequency [%]*	*allele*	*frequency [%]*
***02**	29.96	***44**	12.96
***03**	15.26	***07**	11.61
***01**	13.50	***35**	10.00
***24**	9.20	***51**	8.18
***11**	5.33	***15**	7.96
***26**	4.27	***08**	7.92
		***18**	6.57
		***40**	6.06
		***27**	4.42
HLA C		**HLA DRB1**	
allele	*frequency [%]*	*allele*	*frequency [%]*
***07**	29.03	***11**	14.31
***04**	12.52	***13**	13.50
***03**	10.24	***15**	13.50
***06**	9.94	***07**	13.10
***12**	8.25	***04**	11.13
		***01**	9.85
		***03**	8.87
		***08**	4.71

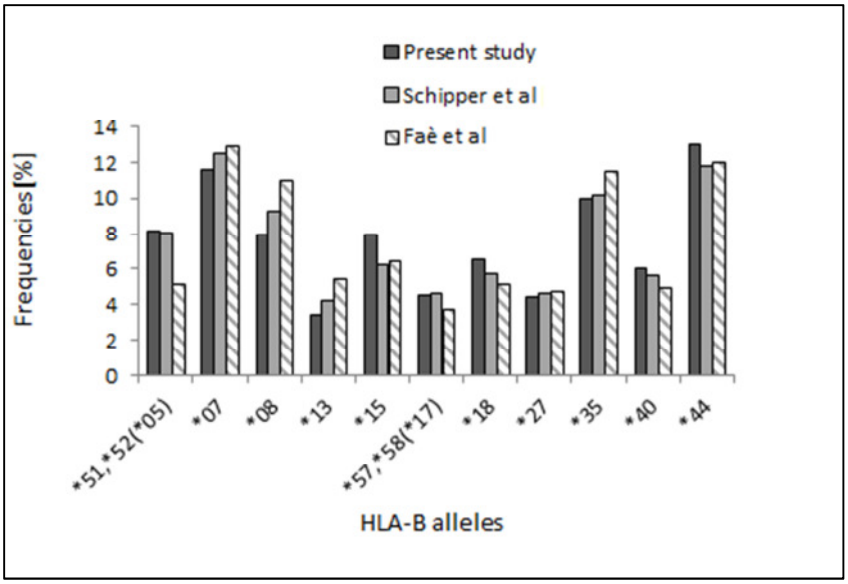
The results were compared to reported allele frequency distributions of the Austrian population. Schipper and co-workers published an overview of allele frequencies based on maximum likelihood estimates of various bone marrow donor registries (Schipper, et al., 1997). Among others they presented frequencies of HLA-A and HLA-B of the Austrian bone marrow donor registry based on 14 842 donors. A Chi² test revealed no significant differences between the HLA-A allele frequency distributions of the two studies.

In addition, the results were compared to a study by Faè S, Kriks D, Fischer M, and Fischer G published at http://www.allelefrequencies.net (Gonzales-Galarza, et al., 2011), who had a sample size of 200 donors and performed sequence based typing as well. Again, no significant differences in the allele frequency distribution of HLA-A alleles between the two studies was detected.

The analysis of HLA-B allele frequency distributions revealed a different situation. Although it is not instantly visible in Figure 10 there was a significant difference between this data and data published in 1997 by Schipper, which is especially noticeable HLA-B*05 and HLA-B*44. This can be explained by the different typing methods: The data presented in this thesis was obtained by SBT, while Shipper did serological HLA-typing. Since HLA B is the locus with far the most different alleles identified so far (Marsh, et al., 2010), a pool of different HLA split antigens in this locus – like it was done in the case of HLA-B*05 and HLA-B*44 – might not lead to accurate results when comparing those two data sets. There was no significant difference between this results and data provided by Faè et al. according to Chi^2 tests.

To my knowledge, there is no data available for HLA-C allele frequency distributions in the Austrian population. However, Faè et al. reported DRB1 allele frequency distribution. A formal Chi² test revealed no significant differences between the two studies.





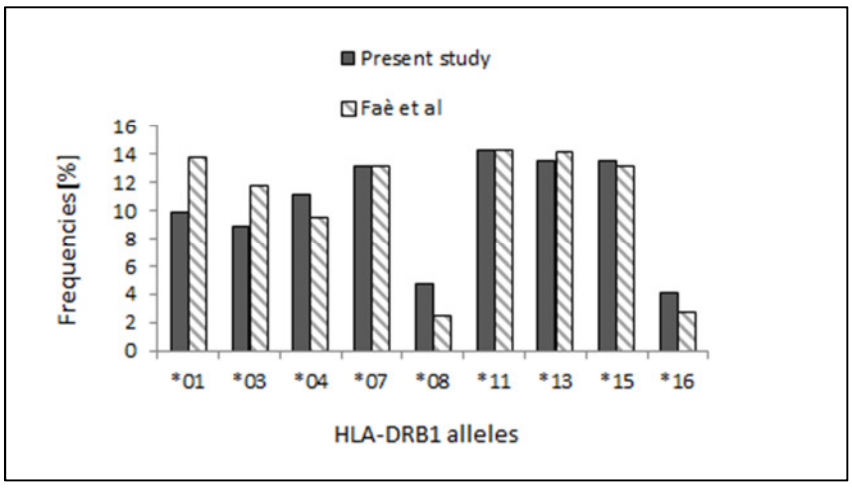


Figure 10: Comparison of the most frequent alleles of HLA-A, HLA-B, and HLA-DRB1 reported in the Austria population.
In some cases only the frequency of broad antigen serotypes were provided by **(Schipper, et al., 1997)**. In such cases the frequency of broad antigen serotypes with the respective split antigen was pooled to one class. The broader specificity was indicated in parenthesis.

3.4. Frequent HLA alleles – high resolution data

Resolving all ambiguities on the subtype level of four loci (HLA A, HLA B, HLA C, HLA DRB1, 77 donors) and three loci (HLA A, HLA B, HLA DRB1, 100 donors) yielded to 15 HLA A, 26 HLA B, 12 HLA C, and 13 HLA DRB1 alleles.

The highest level of allelic heterogeneity was observed within:

- HLA-A*02: five alleles
    - HLA-A*02:01P
    - HLA-A*02:02P
    - HLA-A*02:05P
    - HLA-A*02:06P
    - HLA-A*02:17P
- HLA-B*44: five alleles
    - HLA-B*44:02P
    - HLA-B*44:03P
    - HLA-B*44:05P
    - HLA-B*44:27P
    - HLA-B*44:29P
- HLA-C*07: five alleles
    - HLA-C*07:01P
    - HLA-C*07:02P
    - HLA-C*07:04P
    - HLA-C*07:18P
    - HLA-C*07:39P
- HLA-C*03: five alleles
    - HLA-C*03:02P
    - HLA-C*03:03P
    - HLA-C*03:04P
    - HLA-C*03:38P
    - HLA-C*03:49P
- HLA-DRB1*13: six alleles
    - HLA-C*13:01P
    - HLA-C*13:02P
    - HLA-C*13:03P

  - HLA-C*13:05P
  - HLA-C*13:42P
  - HLA-C*13:08P
- HLA-DRB*04: six alleles
  - HLA-DRB1*04:01P
  - HLA-DRB1*04:02P
  - HLA-DRB1*04:03P
  - HLA-DRB1*04:04P
  - HLA-DRB1*04:07P
  - HLA-DRB1*04:08P

A list with the most frequent alleles resolved on the subtype level is provided in the appendix.

To my knowledge there is no previously published data for high resolution HLA-data for the Austrian or German population, so a comparison with other studies was not possible.

3.5. Frequent HLA haplotypes – low resolution data

HLA-haplotypes were calculated with an EM-algorithm. 755 HLA-A-B-DRB1 haplotypes were detected. The most frequent three loci haplotypes were A*01-B*08-DRB1*03, as well as A*03-B*07-DRB1*15. The most frequent three loci haplotypes are listed in Table 10.

Table 10: Frequent three loci haplotypes (n=1370)

HLA A	HLA B	HLA DRB1	frequency [%]
*01	*08	*03	3.9402
*03	*07	*15	2.9467
*02	*15	*04	1.8274
*02	*07	*15	1.8274
*02	*44	*04	1.7363
*02	*18	*11	1.3762
*29	*44	*07	1.0981
*02	*15	*13	1.0950
*01	*57	*07	1.0762
*23	*44	*07	1.0212

510 HLA-A-B-C-DRB1 haplotypes were found. The most frequent four loci haplotypes were A*03-B*07-C*07- DRB1*15 followed by A*01-B*08-C*07-DRB1*03 as well as A*01-B*35-C*04-DRB1*01. The most frequent four loci haplotypes are listed in Table 11.

Table 11: Frequent four loci haplotypes (n=503)

HLA-A	HLA-B	HLA-C	HLA-DRB1	Frequency [%]
*03	*07	*07	*15	4.013
*01	*08	*07	*03	3.7998
*02	*44	*05	*04	1.5784
*02	*15	*03	*04	1.5381
*02	*07	*07	*15	1.5158
*02	*18	*07	*11	1.3636
*03	*35	*04	*01	1.2922
*01	*57	*06	*07	1.2155
*02	*44	*05	*13	1.1073
*23	*44	*04	*07	1.0934

Linkage disequilibrium (LD) analysis using 1000 permutation for each four pairs (476 samples) as well as for each of the three pairs (1370 samples) revealed the presence of a strong LD (p-values < 0.001) between each of the pairs of loci. The strongest relative linkage disequilibrium was revealed between HLA-B and HLA-C. This could be expected as the spatial distance between the two genes is small (Bodmer, 1987).

A study based on a huge sample of over 13 000 German blood donors listed the most frequent 320 three loci HLA-A-B-DRB1 haplotypes (Müller, et al., 2003). This list reveals similarities as one would expect due to the geographic proximity (see Figure 11: Comparison of the most frequent three loci haplotypes). Nevertheless a slightly more diverse result based on the different population-composition was anticipated, but not proven. Apparently the influence of migrants from Eastern Europe and Turkey do not contribute to varying results from the German data. Even the polymorphism in this data is comparable to the German sample. While the 20 most frequent haplotypes in this sample cover 26% of all haplotypes, Müller et al (Müller, et al., 2003) reported coverage of 27.8%. Since information of the ethnic origin of the umbilical cord blood donors or the country of birth of the respective parents were not obtained for this data or reported for the German population sample, no further comparison can be made.

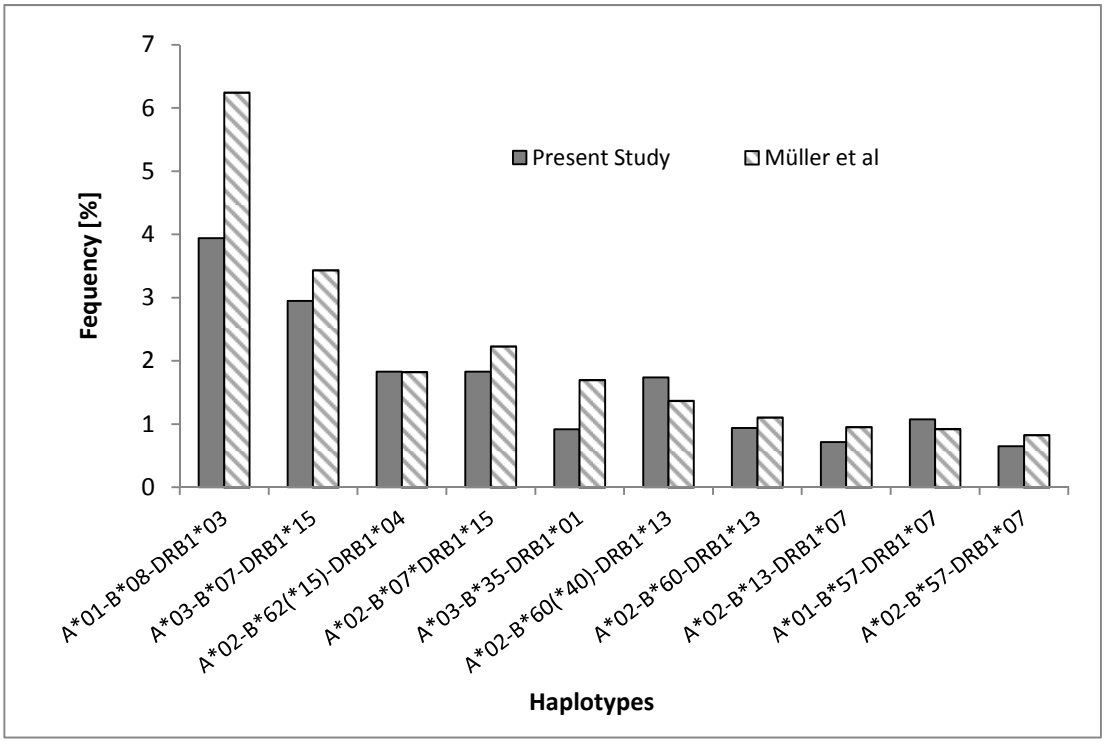


Figure 11: Comparison of the most frequent three loci haplotypes
The figure shows a comparison between the present data and a study by (**Müller, et al., 2003**) – no significant differences between the Austrian and the German data can be reported.

4. CONCLUSION

The aims of this master thesis was to present extensive HLA-data for the Austrian population and determine if the results are on the one hand comparable to previously published Austrian data (Schipper, et al., 1997) (Gonzales-Galarza, et al., 2011) and on the other hands similar to published German population data (Müller, et al., 2003).

Additionally Hardy-Weinberg exact test was done for all the analyzed loci. At locus HLA A a significant deviation was found. Since similar results were reported by a study in European Americans (Maiers, et al., 2006), due to the sample size and the fact that a deviation was only found in one locus, it was not further examined.

The comparison by Chi^2-tests of low resolution HLA-allele frequency data with previous studies showed no significant differences, except for HLA B. This can be explained due to different typing techniques and consequential data pooling.

Frequent HLA-alleles at a high resolution level were presented but due to the lack of previously published data for Austria not compared to other studies.

The most frequent haplotypes were compared to a German population. Contrary to what was expected, no significant differences were detected in the haplotype distribution and a comparison of the reported polymorphism.

To summarize, the results of this master thesis indicate that no previously assumed significant differences in the HLA-data were detected between the Austrian and the German population despite their slight disparities in the population-composition. This presented data is the first extensive HLA-data based on umbilical cord blood samples for the Austrian population and is therefore necessary for stem cell donor databases. Nevertheless prospective studies, where an Austrian population sample is compared to population data from various other countries should be done in order to further broaden the knowledge of matchable populations. This will increase the probability of finding a matching stem cell donor.

REFERENCES

Aderem, A. and Underhill, D. M. 1999. Mechanism of Phagocytosis in Macrophages. *Annu Rev Immunol.* 1999, Vol. 17, pp. 593-623.

Akira, S., Uematsu, S. and Takeuchi, O. 2006. Pathogen Recognition and Innate Immunity. *Cell.* 2006, Vol. 124, 24, pp. 783-801.

Alberts, B., et al. 2008. *Molecular Biology of the Cell.* 5th edition. New York : Garland Science, 2008. S. 1525 pp.

Banchereau, J., et al. 2000. Immunobiology of Dentritic Cells. *Annu Rev Immunol.* 2000, Vol. 18, pp. 767-811.

Beissert, S., Schwarz, A. and Schwarz, T. 2006. Regulatory T Cells. *J Invest Dermatol.* 2006, Vol. 126, pp. 15-24.

Biosystems, Applied. *3500xL Dx Genetic Analyzer User Guide.*

Blasczyk, R. 2003. HLA Diagnostic Sequencing - Conception, Application and Automation. *Journal of Laboratory and Clinical Medicine.* 27, 2003, pp. 359-368.

Bodmer, W. F. 1987. The HLA system: Structure and function. *J Clin Pathol.* 1987, Vol. 40, pp. 948-958.

Bodmer, WF. 1987. The HLA system: Structure and function. *Journal of Clinical Pathology.* 1987, pp. 945-958.

Brand, O., Gough S. and Heward, J. 2005. HLA, TLA-4 and PTPN22: the shared genetic master-key to autoimmunity? *Expert Rev Mol Med.* 2005, Vol. 7, 23, pp. 1-15.

Cairo, M. S. and Wagner, J. E. 1997. Placental and/or Umbilical Cord Blood: An Alternative Source of Hematopoietic Stem Cells for Transplantation. *Blood.* 1997, Vol. 90, 12, pp. 4665-4678.

Clayton, J and Lonjou, C. 1997. Allele and haplotype fequencies for HLA loci in various ethnic groups. 1997, pp. 665-820.

Collins, R. WM. 2004. Human MHC class I chain realted (MIC) genes: Their biological function and relevance to disease and transplantation. *Eur J Immunogen.* 31, 2004, pp. 105-114.

Creswell, P. 1994. Assembly, transport, and function of MHC class II molecules. *Annu Rev Immunol.* 1994, Vol. 12, pp. 259-293.

Crotty, S. and Ahmed, R. 2004. Immunological Memory in Humans. *Sem Immunol.* 2004, Vol. 16, pp. 197-203.

Dausset, J. 1958. Iso-leuko-anticorps. *Acta Haemat.* 1958, Vol. 20, pp. 156-166.

Excoffier, L and Lischer, HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources.* 2010, 10, pp. 564-567.

Excoffier, L and Slatkin, M. 1995. Maximum-lilelihood estimation of molecular haplotype frequencies in diploid population. *Mol Biol Evol.* 1995, 12, p. 921.

Fallin, D and Schork, NJ. 2000. Accuracy of Haplotype Frequency Estimation for Biallelic Loci, via the Expectation-Maxmization Algorithm for Unphased Diploid Genotype Data. *The American Journal of Human Genetics.* 2000, 67, pp. 947-959.

Feder, J. N., et al. 1996. A novel MHC class I like gene is mutated in patients with hereditary Hemochromatosis. *Nature Genet.* 13, 1996, Vol. 4, pp. 339-408.

Gonzales-Galarza, FF, et al. 2011. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acid Research.* 39, 2011, pp. D913-D919.

Guardiola , J., et al. 1996. Functional significance of polymorphism among MHC class II gene promotors. *Tissue Antigens.* 1996, Vol. 48, 6, pp. 615-625.

Guo, SW and Thompson, EA. 1992. Performin the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics.* 1992, Vol. 48(2).

Hamerman, J. A., Ogasawara, K. and Lanier, L. L. 2005. NK cells in innate immunity. *Curr Opin Immunol.* 2005, Vol. 17, pp. 29-35.

Heeg, K. 2008. *The Hypothalamus - The innate immune system.* s.l. : Elsevier, 2008. pp. 87-99.

Horton, R., et al. 2004. Gene map of the extended human histocompatibility complex. *Nature Rev Genet.* 5, 2004, Vol. 12, pp. 889-899.

Jin, P and Wang, E. 2003. Polymorphism in Clinical Immunology - From HLA Typing to Immunogenetic Profiling. *Journal of Translational Medicine.* 2003.

Jung, D. and Alt, F. W. 2004. Unraveling V(D)J Recombination: Insights into Gene Regulation. *Cell.* 2004, Vol. 116, 23, pp. 299-311.

Klein, J. und Sato, A. 2000. The HLA system. First of two parts. *N Engl Jour Med.* 2000, Bd. 343, 10, S. 702-709.

Körbling, M. and Anderlini, P. 2001. Peripheral Blood Stem Cell Versus Bone Marrow Allotransplantation: Does the Source of Hematopoietic Stem Cells Matter? *Blood.* 2001, Vol. 98, 10, pp. 2900-2908.

Lai, A. Y. and Kondo, M. 2008. T and B lymphocyte differentiation from hematopoietic stem cell. *Semin Immunol.* 20, 2008, pp. 207-212.

Ljunggren, H. G. and Thorpe, C. J. 1996. Principles of MHC class I mediated antigen presentation and T cell selection. *Histol Histopathol.* 1996, Vol. 11, 1, pp. 267-274.

Long, JC, Williams, RC und Urbanek, M. An EM-Algorithm and testing strategy.

Loureiro , J. and Ploegh, H. L. 2006. Antigen presentation and the ubiquitin-proteasome system in host-pathogen interactions. *Advanc Immunol.* 2006, Vol. 92, pp. 225-305.

Lu, L. F. and Rudensky, A. 2009. Molecular orchestration of differentiation and function of regulatory T cells. *Genes Dev.* 2009, Vol. 23, pp. 1270-1282.

Maiers, Martin, Gragert, Loren and Klitz, William. 2006. High-resolution HLA alleles and haplotypes in the United States population. *Human Immunol.* 2006, Vol. 68, pp. 779-788.

Marsh, S.G.E., et al. 2010. Nomenclature for factors of the HLA system, 2010. *Tissue Antigens.* 2010, Vol. 75, pp. 291-455.

Marsh, Steven G.E., Parham, Peter and Barber, Linda D. 2000. *The HLA FactsBook.* London, California : Academic Press, 2000.

Medzhitov, R. and Janeway, C. A. 1997. Innate immunity: impact on the adaptive immune response. *Curr Opin Immunol.* 1997, Vol. 9, pp. 4-9.

Mehra, N. K., et al. 2010. *The HLA complex in Biology and Medicine: A resource book.* New Delhi : Jaypee Brotheres Medical Publishers, 2010.

Middleton, D., et al. 2000. Analysis of the Distribution of HLA-A Alleles in Populations from Five Continents. *Human Immunology.* 2000, Vol. 61, pp. 1048-1052.

Müller, Carlheinz R., Ehninger, Gerhard and Goldmann, Shraga F. 2003. Gene and Haplotype Frequencies for the Loci HLA-A, HLA-Bm and HLA-DR based on Over 13 000 German Blood Donors. *Human Immunology.* 2003, Vol. 64, pp. 137-151.

Natarajan, K., et al. 1999. MHC class I molecules, structure and function. *Rev Immunogen.* 1999, Vol. 1, 1, pp. 32-46.

Pangburn, M. K., Ferreira, V. P. and Cortes, C. 2008. Discrimination between host and pathogens by the complement system. *Vaccine.* 2008, Vol. 30, 8, pp. 115-121.

Petersdorf, EW, et al. 2001. Major-Histocompatibility-Complex. Class I Alleles and Antigens in Hematopoietic-Cell Transplantation. *The New England Journal of Medicine.* 2001, pp. 1794-1800.

Robinson, J, et al. 2011. The IMGT/HLA database. *Nucleic Acids Research.* 2011, Vol. 39, pp. D1171-D1176.

Roitt, I. M., Brostoff, J. and Male, D. K. 1995. *Kurzes Lehrbuch der Immunologie.* 3. neubearbeitete Auflage. New York : Thieme, 1995. p. 2 pp.

Schipper, R.F., et al. 1997. HLA Gene and Haplotype Frequencies in Bone Marrow Worldwide Registries. *Human Immunology.* 1997, Vol. 52, pp. 54-71.

Schmidt, AH, et al. 2009. Estimation of high-resolution HLA-A, -B, -C, -DRB1 alleles and haplotype frequencies based on 8862 German stem cell donors and implications for strategic donor registry planning. *Human Immunology.* 2009, Vol. 70, pp. 895-902.

Smith, D. M., et al. 2005. HLA class I null alleles and new alleles affect unrelated bone marrow donor searches. *Tissue Antigens.* 2005, Vol. 66, 2, pp. 93-98.

Spits, H. 2002. Development of T cells in the human thymus. *Nat Rev Immunol.* 2002, Bd. 2, S. 636-645.

Statistik Austria. 2011. *Bevölkerungsstand 2011.* Wien : Verlag Österreich GmbH, 2011.

Stone, K. D., Prussin, C. and Metcalfe, D. D. 2010. IgE, Mast Cells, Basophils, and Eosinophils. *J Allergy Clin Immunol.* 2010, Vol. 125, pp. 73-80.

The HLA FactsBook. **Marsh, Steven G.E., Parham, Peter and Barber, Linda D. 2000.** London, California : Academic Press, 2000.

Thorsby, E. 2009. A short history of HLA. *Tissue Antigens.* 2009, Vol. 74, pp. 101-116.

Zhao, J., et al. 2008. Do adaptive immune cells suppress or activate innate immunity? *Trends Immunol.* 30, 2008, Vol. 1, pp. 8-11.

APPENDIX

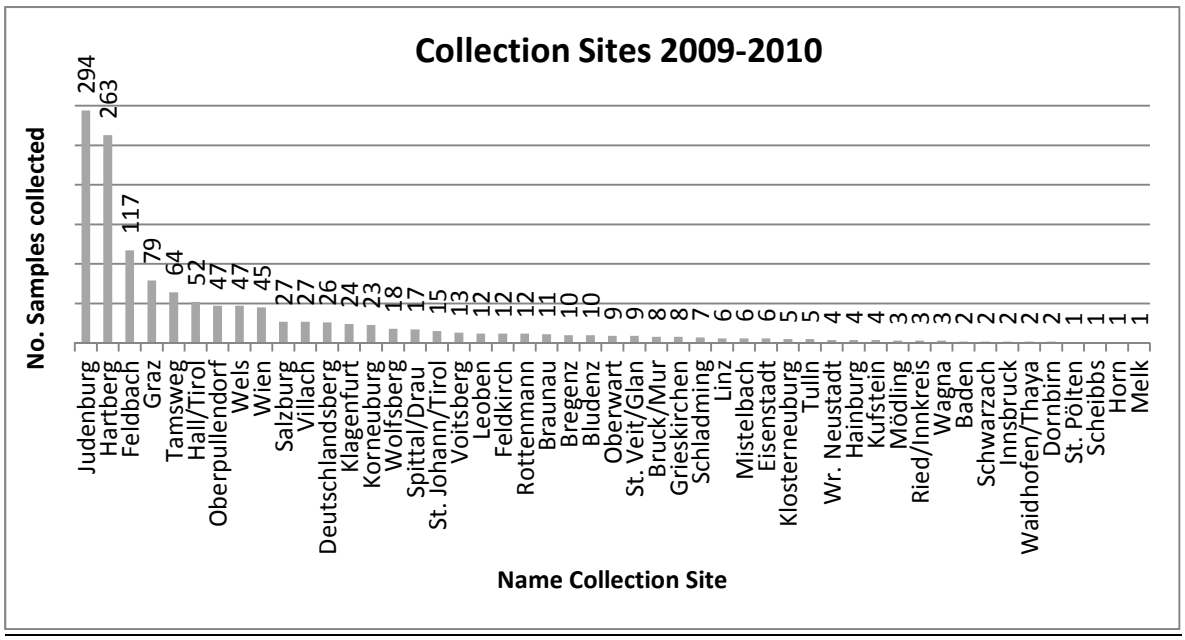


Figure 12: Location of collection sites

Table 12: Complete list of HLA alleles (low resolution)

HLA-A		HLA-B	
allele	*frequency [%]*	*allele*	*frequency [%]*
***02**	29.96	***44**	12.96
***03**	15.26	***07**	11.61
***01**	13.50	***35**	10.00
***24**	9.20	***51**	8.18
***11**	5.33	***15**	7.96
***26**	4.27	***08**	7.92
***68**	4.23	***18**	6.57
***32**	3.36	***40**	6.06
***31**	2.99	***27**	4.42
***29**	2.81	***57**	3.58
***23**	2.74	***13**	3.47
***25**	2.26	***38**	2.52
***30**	2.12	***14**	2.45
***33**	1.20	***49**	1.82
***66**	0.37	***39**	1.79
***34**	0.33	***50**	1.46
***69**	0.07	***55**	1.46
		***37**	1.13
		***56**	0.99
		***58**	0.99
		***52**	0.84
		***41**	0.77
		***47**	0.37
		***45**	0.33
		***53**	0.18
		***46**	0.07

		***48**	0.07
		***42**	0.04
		***73**	0.04
HLA-C		**HLA-DRB1**	
allele	*frequency [%]*	*allele*	*frequency [%]*
***07**	29.03	***11**	14.31
***04**	12.52	***13**	13.50
***03**	10.24	***15**	13.50
***06**	9.94	***07**	13.10
***12**	8.25	***04**	11.13
***02**	7.26	***01**	9.85
***05**	6.96	***03**	8.87
***01**	4.47	***08**	4.71
***015**	3.68	***16**	4,12
***16**	3.58	***14**	3,28
***08**	1.89	***12**	1,97
***14**	1.79	***10**	0,84
***17**	0.40	***09**	0,80

Table 13: Frequent HLA alleles (high resolution)

HLA-A			**HLA-B**			**HLA-C**			**HLA-DRB1**		
allele	*subtype*	*frequency [%]*	*allele*	*subtype*	*frequency [%]*	*allele*	*subtype*	*frequency [%]*	*allele*	*subtype*	*frequency [%]*
***02**	02:01P	28.50	***44**	44:02P	9.00	***07**	07:01P	10.39	***11**	11:01P	1.50
	02:02P	0.50		44.03P	4.50		07.02P	9.09		11:02P	1.00
	02:05P	0.50		44:05P	1.00		07:04P	1.95		11:03P	0.50
	02:06P	0.50		44:27P	1.00		07:18P	0.65		11:04P	6.50
	02:17P	0.50		44:29P	0.50		07:39P	0.65		11:30P	0.50
***03**	03:01P	12.50	***07**	07:02P	7.50	***04**	04:01P	11.04	***13**	13:01P	9.50
***01**	01:01P	14.50		07:10P	0.50	***03**	03:02P	0.65		13:02P	0.50
***24**	24:02P	7.00	***35**	35:01P	4.50		03:03P	3.25		13:03P	2.50
	24:03P	0.50		35:02P	1.00		03:04P	9.09		13:05P	0.50
***11**	11:01P	6.00		35:03P	2.50		03:38P	0.65		13:42P	0.50
***26**	26:01P	5.00		35:08P	0.50		03:49P	0.65		13:67P	0.50
***68**	68:01P	5.00	***51**	51:01P	7.50				***15**	15:01P	11.00
	68:12P	1.00		51:05P	0.50				***07**	07:01P	19.00
			***15**	15:01P	7.50				***04**	04:01P	10.50
				15:08P	0.50					04:02P	0.50
				15:17P	0.50					04:03P	1.00
				15:18P	0.50					04:04P	4.00
			***08**	08:01P	2.50					04:07P	1.00
			***18**	18:01P	7.00					04:08P	0.50
				18:05P	0.50				***01**	01:02P	4.00
			***40**	40:01P	4.00				***08**	08:01P	5.00
				40:02P	1.00					08:02P	1.00
			***27**	27:02P	0.50					08:04P	1.00
				27:05P	3.50					08:10P	0.50
									***16**	16:01P	7.00
										16:02P	1.00

ACKNOWLEDGEMENTS

I sincerely want to thank:

Prof. Dr. Peter Macheroux….....	for not only agreeing to be my supervisor, but also for his patience and willingness to accept this unusual topic
Prof. Dr. med. Peter Wernet…….	for agreeing to be my supervisor so late on in the project and supporting me throughout the end
Vivocell Bisolutions ……	for agreeing to use their data and giving me the opportunity to write this master thesis
Dr. Gudrun Schappacher-Tilp ….	for explaining every statistical problem over and over again by endless cups of coffee until even I understood
Dr. Verena Börger ……	for pushing me these last 100 meters, for all these motivation emails and helpful corrections and suggestions – I honestly think you saved me from going crazy!
Petra, Nina, Martina, Resi and Franz……..	for not only being the best colleagues one could wish for, but also for making every work day fun and challenging
Daisy ……	for always exactly saying what I need to hear and being the wonderful kind person you are
Simone …….	for always understanding me when no one else does and saying out loud what I only dare to think
Mirjam …….	for being a great friend and best study-buddy (☺) one can wish for
Johnny ……..	for all those lunches, dinners, road-trips, serious conversations, and silly chit-chats
Ma & Pa ……...	for simply being who you are and supporting me no matter what

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

19.12.2012	Stefanie Nerstheimer
date	(signature)