UNIVERSITÉ DE GENÈVE
FACULTÉ DE MÉDECINE

Clinical Medicine Section
Department of Medicine
Service of Nephrology

"Optimizing clinical outcome in allogeneic hematopoietic stem cell transplantation: pre-transplant and post-transplant immunogenetic markers"

Thesis submitted to the Faculty of Medicine of
the University of Geneva

for the degree of Privat-Docent

by

Stéphane BUHLER

Geneva

2023

Table of contents

Acknowledgments

I am very grateful to my mentors who introduced me to the exciting field of HLA, histocompatibility and immunogenetics, the Professors Alicia Sanchez-Mazas and Jean-Marie Tiercy. I would also like to thank all their collaborators with whom I worked and learned from throughout the years, even until today for some of them.

More recently and after several years devoted to fundamental research in evolutionary biology and population genetics, Dre Sylvie Ferrari-Lacraz and Professor Jean Villard welcomed me to the Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility (UIT-LNRH) in Geneva University Hospitals. I am very thankful for their continuous support and their teaching of the management of a routine clinical laboratory in transplantation immunology. I am also thankful for their academic support and expertise, notably concerning the publications that are presented in this thesis and our ongoing research projects.

I would also like to thank the technicians and administrative staff at UIT-LNRH and the colleagues from our research team for the daily interactions and fruitful collaboration.

Science is about sharing knowledge and networking. The HLA community is a nice example of these characteristics, as the discovery and further characterization of such a complex genetic system was made possible thanks to the early pioneers and the establishment of international workshops taking place regularly since 1964 until today. I had the privilege to participate to several of them and also to other international conferences (such as the annual European Federation of Immunogenetics meetings) to collaborate with and learn from many colleagues and experts abroad.

Last but not least, I would like to thank my lovely family for their presence and support during all these years.

Summary

Allogeneic hematopoietic stem cell transplantation (alloHSCT) is a standard treatment for hematological disorders like acute leukemia and primary immune deficiencies. Various parameters at the pre, peri and post-transplant stages have been shown to significantly influence clinical outcome and patient's prognosis. In this thesis submitted to the University of Geneva for the degree of Privat-Docent, several of the latest advances in transplantation immunology, histocompatibility and immunogenetics are examined by integrating five of my recent studies to a review of the scientific literature on this thematic. Two main aspects are discussed more specifically.

At the pre-transplant stage, the importance of optimal genetic compatibility between donor and patient and the development of integrative donor selection algorithms that allow to better predict post-transplant risks are discussed through the lens of the HLA-DPB1 paradigm. Different contributions and models to adjust and fine-tune the degree of HLA matching, such as the concept of permissive mismatches, are presented, including a perspective on the availability of new therapeutic protocols and alloHSCT platforms. Indeed, the increasing choice among different types of donors means that almost every patient can now benefit from a suitable donor. However, several questions remain open and are revisited regularly with the publication of new data and results. For instance, what is the best option to consider when several potential donors are available within and outside the patient's family? If the current gold standards for matching are not achievable, which mismatch constellation(s) should be favored? Because the therapy is highly personalized, different approaches are not mutually exclusive and can be tailored to the patient's needs.

At the post-transplant stage, immune recovery of the patient is central to the success of this cellular therapy. Notably, the reconstitution of a broad and tolerant T-cell repertoire is required to alleviate the burden of post-transplant complications such as graft-versus-host disease and infections. The development of high throughput sequencing approaches has enabled the in-depth characterization of the repertoire and has uncovered the complexity of the adaptive immune response to foreign (and self) antigens. In this context, the interest of alloHSCT as a unique model in humans to explore the dynamics and architecture of the T-cell receptor (TCR) repertoire is presented. In addition, the hurdles at characterizing and predicting the alloimmune response and the promises of translating these results to an optimized and personalized post-transplant monitoring of patients are discussed.

Introduction

1. An overview of allogeneic hematopoietic stem cell transplantation

Since the first clinical trials in the late sixties-early seventies, allogeneic hematopoietic stem cell transplantation (alloHSCT) has been established as a curative and standard of care treatment for a large variety of hematologic disorders (Appelbaum 2007, Jenq and van den Brink 2010). This includes life-threatening malignant diseases, primary immune deficiencies and some other hereditary conditions and defective immune states (Snowden et al. 2022). Among these, acute myeloid and lymphoblastic leukemia (AML and ALL, respectively) represent the two most common indications and account together for more than 50% of allografts in Europe (Snowden et al. 2022). The number of allografts performed worldwide keeps growing on a yearly basis, thanks to new indications and a broaden access to this therapy for older patients (Passweg et al. 2021). In autologous HSCT, the stem cells are directly obtained from the patient, processed, stored and subsequently reinfused after the patient completes conditioning (Blume and Thomas 2000). By contrast, the principle of alloHSCT is to replace the hematopoietic/immune system of a patient by the one of a healthy donor carefully selected according to specific and stringent criteria, foremost his/her genetic compatibility with the recipient (described in chapter 2). Other important steps in the procedure comprise conditioning of the recipient, stem cell collection in the donor and reinfusion of cells in the recipient (briefly described hereafter), engraftment and immune reconstitution, including post-transplant interventions and monitoring of post-transplant complications (described in chapter 3).

The key distinctive characteristics of alloHSCT in comparison to autologous transplantation are that the graft is not potentially contaminated by residual malignant cells and contains immunocompetent alloreactive lymphocytes from the donor that are capable of eliciting a curative response (Giralt and Bishop 2009). This response will be mediated by the recognition of foreign antigens on tumor cells leading to their eradication and a decreased risk of relapse, the so-called graft-versus-leukemia (GVL) effect. This latter characteristic is a major advantage provided by alloHSCT as a cellular therapy. However it is accompanied by the risk that the immune response will also be directed against normal tissues of the recipient, causing the severe condition known as graft-versus-host disease (GVHD).

a. Conditioning

The main purpose of conditioning is to allow sufficient graft space for the incoming donor stem cells and is also crucial in order to prevent graft rejection by the host immune system (Giralt and Bishop 2009, Juric et al. 2016). Another contribution is to eradicate tumor cells in case of an underlying oncologic disease. Several conditioning regimens are available mainly based on total body irradiation (TBI), chemotherapy (e.g., alkylating agents such as busulfan or cyclophosphamide) or a combination of both (Juric et al. 2016). The choice depends on the disease to be treated and on patient's characteristics. The regimens can be subdivided into several broad categories regarding intensity of the treatment such as myeloablative (MA), non-myeloablative (NMA) and reduced intensity conditioning (RIC). MA conditioning provides a better control of tumor cells but this is negatively counterbalanced by a significant toxicity, morbidity and transplant-related mortality (TRM). By contrast, NMA and RIC have allowed access to this therapy to older patients because of decreased TRM risks compared to MA. However, these regimens are accompanied by an increased risk of relapse. Notably, NMA provides only a limited control of tumor cells and relies principally on the immune cells infused with the graft for the subsequent eradication of cancer. RIC is somewhere in between MA and NMA, with the goal to limit the toxicity and mortality associated with the therapy but still providing sufficient immunoablation and control of the underlying malignancy (Juric et al. 2016).

The conditioning is a major cause of epithelial damage, especially the mucosa of the skin, the liver and the gut, and of inflammation notably by stimulating the release of pro-inflammatory cytokines such as tumor necrosis factor (TNF)-α and interleukins (IL)-1 and -6. Various endogenous and exogenous danger signals released by damaged tissues (i.e., damage and pathogen associated molecular patterns, DAMPs like alarmins or PAMPs derived from the microbiota, respectively) will contribute in amplifying the innate immune response which will in turn activate professional antigen presenting cells (APCs) such as dendritic cells of the recipient (Ramadan and Paczesny 2015). This will lead to an increased presentation of major and/or minor histocompatibility antigens (see chapter 2) and the risk of activating mature alloreactive donor T cells infused with the graft. This will stimulate their expansion and attack on target tissues, causing GVHD.

b. Source of stem cells

Donor hematopoietic stem cells (SC) can be harvested from the bone marrow (BM), the peripheral blood or umbilical cord blood units (CBUs). Peripheral blood stem cells (PBSCs)

are the preferred source since many years and their usage shows a steady increase annually (Passweg et al. 2021). To harvest PBSCs, the administration of growth factors such as granulocyte colony-stimulating factor (G-CSF) to the donor is necessary in order to mobilize a sufficient amount of $CD34^+$ cells (i.e., a main marker for SC) at the periphery before the collection. The use of PBSCs has been associated with a quicker recovery of hematopoiesis, a decreased rate of relapse for hematological malignancies, principally due to the presence of a higher proportion of mature T cells providing a GVL effect, and to a higher incidence of chronic GVHD compared to BM (Juric et al. 2016). By contrast, the use of CBUs has continued to diminish through time, notably since new advances in therapeutic strategy for haploidentical transplantation (Ciurea and Bayraktar 2015, McCurdy and Fuchs 2016, Baumeister et al. 2020) have allowed a remarkable increase in the selection of such related donors (Passweg et al. 2017), see hereafter in chapter 2.

For further technical details on conditioning and stem cell sources, the reader is referred to two reviews (Giralt and Bishop 2009, Juric et al. 2016).

2. Genetic compatibility in alloHSCT

In alloHSCT, the genetic compatibility between the patient and his/her donor is of paramount importance for global outcome. It is crucial for avoiding graft rejection by the recipient immune system (in conjunction with conditioning) and for minimizing the risks of relapse and development of severe GVHD (Petersdorf 2013, Fleischhauer and Beelen 2016). Indeed, alloreactive donor T cells infused with the graft will target genetic mismatches expressed on the recipient tissues, mediating GVHD on the one hand, but they will also potentially recognize these mismatches and/or neoplastic antigens on tumor cells, mediating GVL on the other hand. Promoting a strong T-cell mediated GVL effect while specifically controlling the deleterious consequences of GVHD has often been considered the Holy Grail of alloHSCT as a cellular therapy (Fleischhauer and Shaw 2017) and many breakthrough progresses have been made (Chabannon et al. 2018).

a. The Major Histocompatibilty Complex (MHC)

A main objective when selecting a potential donor for alloHSCT is to achieve compatibility at several key classical HLA genes (the acronym for Human Leukocyte Antigen) located within the extended MHC. This represents a genomic region of 7.6 MB located on the short arm of

chromosome 6 (i.e., the 6p21.3 region) encompassing 421 loci with 252 of them expressed as proteins and close to 25% being associated with immune functions (1999, Horton et al. 2004). This genomic region is subdivided into three parts. The class II sub region located closer to the chromosome centromere and which contains HLA classical (DR, DQ and DP) and non-classical (DM and DO) genes as well as genes coding for proteins involved in antigen processing and presentation. The class III sub region which contains many genes, notably encoding molecules involved in inflammation and the complement cascade, but no HLA genes. The class I sub region located closer to the telomere and which contains HLA classical (A, B and C) and non-classical (E, F and G) genes as well as genes involved in the innate immune response or having other functions.

The HLA classical genes: polymorphism, structure and function

HLA genes segregate together on chromosome 6 as haplotypes, one inherited from the mother and one from the father. These are the most variable genes in the human genome. Currently, more than 25,000 HLA class I alleles have been described, while class II genes are slightly less variable with more than 10,000 alleles characterized so far (https://www.ebi.ac.uk/ipd/imgt/hla/about/statistics/, (Robinson et al. 2020)). Besides the extremely high levels of polymorphism observed in human populations (Buhler and Sanchez-Mazas 2011, Sanchez-Mazas and Nunes 2018), some pairs of genes such as B~C and DRB1~DQB1 are in significant linkage disequilibrium (Bugawan et al. 2000), with some combinations of alleles segregating together more frequently on haplotypes than what would be expected randomly from their respective allele frequencies. By contrast, some other loci like HLA-A and HLA-DPB1 are defined by a weaker association or by linkage equilibrium with other HLA genes, respectively (Cullen et al. 1997, Kauppi et al. 2005).

The expression of HLA genes is codominant, meaning that individuals who are heterozygotes at a given locus will carry molecules at the cell surface encoded by both alleles. Regarding HLA class I molecules, they are expressed on most tissues and nucleated cells and are composed of one alpha chain encoded by the highly polymorphic HLA-A, B and C genes that is non-covalently bound to β2 microglobulin, an invariant protein encoded by the B2M gene located on chromosome 15 (Figure 1). HLA class II molecules, constitutively expressed on a limited number of specialized cells (Roche and Furuta 2015), are heterodimers composed of one alpha subunit encoded by the HLA-DRA, DQA1 and DPA1 genes and one beta subunit encoded by the HLA-DRB1/3/4/5, DQB1 and DPB1 genes (Figure 1), respectively. However, the cell surface expression of HLA class II molecules is inducible in many cell types, notably during

inflammatory processes (Ting and Trowsdale 2002). The structure of HLA class I and II molecules shares several features (Wu et al. 2021). The most important one is the presence of an antigen recognition domain, where a large proportion of the polymorphism is located, and which is composed of two antiparallel alpha helixes and a β-sheet constituting a groove that accommodates peptide antigens and that is therefore called the peptide binding region (PBR, Figure 1).

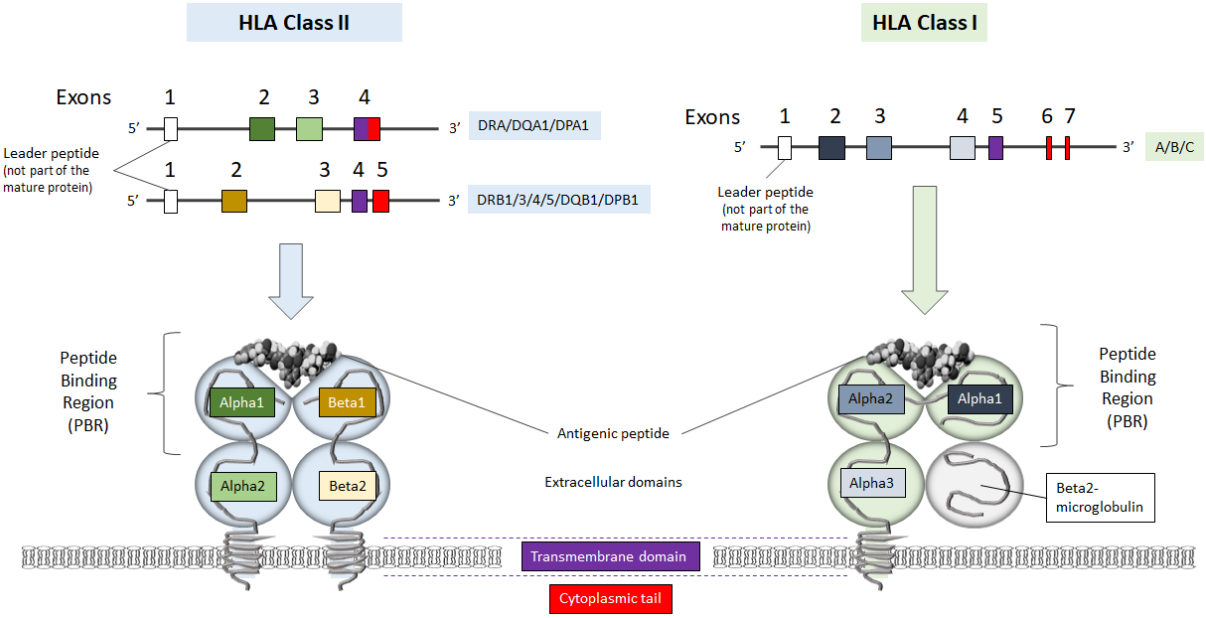

Figure 1: Schematic representation of the structure of HLA class I and class II genes and the corresponding encoded cell surface molecules (personal contribution/material).

The MHC genomic region and yet unknown H-2 and HLA genes in the mouse and human, respectively, were first discovered for their role in histocompatibility during experiments on tumors and tissue engraftment and by using sera from multitransfused patients and multiparous women (Thorsby 2009, Terasaki 2013). It took many more years to characterize the molecular structure of HLA proteins by X-ray crystallography and to understand their main biological function which is to present small antigenic peptides to T cells (Bjorkman 2006). The general rule is that HLA class I molecules present peptides of endogenous cellular origin to $CD8^+$ cytotoxic T cells, while HLA class II molecules bind peptides derived from the extracellular environment for presentation to $CD4^+$ helper T cells (Gfeller and Bassani-Sternberg 2018, Pishesha et al. 2022). In addition, dendritic cells have a unique ability to present internalized antigens on HLA class I molecules through cross-presentation pathways (Embgenbroich and Burgdorf 2018). It has also been shown that up to 20-30% of peptides presented by HLA class II molecules are derived from cytosolic and nuclear proteins through a pathway of autophagy (Roche and Furuta 2015).

In order for each individual to generate a diverse and functional T-cell repertoire, recognition and tolerance of self HLA molecules presenting peptides derived from endogenously produced proteins (i.e., self peptides) are important processes taking place during T-cell maturation and education in the thymus (Klein et al. 2014, Cosway et al. 2021). At the periphery, recognition of foreign (i.e., non-self) peptides presented by self HLA during an infection or of allogeneic peptides:HLA (pHLA) complexes in the context of a transplantation will lead to the specific activation of effector T cells, see also in chapter 3.

Codominance, polygeny and polymorphism are three important characteristics of HLA genes. These endow each individual in the population with the capacity to present a large and diverse repertoire of peptides at the cell surface and thus the ability to mount effective adaptive immune responses against most pathogens. It is well documented that HLA genes are the target of natural selection, especially balancing selection acting on the maintenance of high levels of genetic diversity (Meyer and Thomson 2001, Buhler and Sanchez-Mazas 2011, Sanchez-Mazas and Meyer 2014, Marostica et al. 2022).

In addition to the presentation of antigens to T cells, some HLA-A and B and all HLA-C molecules are ligands for Killer-cell immunoglobulin-like receptors (KIR) expressed at the surface of Natural Killer (NK) cells. Thus, HLA class I molecules fulfill important biological functions at the crossroad between innate and adaptive cellular immune responses (Horowitz et al. 2016, Pende et al. 2019).

The characterization of HLA polymorphism first by serology and then using molecular typing techniques has benefited from many technological developments (Baxter-Lowe 2021), including the recent advent of next generation and third generation sequencing, i.e., NGS and TGS, respectively (Bravo-Egana et al. 2021, Liu 2021). The choice of a genotyping methodology will determine the level of resolution that can be attained, i.e., defined as low, intermediate, high or allelic (Nunes et al. 2011). Because the HLA system is highly complex, a nomenclature was established very early after its discovery to standardize the naming of alleles and thus their usefulness for research and clinical applications (Hurley 2021).

Selection of donors and HLA matching

HLA incompatibilities represent the major genetic barrier in alloHSCT. In practice, donors can be chosen among several alternatives that will determine the degree of compatibility and also the probability to find such a donor (Tiercy 2016, Petersdorf 2017). In every family, Mendelian inheritance of HLA haplotypes means that siblings have a 25% chance to be genotypically identical to each other. HLA identical siblings are still considered the conventional optimal

choice for transplantation purposes (Tiercy 2016, Spellman 2022). In Europe or in the USA, however, only about 30% of patients can benefit from such a donor. For the other patients, alternative donors can be chosen among one or several of the following: a phenotypically compatible unrelated adult volunteer donor (i.e., a matched unrelated donor or MUD), a mismatched unrelated donor (MMUD), a haploidentical related donor (i.e., sharing one inherited HLA haplotype with the patient) or a CBU. Recent overall survival data from the Center for International Blood and Marrow Transplant Research (CIBMTR) suggest a hierarchy for donor selection that prioritize HLA identical siblings, MUDs and then the different mismatched donor sources (Spellman 2022). Similar trends are observed in Switzerland as reported annually by the Swiss Blood Stem Cell Transplantation (SBST) group to the Federal Office of Public Health, see also (Passweg et al. 2018). However, some practical considerations such as rapid access to a suitable donor, stage of the underlying disease or other relevant clinical parameters imply that the search strategy is often personalized for each patient.

Usually, when a matched relative cannot be proposed, the first choice is to consider a MUD among the close to 40 million volunteer donors that are currently available in donor registries around the world (https://statistics.wmda.info/). Indeed, despite the high levels of polymorphism of HLA genes, it is very often possible to identify suitable donors outside of the patient's family, notably because some alleles are common in populations and because of the significant linkage between loci (Sanchez-Mazas et al. 2017, Hurley et al. 2020). The current gold standard for an unrelated donor is to consider a 8/8 or 10/10 match defined by high resolution typing at HLA-A, B, C, DRB1 ± DQB1, the latter locus usually not being considered in the USA in contrast to Europe (i.e., in fact matching for DRB1 usually bring along a DQB1 match because of the strong linkage disequilibrium). When several equally matched donors are available, non-HLA characteristics such as age, sex, CMV matching and ABO matching are considered in the selection process (Booth et al. 2013, Kollman et al. 2016). In Switzerland, the National Reference Laboratory for Histocompatibility (LNRH) based in Geneva is responsible for providing a prediction to find a 10/10 MUD and a search strategy for every patient in need of alloHSCT with an unrelated donor (Figure 2). As shown on the plot, more than 50% of patients can benefit from at least one 10/10 MUD and for many patients the choice is even possible among several equally well-matched donors.

For the remaining patients, novel strategies and protocols have improved the safety of transplant procedures across the HLA barrier (Luznik et al. 2008, Al Malki et al. 2021, Shaw et al. 2021). A 9/10 MMUD is thus often a possible alternative and is generally associated with good clinical

outcome (Petersdorf 2016). It is also an option to consider a cord among the more than 800'000 CBUs registered across the world. One advantage of CBUs is that more HLA mismatches are allowed due to the immaturity of the immune cells contained in the graft, but the limited quantity of cryopreserved stem cells precludes this option for many adult patients (Gluckman and Rocha 2009). As already mentioned, haploidentical donors have now supplanted CBUs in the number of transplants, at least in Europe (Passweg et al. 2017). For any given patient, both parents, one or several siblings (i.e., the chance of siblings sharing one haplotype is of 50% according to Mendelian inheritance), a child or sometimes selected members of the extended family can represent a suitable haploidentical donor depending on the actual consensus recommendations (Ciurea et al. 2020). Yet, the optimal choice between a 9/10 MMUD, a haploidentical donor or a CBU still remains a controversial issue (Spellman 2022).

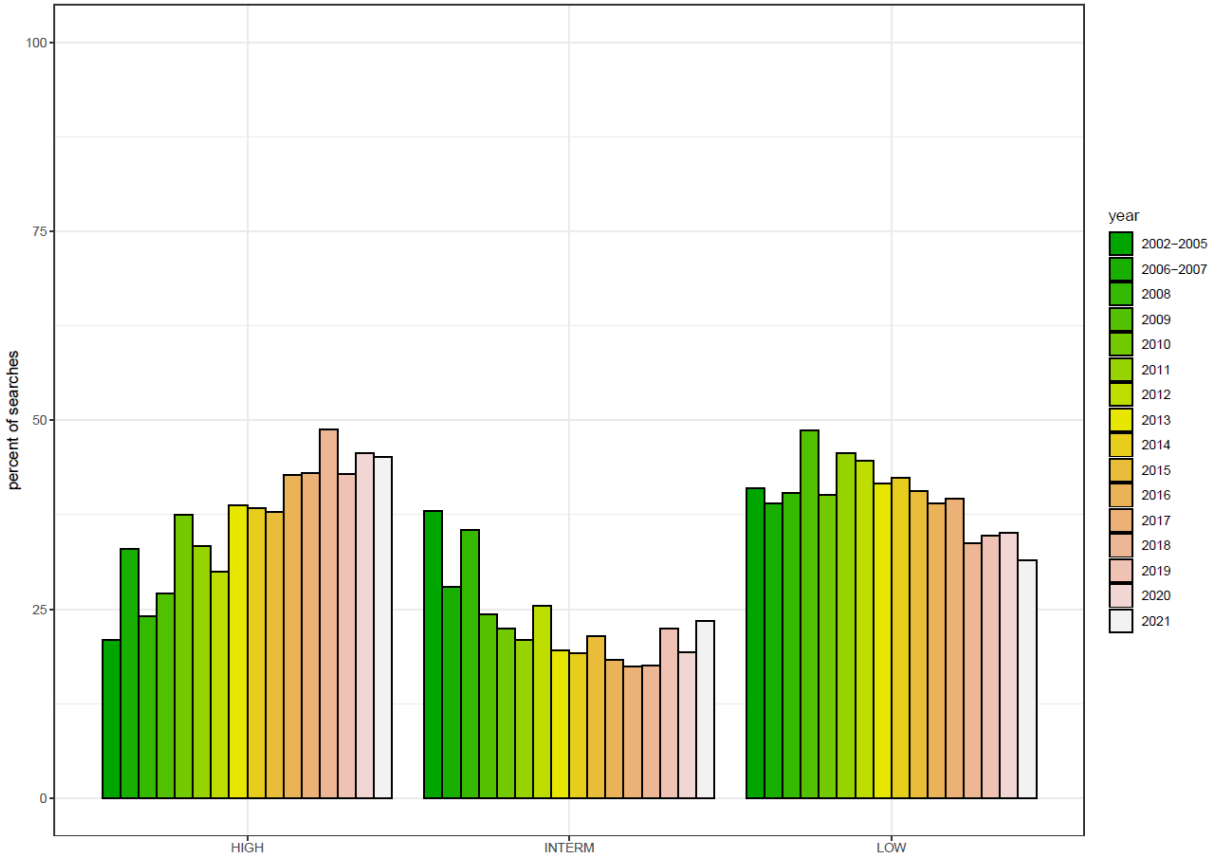

Figure 2: Probability of finding 10/10 unrelated donors in the WMDA database for alloHSCT recipients in Switzerland between 2002 and 2021. The probability is either low (no 10/10 donor is expected), intermediate (1 to 5 donors are expected to be 10/10), or high (more than 5 donors are expected to be 10/10). The percent of searches in each category is shown (personal contribution/material).

National and international guidelines for selecting the optimal donor among the different alternatives described above have been provided and are regularly updated with new advances in the field (Spellman et al. 2012, Little et al. 2016, Dehn et al. 2019, Little et al. 2021, Spellman

2022). Actually, the different donor options should be considered as complementary to each other as they increase the likelihood that every patient in need of alloHSCT will have access to a potentially suitable donor.

b. Minor histocompatibility antigens (MiHAs)

Even in the context of HLA compatible alloHSCT, achieving a proper balance between GVL and GVHD remains a challenge (Warren and Deeg 2013). Indeed, donor alloreactive T cells contained in the graft are able to recognize allogeneic peptides derived from minor histocompatibility antigens that are loaded at the cell surface by self/compatible HLA class I and II molecules (Spierings 2014). For instance, acute GVHD occurs in more than 30% of grafts with HLA identical siblings (Kanda et al. 2016, Martin et al. 2017), which should be attributable to MiHA mismatching (Roy and Perreault 2017). The presence of one or several MiHAs is anticipated between most if not all individuals (Martin et al. 2017), at the notable exception of monozygotic twins (i.e., syngeneic graft), because of the extensive variation characterizing the human genome (Genomes Project et al. 2015). More than 100 MiHAs have now been identified (Oostvogels et al. 2014, Roy and Perreault 2017) and this number is expected to increase in the future with the use of discovery approaches like whole genome association (WGA) scanning (Griffioen et al. 2016). MiHAs can be encoded by genes located both on autosomal and sexual chromosomes and are generated by diverse molecular mechanisms (Griffioen et al. 2016). Of interest, the tissue distribution of MiHAs is an important factor regarding the possible differential induction of GVL and GVHD and this has been considered as a potential immunotherapeutic approach to treat hematological cancers (Warren and Deeg 2013, Griffioen et al. 2016). In this regard, several clinical trials studying MiHAs and alloreactive T-cell responses in alloHSCT have been performed (Spierings 2014). However, MiHA matching is not part of the current strategies and guidelines for donor selection as it is practically not feasible in most situations.

c. Matching and mismatching beyond the HLA gold standards

As described above, the type of donor selected for a graft will define the degree of compatibility that can be achieved with the patient, but also the risks of complications post-alloHSCT. For instance, well-matched unrelated donors present a higher risk of transplant-related mortality than HLA identical siblings because they are often carriers of one or more HLA mismatches outside of the 10/10 compatibility (e.g., HLA-DPB1 mismatches). Moreover, MiHAs spread across the whole genome (see previous section) or non-HLA linked polymorphisms in the

extended MHC region could play a more important role in unrelated grafts than within the family where chromosome 6 is identical between matched siblings and the rest of the genome is expected to share more homology by descent. Indeed, it has been estimated that grafts from unrelated donors harbor a twofold increase in the number of genome-wide MiHA disparities compared to transplantation with matched siblings (Martin et al. 2017). Genome-wide and non-HLA variation can be observed in the form of single nucleotide polymorphisms (SNPs), expression quantitative trait loci (eQTL) or microsatellites (Mullally and Ritz 2007, Dickinson 2008, Bettens et al. 2012) and can affect gene products with immune related functions (Ishikawa et al. 2002, Keen et al. 2004, Parmar et al. 2009, Carapito et al. 2016). Some studies have reported a significant effect of SNPs within the MHC region on different clinical outcomes after alloHSCT both with matched and mismatched unrelated donors (Petersdorf et al. 2012, Petersdorf et al. 2013).

Looking at HLA genes, most algorithms for selecting unrelated donors now integrate matching at HLA-DPB1 and sometimes also consider HLA-DRB3/4/5. Future extensions to other loci like HLA-DQA1 and DPA1 and systematic inclusion of DRB3/4/5 that are now routinely typed in laboratories with the advent of NGS have been discussed (Figure 3). The first step will be to confirm initial findings suggesting that differences at these loci could be relevant on a clinical level (Detrait et al. 2015, Passweg et al. 2015, Petersdorf et al. 2022), similarly to what was previously demonstrated and is still under examination for HLA-DPB1 (Petersdorf et al. 2001, Shaw et al. 2007, Fernandez-Vina et al. 2013, Fleischhauer and Shaw 2017, Mytilineos et al. 2020, Petersdorf et al. 2020, Buhler et al. 2021).

When a 10/10 match is not possible, several considerations prevail (Petersdorf 2016, Tiercy 2016, Bertaina and Andreani 2018). Ideally, the number of HLA disparities should be restricted to one, as the risk is incremental with each additional mismatch (Petersdorf 2008). In the 9/10 setting, HLA-DQB1 mismatches should be prioritized as these disparities have been associated with better outcome than mismatches at HLA-A, B, C and DRB1 (Lee et al. 2007). Very recent data suggest that the risk can be further stratified by looking at HLA-DQ heterodimers, thus accounting for the variation presents both in HLA-DQA1 and DQB1 genes (Petersdorf et al. 2022). Among HLA genes, HLA-B disparities are associated with the worst outcome, but the risk can be stratified and minimized by considering matching for a dimorphism in the HLA-B leader peptide or B leader (Petersdorf et al. 2020, Petersdorf et al. 2020, Sajulga et al. 2022). The leader peptide is encoded by exon 1 (Figure 1) and is not expressed at the cell surface as an integral part of HLA class I molecules. Rather, it is bound (or not depending on the

dimorphism) to the peptide binding groove of non-classical HLA-E molecules and presented to receptors of the C-lectin superfamily (especially NKG2A) expressed on NK cells (Horowitz et al. 2016, Rolle et al. 2018).

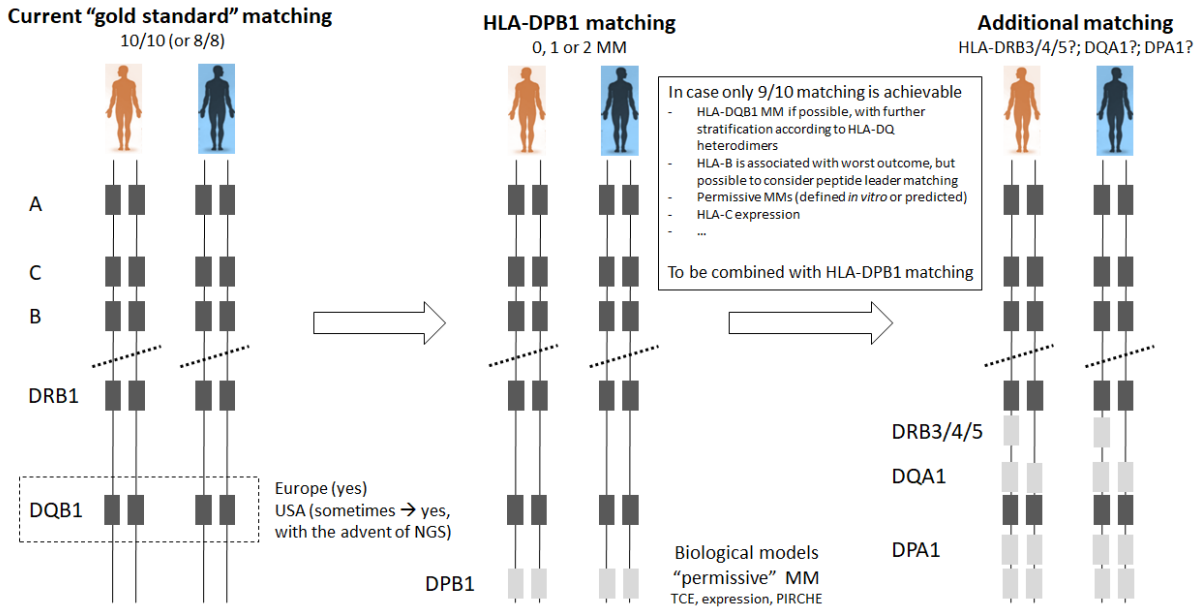

Figure 3: Ongoing and possible future extensions of unrelated donor search algorithms to match additional loci than the current HLA-A, B, C, DRB1 and DQB1 gold standard matching. In case it is only possible to achieve a 9/10 matching, several considerations prevail and are summarized in the figure (see the main text for more details). MM: mismatch (personal contribution/material).

In addition, some HLA mismatches have traditionally been considered as permissive (i.e., they induce no or only weak T-cell alloreactivity). The permissiveness has sometimes been validated using *in vitro* cellular assays (Bettens et al. 2013, Joris et al. 2014, Bettens et al. 2016) or has been predicted based on the localization of amino acid change(s) within the mismatched HLA molecules and the retrospective analysis of their impact on clinical outcome in large cohorts (Pasi et al. 2011, Fernandez-Vina et al. 2014, Tiercy 2014, Passweg et al. 2015, Petersdorf 2016). For instance, mismatches involving residues located outside of the peptide binding region, having little influence on the peptide repertoire or thought not to be seen by the T-cell receptor are usually considered as permissive in daily practice (Pidala et al. 2013). Other permissive mismatches have been defined by biological models derived from *in vitro* experiments (Zino et al. 2004, Crivello et al. 2015, Meurer et al. 2021) and then confirmed in large retrospective clinical studies, such as the T-cell epitope (TCE) matching for HLA-DPB1 (Crocchiolo et al. 2009, Fleischhauer et al. 2012, Mytilineos et al. 2020).

Polymorphisms at close vicinity of HLA genes or located within untranslated regions can influence the mRNA expression level of different alleles and thus possibly affect T-cell

recognition of the encoded allotypes at the cell surface (Bettens et al. 2016, Petersdorf 2016, Rene et al. 2016, Bettens et al. 2022, Johansson et al. 2022). Strong evidence supports a role of HLA-DPB1 allelic expression in the balance of risks between GVHD and relapse after HLA-DP mismatched alloHSCT (Petersdorf et al. 2015, Petersdorf et al. 2020, Buhler et al. 2021), while the possible relevance of HLA-C expression is more controversial (Petersdorf et al. 2014, Morishima et al. 2016). HLA loci have been classified into high (HLA-A, B, C and DRB1) and low (HLA-DRB3/4/5, DQ and DP) expression loci (HEL or LEL), respectively, with results showing a deleterious effect on clinical outcome when a single HEL mismatch or several LEL mismatches were present (Fernandez-Vina et al. 2013).

Other contributions in order to enhance HLA compatibility and reduce the burden of post-alloHSCT complications have investigated matching for phased haplotypes (Petersdorf et al. 2007), frequent haplotypes (Morishima et al. 2010, Joris et al. 2013, Buhler et al. 2019), the role of indirect presentation of HLA mismatched antigens to T cells by using the predicted indirectly recognizable HLA epitopes algorithm or PIRCHE (Thus et al. 2014, Thus et al. 2014, Thus et al. 2014, Geneugelijk et al. 2019, Buhler et al. 2021), as well as the use of ultrahigh resolution typing (Mayor et al. 2015, Mayor et al. 2019).
Further considerations on improving clinical outcome with optimized HLA matching are provided in two of the enclosed articles (Buhler et al. 2019, Buhler et al. 2021), see the summaries on pages 23, 26-27 and the articles in the annexes (see also the aims of the thesis in chapter 4). A perspective on this topic with additional references is included in the conclusion chapter.

3. Immune reconstitution post-alloHSCT

a. Immune cell subsets, timeline of reconstitution and influential parameters

The recovery of a fully functional donor-derived immune system after alloHSCT is critical for patient prognosis and central to the success of this cellular therapy. This will enable the immunocompromised patient with protection against various and often opportunistic infections and also possibly elicit a GVL curative response. Following the initial steps of conditioning and graft infusion, immune cells recover according to different timelines that are influenced by clinical variables like patient, donor and transplant characteristics as well as post-transplant complications and interventions (Figure 4). This has been described extensively (Storek et al. 2008, Bosch et al. 2012, Ogonek et al. 2016), including with the use of novel and highly

sensitive techniques (Stern et al. 2018). After an aplastic phase of about two weeks caused by the conditioning and characterized by severe neutropenia (pre-engraftment phase), innate cells reconstitute from engrafted donor $CD34^+$ stem cells in the following timeframe: neutrophils (~15-30 days), monocytes, dendritic cells and NK cells (~1 month). NK cells represent the dominant circulating lymphocyte population in the first three months following engraftment and although they reconstitute quickly from a quantitative aspect, their functional recovery can take up to several months (Stern et al. 2018). Regarding adaptive immunity, $CD8^+$ T cells usually attain physiological values within the first year post-transplant, while CD4+ T cells do not reach standard counts and full function before two years or even longer. The reconstitution of the T-cell compartment is described more in details in the following chapter sections and has been recently reviewed (Dekker et al. 2020). Humoral immunity provided by B cells is the slowest to recover. Normal counts of B cells can be reached within 1-2 years, although different populations recover according to different period of times, but their function often remain compromised for much longer (Ogonek et al. 2016). This is partly due to the delayed recovery of T cells and also because of long-term deficiencies in antibody class switching and diminished somatic hypermutation in some patients.

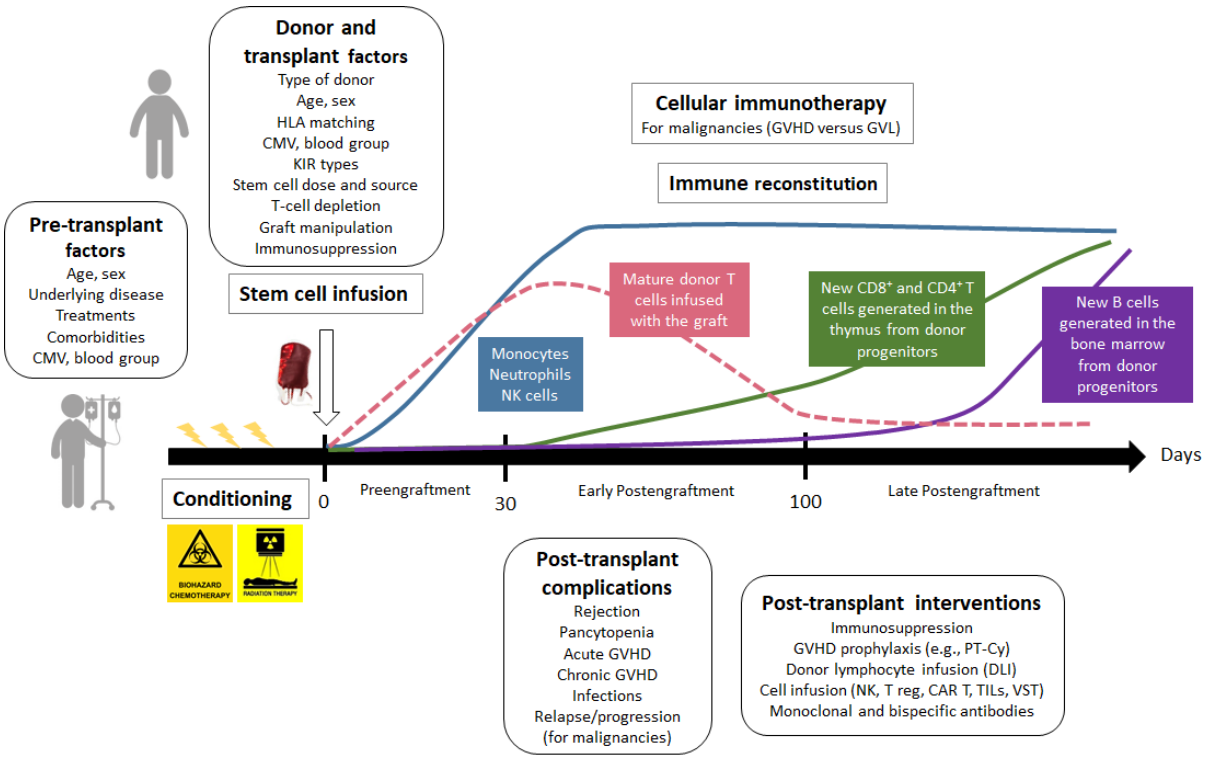

Figure 4: Schematic representation of the main steps of allogeneic hematopoietic stem cell transplantation and the influential parameters associated with immune reconstitution after the procedure (personal contribution/material).

The kinetics of immune recovery as well as immunological outcomes are influenced by the clinical profile of the patient (e.g., age, sex, type of disease, treatments and comorbidities). They

also depend on the type of donor (Anasetti et al. 1989, Kollman et al. 2001, Baron et al. 2006, Kollman et al. 2016, Baumeister et al. 2020) and are associated to some transplant factors (Figure 4). Notably, it is well established that the source of stem cells, including total nucleated cell (TNC) dose and the number of $CD34^+$ cells within the graft, is an important factor contributing to the pace and pattern of reconstitution (Seggewiss and Einsele 2010, Martin et al. 2016). Post-transplant complications like GVHD, viral reactivations, infections and relapse of malignancy will also depend on the type of graft, as PBSC, BM and CBU will differ regarding infused T cells both in quantitative and qualitative aspects (Ogonek et al. 2016). The conditioning regimen (Gyurkocza and Sandmaier 2014), graft manipulations such as T-cell depletion (Vadakekolathu and Rutella 2017) and immunosuppression are other important parameters that can impact the post-transplant course. In the context of haploidentical grafts, a limited number of studies have shown that the immune reconstitution of distinct cell subsets sometimes differ significantly from the kinetics well-established in the HLA matched setting. Both faster and delayed recoveries have been reported depending on the type of cells analyzed or on the study, as reviewed in Baumeister et al. (2020). However, the authors also argue that the type of haplo-platform used as well as the GVHD prophylaxis could explain a substantial part of the observed differences (e.g., T-cell replete grafts combined with subsequent *in vivo* T-cell depletion or T-cell depleted grafts with *ex vivo* manipulations).

Assessment of immunological recovery is an integral part of the clinical follow up of patients after transplantation. This allows to monitor the risks of complications and guide the need for interventions (Stern et al. 2018). Recent technological developments such as mass cytometry (cytometry by time of flight or CyTOF) that enables the simultaneous monitoring of a large number of immune cells, in combination with protein profiling, are paving the way to improve standard of care and post-HSCT follow-up (Lakshmikanth et al. 2017). In addition, a complementary approach to the system-level analyses provided by CyTOF is the in-depth examination of the T-cell repertoire and its reconstitution after alloHSCT using high throughput sequencing of the T-cell receptor (TCR) also called immunosequencing (Robins 2013, Six et al. 2013, Heather et al. 2018).

b. Reconstitution of the T-cell compartment

Several studies have shown that lymphocyte recovery in terms of absolute count in the first trimester after alloHSCT is a major determinant of patient outcome (Kim et al. 2004, Savani et al. 2007, Le Blanc et al. 2009). After the conditioning regimen, homeostatic proliferation of

mature donor T cells infused with the graft transiently fills the void in the T-cell compartment (Figure 4). Homeostasis controls the size of the T-cell pool and is a tightly regulated process mediated by internal stimuli including cytokines and the presentation of self pHLA complexes to the TCR (Jameson 2002). The composition of the T-cell repertoire at that stage is also reflecting peripheral cytokine-driven expansions of mature donor T cells (and possibly of some residual recipient cells), both from the naïve and memory compartments, after their encounter with foreign antigens (Berard and Tough 2002, Macedo et al. 2009, Allen et al. 2011). Notably, alloreactive T cells can recognize alloantigens (i.e., mismatched HLA or MiHAs) and pathogen-specific T cells can react to latent virus present in the recipient such as Cytomegalovirus (CMV) and Epstein-Barr virus (EBV). This thymus-independent pathway is linked to the so-called functional diversity of T cells (Figure 5). It usually results in a skewed and constricted repertoire that is closely associated with infections and GVHD (Mackall et al. 1996, Suessmuth et al. 2015, Ogonek et al. 2017, Inman et al. 2018, Phan et al. 2018).

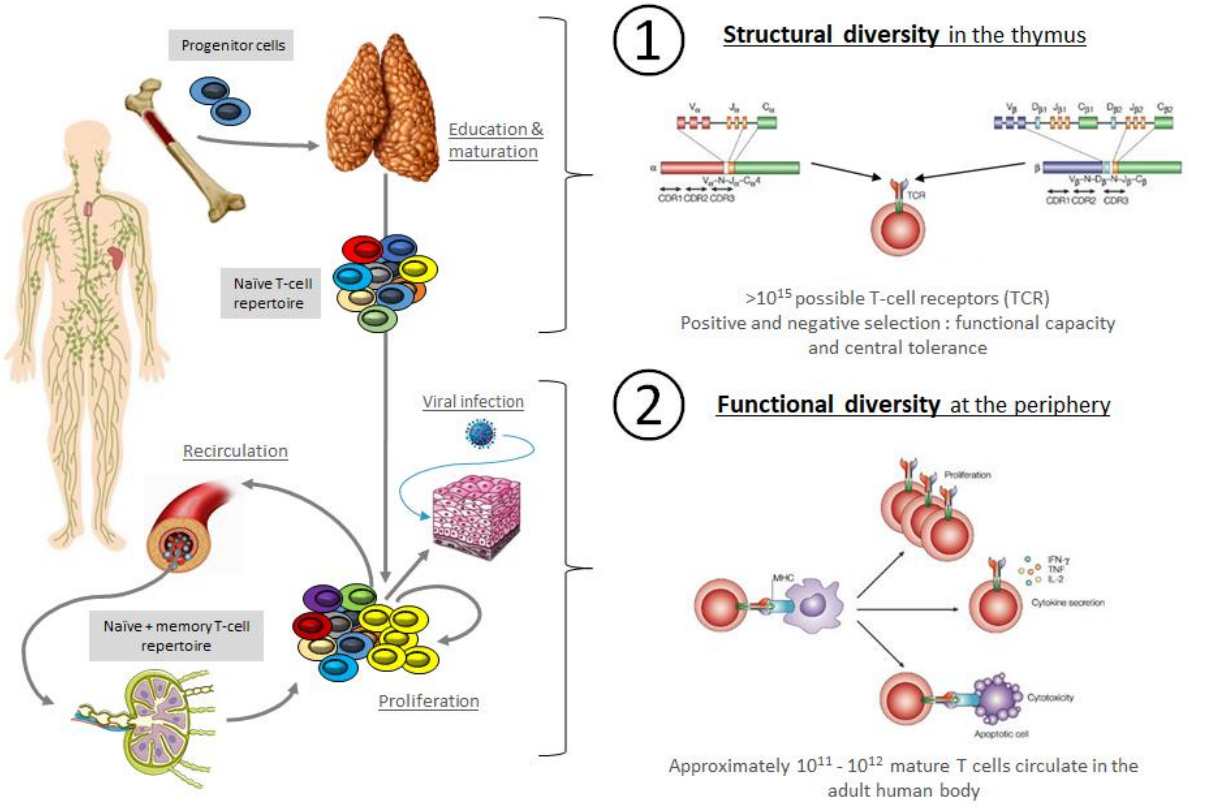

Figure 5: The T-cell repertoire and the main mechanisms shaping its diversity, partly adapted from Nikolich-Zugich et al. (2004).

Subsequently, *de novo* production of naïve T cells will start in the thymus following the migration of lymphoid precursors derived from the graft (Cosway et al. 2021). This thymus-dependent pathway is expected to restore a structurally broad, tolerant and fully responsive repertoire (Figure 5). However, the thymus is highly sensitive to the independent or combined

effects of conditioning, immunosuppression, infectious episodes and acute GVHD, which could impair the proper reconstitution of the repertoire (Weinberg et al. 2001, Chaudhry et al. 2017, Velardi et al. 2021). It is also anticipated that age of the recipient affects T-cell neogenesis (Hakim et al. 2005, Mittelbrunn and Kroemer 2021). Age of the donor is another important parameter that has been shown to delay immune recovery (Baron et al. 2006). In addition, clinical interventions such as antithymocyte therapy (Link-Rachner et al. 2018) or prophylactic drugs like letermovir inducing a reduced CMV exposure (Zamora et al. 2021) could have some influence on the shape of the repertoire.

The mechanisms of T-cell reconstitution post-alloHSCT and the influence of clinical parameters (including HLA matching) were recently investigated by immunosequencing in a large cohort of 116 patients transplanted in Geneva and in their corresponding donors (Buhler et al. 2020). More details on this study can be found in the summary on pages 24-25 and in the annexed article (see also the aims of the thesis in chapter 4).

c. The alloreactive T-cell repertoire

The different mechanisms driving the constitution and composition of the T-cell repertoire at the periphery have been described (Berard and Tough 2002, Allen et al. 2011). In the setting of alloHSCT, alloreactive donor T cells can proliferate following direct or indirect recognition of HLA mismatches (Geneugelijk et al. 2014) or after recognizing peptides derived from MiHAs and presented by self/compatible HLA molecules. The cognate interactions involving highly polymorphic HLA molecules presenting very diverse repertoires of self and non-self peptides on antigen presenting cells to somatically rearranged TCRs expressed on T lymphocytes underlies the potential of T cells for alloreactivity (Felix and Allen 2007). Up to 10^{12} T cells circulate in the adult body with a naïve repertoire constituted by more than 10^{8} distinct TCR clones (Nikolich-Zugich et al. 2004, Attaf et al. 2015). Yet, this diversity is dwarfed by the vast array of peptides that T cells must recognize (Sewell 2012). Thus, in order for adaptive immunity to operate, the basis of pHLA-TCR interactions has been described as flexible and crossreactive (Yin and Mariuzza 2009). For instance, it has been shown that up to one million peptides can be specifically recognized by a given TCR (Wooldridge et al. 2012), while many distinct TCRs should be able to recognize the same antigenic peptide (Vujovic et al. 2020). A delicate balance between specificity and degeneracy defines T-cell epitopes, with key peptide residues in contact with the TCR constraining cross-reactivity and flexibility being allowed at

other less important positions (Birnbaum et al. 2014). This promiscuity of pHLA-TCR interactions is the main culprit of T-cell mediated alloreactivity in combination with the fact that TCRs potentially reactive to HLA allotypes are not eliminated by negative selection in the thymus (DeWolf and Sykes 2017). Studies have demonstrated that virus-specific memory T cells can strongly cross-react with alloantigens (Degauque et al. 2016, D'Orsogna et al. 2017) and that a high proportion (probably more than 10%) of TCRs are alloreactive (DeWolf et al. 2018), with both memory antiviral and naïve T cells contributing significantly (Macedo et al. 2009, Benichou et al. 2017, Dekker et al. 2022). Interestingly, the breadth of the alloreactive repertoire against different HLA allotypes was very similar and diverse despite different levels of alloreactivity as measured by the frequency of $CD4^{+}CD137^{+}$ lymphocytes in cell cultures (Arrieta-Bolanos et al. 2018).

Other recent data suggest that the alloimmune response occur very early after transplantation, with a polyclonal TCR repertoire observed after two weeks in autografts compared to oligoclonal profiles in allografts. Moreover, the patients that subsequently developed acute GVHD were characterized by a restricted TCR repertoire in term of diversity (Inman et al. 2018).

By combining one-way mixed lymphocyte reactions (MLR) with selected HLA class I mismatches and immunosequencing of the responding T cells, we have investigated the alloimmune response and whether it could be predictable as described on page 28 and in the annexed article (see also the aims of the thesis just below).

4. Aims of this thesis

As mentioned above, a main mission of the National Reference Laboratory for Histocompatibility is to provide a report to the Swiss Blood Stem Cell (SBSC) registry and the four allogeneic transplant centers located in Geneva, Basel and Zurich for the optimal HLA based selection of unrelated donors for residents in need of alloHSCT. The laboratory is thus actively involved since many years in groups like the International Histocompatibility Working Group in Hematopoietic Cell Transplantation (IHWG-HCT) which purpose is a collaborative effort to increase the availability and efficacy of hematopoietic stem cell transplantation from alternative donors through an improved understanding of the genetic barrier (Petersdorf et al. 2013, Petersdorf et al. 2020, Petersdorf et al. 2020, Petersdorf et al. 2022). The laboratory also maintains a large database including all patients from Switzerland who have been transplanted

with unrelated donors. This large cohort has allowed to perform retrospective analyses on clinical outcome and HLA matching. This has led to improvements in the strategy for unrelated donor selection in Switzerland such as the early introduction of prospective HLA-DPB1 genotyping (Bettens et al. 2012) or the prioritization of mismatches considered as more permissive (Passweg et al. 2015). More recently, the cohort was investigated to validate algorithms that were proposed in the literature regarding matching for HLA haplotypes (Buhler et al. 2019) and biological models to predict HLA-DPB1 permissive mismatches (Buhler et al. 2021). Both articles are summarized below and are enclosed in the annexes.

Thanks to the access to samples from all recipients and donors and its expertise in cellular and molecular biology, the laboratory is exploring other aspects of the HLA polymorphism and their translational application to alloHSCT. Notably, we investigated the regulation of HLA class I expression using RNA sequencing and by performing unstimulated and cytokine-stimulated cell cultures (Bettens et al. 2022). We also examined the alloreactive T-cell repertoire using *in vitro* stimulation of T lymphocytes with a set of selected HLA disparities and by performing immunosequencing of the responding alloreactive clonotypes (Bettens et al. 2020). Both articles are summarized below and are enclosed in the annexes.

In addition, since several years the laboratory is interested by the TCR repertoire reconstitution after alloHSCT. Allografts represent a unique model in humans to study the dynamics of the reconstituting repertoire and to disentangle the different mechanisms that drive and shape its composition. On the clinical side, it is important to characterize the main factors associated with the repertoire's diversity and the parameters that could impact alloHSCT outcome (Buhler et al. 2020). The article is summarized below and is enclosed in the annexes.

To resume, the thesis explores several of the latest advances in transplantation immunology, histocompatibility and immunogenetics, both at the pre and post-transplantation stages, with the goal to better monitor and predict clinical outcome as well as to translate these results to the standard of care for alloHSCT patients.

Summary of article 1

High resolution HLA phased haplotype frequencies to predict the success of unrelated donor searches and clinical outcome following Hematopoietic Stem Cell Transplantation

S. Buhler, H. Baldomero, S. Ferrari-Lacraz, J.M. Nunes, A. Sanchez-Mazas, S. Massouridi-Levrat, D. Heim, J. Halter, G. Nair, Y. Chalandon, U. Schanz, T. Güngör, G. Nicoloso, J.-M. Tiercy, J. Passweg, J. Villard

Bone Marrow Transplantation (2019); 54: 1701–1709

The current strategy for selecting unrelated donors in alloHSCT is to achieve a 10/10 HLA phenotypic match based on high resolution typing at HLA-A, -B, -C, -DRB1 and -DQB1. Some studies have proposed that further matching at the genotypic level (i.e., matching for both haplotypes carried by the donor and recipient) could be beneficial to improve patient prognosis. The underlying concept is that by matching for haplotypes the probability is increased that other polymorphic genes and minor histocompatibility antigens located in the MHC region will also be matched because of the strong linkage disequilibrium in this part of the genome, thus improving the genetic compatibility between donor and recipient.

In this study, we determined the high resolution phased HLA haplotypes of 291 patients from Geneva by segregation analysis with family relatives. We then assessed whether haplotypes could be used as a surrogate predictor of a successful unrelated donor search. A sample of 211 patients who received a graft from a 10/10 matched unrelated donor (i.e., comprising 101 patients from the initial cohort and additional patients from two other Swiss transplant centers) was subsequently analyzed to investigate a putative impact of haplotypes on overall survival, GVHD and relapse/progression.

We could confirm that the probability of finding a 10/10 matched donor in the World Marrow Donor Association registry is determined by the ranks of the haplotypes carried by the patient. Such predictions are important to define the optimal strategy to select the best suitable donor among several alternatives (i.e., a matched or mismatched unrelated donor or a haploidentical donor). Regarding clinical outcomes, we did not find a significant impact of haplotypes. The controversial effect of haplotypes in different cohorts could be related to heterogeneities between studies at defining common haplotypes, differential impacts of haplotypes in populations with variable levels of HLA diversity, HLA loci not considered in the phased haplotypes such as HLA-DPB1 or a lack of statistical power due to the extreme polymorphism observed in the HLA region.

Summary of article 2

Genetic T-cell receptor diversity at 1 year following allogeneic hematopoietic stem cell transplantation

S. Buhler, F. Bettens, C. Dantin, S. Ferrari-Lacraz, M. Ansari, A.C. Mamez, S. Masouridi-Levrat, Y. Chalandon, J. Villard

Leukemia (2020); 34: 1422-1432

As stated in the introduction, the recovery of a fully functional immune system after alloHSCT is essential for patient prognosis. Several events such as conditioning, infections, especially by viruses, and graft-versus-host disease can influence the reconstitution of a new T-cell repertoire, a process that can take up to one or two years after engraftment. Recently, post-transplant monitoring of the TCR repertoire has gained ground as a strategy to improve standard of care of allograft recipients. This has notably been made possible thanks to the development of high throughput sequencing protocols that allow the simultaneous analysis of tens or hundreds of thousands of T-cell clonotypes.

In this study, we analyzed the TCR CDR3β region by immunosequencing in a large cohort of 116 full chimeric recipients at one year post-HSCT and their donors collected just before transplantation. We first investigated the extent of TCR repertoire overlap between donor and recipient. We could show that the overlap is actually very low, an observation suggesting that the repertoire is principally reconstituted *de novo* despite the presence of large amounts of donor T cells in the graft. We then assessed which clinical parameters were associated with the significant reduction in diversity that was observed between pre and post-HSCT repertoires in most donor/recipient pairs. Among the more than 20 parameters surveyed, only four yielded a significant association: age of the recipient, age of the donor, the CMV serologic status of both donor and recipient and CMV infection/reactivation in the recipient after transplantation. Indeed, increasing age, CMV seropositivity of donor and recipient (i.e., D^+/R^+) and the detection of CMV DNA in plasma of recipients with or without symptoms independently led to a shift of the repertoire diversity toward oligoclonality at one year post-HSCT. We could also show that CMV-specific clonotypes were enriched after HSCT, especially in case of a CMV infection/reactivation. Another outstanding feature of the post-HSCT repertoire was the presence of dominant clonotypes of unknown specificity exhibiting highly expended frequencies. These clonotypes were overrepresented in D^+/R^+ pairs and in recipients suffering from severe acute GVHD. We speculated that these clonotypes of unknown specificity can be

truly CMV specific but are not yet characterized. Alternatively, they could represent bystander nonspecific T cells triggered by inflammatory cytokines secreted during the alloimmune response or a CMV infection. Interestingly, the diversity of the repertoire at one year was not predictive of subsequent clinical events.

Summary of article 3

Analysis of biological models to predict clinical outcomes based on HLA-DPB1 disparities in unrelated transplantation

S. Buhler, H. Baldomero, S. Ferrari-Lacraz, A.-C. Mamez, S. Masouridi-Levrat, D. Heim, J. Halter, G. Nair, Y. Chalandon, U. Schanz, T. Güngör, G. Nicoloso, J.R. Passweg, J. Villard

Blood Advances (2021) 5: 3377-3386

As described in the introduction, the current gold standard in unrelated HSCT is to achieve a 10/10 match between donor and recipient (i.e., compatibility at HLA-A, -B, -C, -DRB1 and -DQB1). Nowadays, prospective typing has allowed to optimize donor selection for another important HLA gene, the DPB1 locus located in the class II region. Several studies have shown that HLA-DPB1 mismatches are associated to an increased risk of GVHD but this risk is counterbalanced by a reduced occurrence of relapse. Although for many patients it is now possible to identify donors that are matched for both DPB1 alleles (i.e., 12/12 match), a substantial number of patients don't have a donor that is fully matched at this locus. For this reason, several biological models have been proposed to identify so-called permissive mismatches (i.e., inducing no or less harmful effects such as acute GVHD). One model is based on the immunogenicity of DPB1 mismatches defined by T-cell epitopes (TCE). Another model is based on the level of cell surface expression of DPB1 molecules defined by a polymorphism located in the 3' untranslated region of the gene. Finally, a third model is based on the indirect recognition of allogeneic peptides derived from mismatched HLA molecules (Predicted Indirectly ReCognizable HLA Epitopes, PIRCHE).

In this retrospective study, we compared clinical outcomes for the classical allele matching model and the three biological models in a large number of patients (n=909) of the Swiss cohort who received a 10/10 matched allograft between 2008 and 2018. We could demonstrate that the best option to prevent acute GVHD remains the selection of a 12/12 donor. In addition, when a donor with one DPB1 allele mismatch was considered, the lowest risk of acute GVHD was when the mismatch involved an allele with a low expression profile in the recipient, followed by a permissive TCE mismatch and the absence of PIRCHE II (i.e., allogeneic HLA-derived peptides presented in HLA class II). For situations with two DPB1 mismatches, a permissive TCE mismatch and no potential for PIRCHE II against the recipient were favorable constellations. The expression model was not considered as it has been conceived only for situations involving one mismatch. In addition, the balanced risk between acute GVHD and

relapse was observed in all statistical models that were explored. In conclusion, this study points toward an integrative donor selection strategy that combine biological models with classical matching in order to improve post-HSCT complications and patient prognosis.

Summary of article 4

CD8+ T-cell repertoire in HLA class I-mismatched alloreactive immune response

F. Bettens, Z. Calderin Sollet, S. Buhler, J. Villard

Frontiers in Immunology (2021); 11, 588741

Direct allorecognition of nonself pHLA complexes at the cell surface by T cells involves multifaceted ligand-receptor cognate interactions. It is proposed that up to one-tenth of circulating T cells are alloreactive, notably due to their intrinsic crossreactive potential. In alloHSCT, this dynamic process can be associated to the beneficial GVL effect but can also lead to GVHD.

In this study, we used one-way MLRs to define the TCR CDR3β repertoire of alloreactive cytotoxic $CD8^+$ responder T cells. These cells were defined by the cell surface expression of CD137 and were triggered with different combinations of HLA mismatched stimulatory cells. In order to control for HLA disparities, we used responder and stimulatory cells collected from distinct healthy blood donors with the same mismatched HLA molecule(s) as well as MLR replicates. To mimic *in vivo* inflammatory processes, we upregulated the cell surface expression of HLA class I molecules with pro-inflammatory cytokines TNFα and IFNβ.

We could show that stimulatory cells carrying the same HLA mismatch induced alloreactive $CD8^+CD137^+$ cells with very different repertoires in distinct responders phenotypically matched for HLA. Moreover, stimulatory cells with the same HLA background induced very different repertoires of alloreactive cells $CD8^+CD137^+$ in the same HLA mismatched responder. Also, the repertoires differed significantly when the same responder cells were induced with stimulatory cells with or without upregulated cell surface expression of HLA class I antigens. Interestingly, the presence of many HLA mismatches induced a more diverse alloreactive repertoire than stimulatory cells carrying only one or a few alloantigens. In conclusion, our data demonstrate that the repertoire of alloreactive cytotoxic T cells is very diverse and that their expansion is hardly predictable following stimulation with alloantigens, even when controlling for HLA disparities. Tailoring the alloimmune response remains a major challenge in HLA mismatched alloHSCT.

Summary of article 5

Regulation of HLA class I expression by non-coding gene variations

F. Bettens, H. Ongen, G. Rey, S. Buhler, Z. Calderin Sollet, E. Dermitzakis, J. Villard

PloS genetics (2022); 18: e1010212

The degree of matching at highly polymorphic HLA genes is critical for different outcomes after solid organ and hematopoietic stem cell transplantation. The impact of allelic diversity on the cell surface expression of MHC molecules and the putative role of expression on clinical outcomes remain under investigation in the context of the alloimmune response between donors and recipients.

In this study, we quantified the allelic expression of HLA class I loci (HLA-A, B and C) by RNA sequencing and conducted an analysis of expression quantitative traits loci (eQTL) to investigate whether HLA expression regulation could be associated with non-coding gene variations. To mimic the pro-inflammatory environment that underlies the post-transplantation setting, we stimulated peripheral blood mononuclear cells (PBMCs) with cytokines TNFα and IFNβ before quantifying the expression.

We could demonstrate that HLA-B alleles exhibit the highest levels of expression followed by HLA-C and HLA-A alleles. Interestingly, the results suggested a coordinated and paired expression of both alleles at each locus in every individual. The analysis of individuals carrying common and conserved HLA class I haplotypes showed that the differences in expression were mostly determined at the individual level. The cytokine-induced upregulation of HLA class I RNA and cell surface expression did not change the coordinated profiles of allelic expression observed in all individuals. We identified cis eQTLs explaining 29%, 13%, and 31% of the respective variance in HLA-A, B, C expression in unstimulated cells and 9%, 23%, and 50% of the variance in cytokine-stimulated cells, respectively. The identified eQTLs were shown to be independent from the coding variation in HLA alleles and thus may influence the intra-allelic variability in expression although they might not represent the causal eQTLs.

Conclusions and perspectives

The field of alloHSCT has greatly benefited from progresses in medicine and new technological developments. This encompasses the possibility of fast and multiplexed allelic/high resolution typing of all HLA genes in histocompatibility laboratories (Barone et al. 2015, Mayor et al. 2015, Monos and Maiers 2015, Bravo-Egana et al. 2021, Liu 2021), the accessibility to millions of well-typed volunteer adult donors with improved matching prediction algorithms, see (Dehn et al. 2016) or https://searchmatch.wmda.info/, as well as new tailored post-transplantation interventions like chimeric antigen receptor (CAR) T cells, donor lymphocyte infusion (DLI), virus specific T cells (VST), regulatory T cells (Treg), tumor infiltrating lymphocytes (TILs) and new drugs like monoclonal and bispecific antibodies (Chabannon et al. 2018, Brown and Mackall 2019, Ma et al. 2021). Furthermore, a better understanding of the genetic barrier, as introduced in the first chapter of this thesis and discussed hereafter, and new strategies and protocols to control post-transplant complications like the use of post-transplant cyclophosphamide (PT-Cy) as GVHD prophylaxis have considerably improved the safety of alloHSCT.

The introduction of an intensive PT-Cy based prophylaxis in the setting of haploidentical grafts (O'Donnell et al. 2002, Luznik et al. 2008), as a way to induce a strong immune tolerance by "erasing" the HLA barrier, is frequently considered the most important advance within the past 15 years for alloHSCT (Mussetti et al. 2021). It enabled haploidentical transplantation to achieve comparable results to grafts from matched unrelated donors receiving a traditional GVHD prophylaxis (Ciurea et al. 2015). This opened a debate regarding which type of donor is a better choice. Notably, it was argued that clinical endpoints like engraftment, relapse and GVHD or parameters such as donor availability and rapid access to a suitable donor are also important to consider besides the initial good results on overall survival, with some favoring MUDs and others supporting related haploidentical donors as a better alternative (Fuchs 2017, Shaw 2017). This remains a hot topic in the community to these days (Lalli 2023). Furthermore, since the very promising results led to the widespread use of haploidentical donors (Passweg et al. 2017), a legitimate question was what to expect from this efficient protocol in grafts with MUD, in whom significant risks of acute GVHD persist despite standard GVHD prophylaxis (Bacigalupo 2021). A recent study provided some clues by comparing both types of donors in patients with acute leukemia or myelodysplastic syndrome and receiving the same PT-Cy prophylaxis (Gooptu et al. 2021). The main conclusion was that HLA compatibility remains a

crucial parameter, as the risks of acute GVHD were decreased, while disease-free and overall survival were higher with MUD compared to haploidentical donors. Interestingly, the reduction of GVHD was not accompanied by a higher risk of relapse in this cohort. In two other studies, the data were partially but not fully concordant with the results of Gooptu et al (2021). Haploidentical grafts were associated with increased risks of GVHD and nonrelapse mortality (NRM), but with lower relapse incidence resulting in no difference in survival compared to MUD or matched sibling donors (MSD) in patients with AML (Sanz et al. 2020), while no differences among the three donor groups were observed in patients with ALL (Sanz et al. 2021). In another cohort, the impact of the source of stem cells was investigated and associated with a similar survival but lower risks for acute and chronic GVHD in haplo-BM compared to MUD-PBSC grafts (Nagler et al. 2021). Clinical trials with PT-Cy have also been extended to mismatched unrelated donor transplants, with the goal to increase access to alloHSCT for more patients (Al Malki et al. 2021, Shaw et al. 2021, Battipaglia et al. 2022).

Although the type of disease, source of stem cells, studied cohort and probably many other parameters are influential on the clinical endpoints as illustrated above, HLA still matters for predicting and minimizing post-transplant risks in alloHSCT despite the availability of PT-Cy prophylaxis (Bacigalupo 2021, Gooptu et al. 2021, Fleischhauer 2022, Fuchs et al. 2022, Spellman 2022). Besides the gold standard matching at the genotype (i.e., in HLA identical grafts with MSD) or phenotype levels (i.e., in 10/10 grafts with MUD), several other possible approaches to achieve a better genetic compatibility have been investigated in recent years, as described in the introduction. Some approaches are not straightforward or feasible to implement in the routine or have simply not led to consensual recommendations, like matching for HLA haplotypes (Tay et al. 1995, Petersdorf et al. 2007, Morishima et al. 2010, Joris et al. 2013, Buhler et al. 2019).

By contrast, the importance of HLA-DPB1 matching has been and still is comprehensively investigated with contributions of various teams and under the guidance of international collaborative efforts. A significant role of mismatches at this locus has been described since a long time (Petersdorf et al. 2001, Shaw et al. 2007, Shaw et al. 2007, Fleischhauer and Shaw 2017). Although upfront HLA-DPB1 typing now allows to identify 12/12 matched unrelated donors for a substantial number of patients, around 60 percent of patients currently don't benefit from such donors (Buhler et al. 2021). This has led to other strategies to define some levels of biological permissiveness when one or two DP mismatches cannot be avoided. Three main models have been examined and are discussed hereafter (for additional details see also our

annexed article). Currently, the TCE model is the most used and it has been integrated in several donor algorithms such as HapLogic (Dehn et al. 2016), OptiMatch or Hap-E (https://searchmatch.wmda.info/). It is applicable to every DP mismatch constellation (i.e., one or two mismatches, either bidirectional or in the graft-versus-host and host-versus-graft direction). It has been validated in large retrospective cohorts (Crocchiolo et al. 2009, Fleischhauer et al. 2012, Pidala et al. 2014, Lorentino et al. 2019, Mytilineos et al. 2020, Petersdorf et al. 2020, Buhler et al. 2021) and has been continuously upgraded with new data, notably coming from *in vitro* experiments. For instance, the initial model was mainly based on the concept of direct allorecognition of T-cell epitopes on different DP molecules (Zino et al. 2004, Zino et al. 2007) and the permissiveness of groups of alleles defined by amino acid sequence similarity at specific positions of the peptide binding region (Crivello et al. 2015, Crivello et al. 2016). More recent analyses suggest a significant role of the immunopeptidome presented by HLA-DP molecules (Meurer et al. 2021). Moreover, a core group of alleles with structural similarity in the peptide binding region was defined as a main driver of permissiveness in DP mismatched allografts (Arrieta-Bolanos et al. 2022). The TCE model has been significantly associated to different clinical endpoints such as survival, TRM, acute GVHD and relapse, although with some heterogeneity among the studies cited above.

In parallel, the HLA-DPB1 expression model has been gaining more attention as it can now be predicted for most DPB1 alleles (Schone et al. 2018) and has proven to be informative of the balanced risks between relapse and GVHD in several independent cohorts (Petersdorf et al. 2015, Lorentino et al. 2019, Petersdorf et al. 2020, Buhler et al. 2021). Actually, in the Swiss cohort, the expression model was more informative than the TCE model to predict post-transplant risks and is considered first in the LNRH algorithm for the selection of unrelated donors with a mismatch at the HLA-DPB1 locus. However, one limitation is that it can only be considered in situations of a single mismatch with a graft-versus-host vector (i.e., in practice it is only applicable to patients who have one high expressed allele matched and one low expressed allele mismatched with their respective donors). In addition, the PIRCHE model is intriguing for predicting outcome in DP mismatched allografts (both with one or two mismatches, but only in the graft-versus-host direction), as shown by our study and the initial analysis performed in a small cohort of patients (Thus et al. 2014).

Since there are several possible and non-mutually exclusive ways of looking at the permissiveness of HLA-DPB1 disparities, this has raised several questions. Which approach is the most informative, if any? Are the models acting synergistically or independently from each other? Are the effects additive and is it possible to combine the different models to improve

risk prediction? One difficulty to answer these questions is that there is a non-negligible overlap between the three models (i.e., mismatches conferring a high risk in a given model are often expected to do the same in the other models), as shown by us and others (Fleischhauer 2015, Meurer et al. 2018, Buhler et al. 2021). Individual analyses of both the TCE and expression models and their direct comparison in several studies have not allowed to clarify whether they can be integrated and if their integration results in a better prediction of post-transplant risks (Morishima et al. 2018, Lorentino et al. 2019, Mytilineos et al. 2020, Petersdorf et al. 2020, Buhler et al. 2021). A very recent study has tried to address these questions and suggest that donor/recipient pairs belonging to the subset carrying a TCE-permissive mismatch combined with a high expressed mismatch (abbreviated as TPHE) are associated with better relapse-free survival compared to matched and non-TPHE mismatched pairs (Ruggeri et al. 2023). Here the hypothesis is that both models synergize in the context of TPHE mismatches. Alloreactive T cells in the TCE-permissive setting, i.e., characterized by a constrained TCR repertoire and having a mitigated potential for causing GVHD but some capability to promote GVL (Fleischhauer and Beelen 2016, Fleischhauer and Shaw 2017, Meurer et al. 2021), are most effective in presence of the high cell surface expression of a DP mismatch on malignant cells. However, this hypothesis remains to be formally demonstrated in additional studies (Pidala and Anasetti 2023).

A significant part of this discussion is focusing on the HLA-DPB1 paradigm. Looking from the histocompatibility perspective, we believe that it nicely illustrates the different innovations and progresses that have been made and are still needed to overcome HLA disparities and improve patient's prognosis. It is also paving the way for future works aimed at defining donors with optimized pre-transplant immunogenetic markers. Indeed, once the importance of genetic compatibility was demonstrated at this locus and as we have seen, the biology of HLA-DPB1 disparities was dissected under different angles and by integrating several theoretical aspects of allorecognition, notably the direct and indirect pathways. The different models have also been continuously updated with new data in the field, improved by new technological developments and comprehensively investigated using elegant *in vitro* assays. This has allowed to contemplate an integrative algorithm of permissive matching at the HLA-DPB1 locus (Ruggeri et al. 2023). This algorithm is still under refinement and will certainly be even more tailored to patient's benefits with validation data from future international collaborations (Pidala and Anasetti 2023). The consideration and integration of complementary matching/permissiveness models like PIRCHE should also follow soon as suggested by our study in the Swiss cohort and

the initial study of Thus et al (2014). Several studies supporting the use of PIRCHE in other alloHSCT settings like 9/10 allografts (Thus et al. 2014, Ayuk et al. 2019, Geneugelijk et al. 2019, Stenger et al. 2020) and haploidentical grafts (Rimando et al. 2020, Grubic et al. 2022) have already been published.

Recently, a study in a cohort of patients treated by haplo-PT-Cy reported no effect of the number of HLA mismatches but a significant association of the TCE and B leader models as well as the choice of the mismatched locus (i.e., DRB1 mismatch and DQB1 match) with different clinical endpoints (Fuchs et al. 2022). A web tool where relevant clinical and HLA characteristics can be entered to predict disease-free survival (DFS) for any given patient was developed and is publicly available (http://haplodonorselector.b12x.org/v1.0/). This latest study is in line with the prevailing concept of adjusting and optimizing but not totally erasing the HLA barrier in an integrative way and in different transplant settings to support donor selection, here haploidentical grafts where patients often have several potential family donors (Fleischhauer 2022). Of note, the immune escape mechanism of acute myeloid leukemia (but also of other cancer cells) called HLA loss that has been described after alloHSCT with haploidentical relatives but also with MSD, MUD and MMUD will have to be somehow accounted for by future donor selection algorithms (Christopher et al. 2018, Toffalori et al. 2019). Notably, the role of individual HLA heterogeneity in the context of cellular immune evasion in cancer has been discussed (Pagliuca et al. 2022). These different data also show that models initially developed to answer specific needs and questions like the TCE, PIRCHE or the B leader can be extended to other clinical settings and are informative too. Another example is the importance of the immunopeptidome at determining T-cell alloreactivity against HLA-DPB1 mismatches that has just been translated to HLA class I mismatched allografts and shown to be clinically relevant (Crivello et al. 2023).

Regarding the B leader matching model (Petersdorf et al. 2020, Petersdorf et al. 2020), it is expected to influence the alloreactivity of another important immune cell subset, namely the NK cells. As briefly introduced in the first chapter of this thesis, the NK alloimmune response in such a constellation could be driven through HLA-E:CD94/NKG2A interactions or lack thereof (Rolle et al. 2018). Indeed, NK cells are capable of detecting the absence of HLA class I molecules on the surface of target cells (i.e., missing self) or alterations in the peptidome presented by HLA class I molecules (i.e., altered self) via a large set of germline–encoded activating and inhibitory receptors (Hilton and Parham 2017, Carrillo-Bustamante et al. 2018). Two schools of NK cells education have been proposed (Horowitz et al. 2016), one implicating

HLA-E and CD94/NKGDA and the other involving KIRs and their HLA class I ligands (Hilton and Parham 2017, Parham and Guethlein 2018). The different molecular mechanisms underlying NK cell alloreactivity are complex and not fully understood at the moment (Hamada et al. 2021). In alloHSCT, this concept has predominantly been observed with haploidentical grafts (Ruggeri et al. 2002, Zhao et al. 2014, Locatelli et al. 2018, Pende et al. 2019, Shimoni et al. 2019), but also in transplantation with other types of donors (Cooley et al. 2009, Cooley et al. 2010). HLA-E and NKG2A, which clusters in the Natural Killer Complex on chromosome 12, are highly conserved and present a system for the "gross" detection of HLA class I cell surface expression (Carrillo-Bustamante et al. 2016). By contrast, the Leukocyte Receptor Complex containing KIR genes on chromosome 19 is highly diverse in terms of gene content, copy number variation and allelic polymorphism (Parham and Guethlein 2018). Until recently, the incapacity to fully explore KIR genetic diversity with the available genotyping methods has limited the predictive value of KIRs for alloHSCT. However, thanks to high throughput sequencing protocols and bioinformatics tools specifically developed to characterize this extremely complex genetic system (Norman et al. 2016, Beziat et al. 2017, Wagner et al. 2018, Marin et al. 2021, Downing and D'Orsogna 2022), mounting evidence indicates the contribution of KIR allelic functional diversity in the clinical course after transplantation (Venstrom et al. 2012, Bari et al. 2013, Boudreau et al. 2017, Guethlein et al. 2021). Yet, a consensual view is still missing (Schetelig et al. 2020). Furthermore, KIRs are stochastically expressed at the cell surface with thousands of different NK cell subsets circulating in the peripheral blood of any given individual (Horowitz et al. 2013). This means that besides KIR allelic genotyping, the determination of KIR phenotypes by flow cytometry or CyTOF should lead to a better understanding of NK cell alloreactivity (Horowitz et al. 2015, Romee et al. 2016, Li et al. 2019). While not the direct focus of this thesis, our group has been interested by NK cell biology since several years (Hadaya et al. 2008, Buhler et al. 2009, de Rham et al. 2014, Buhler et al. 2016, de Rham et al. 2020). We are currently investigating the putative combinatorial effects of different NK cell receptors and their HLA ligands defined at high resolution in the Swiss alloHSCT cohort. Indeed, integrating other genetically diverse systems such as KIR as pre-transplant markers in addition to HLA in the process of donor selection could help improve patient's global outcome following alloHSCT.

To further illustrate the attention gathered by this topic, a study involving another receptor expressed on NK cells has just been published (Petersdorf et al. 2023). It shows the clinical relevance of polymorphisms in both the NKG2D receptor, which sense "kill me" signals, and

its ligands, the stress-induced MHC class I chain related molecules (encoded by MICA and MICB) that are expressed by many tumor cells.

With the increasing number of HLA mismatched allografts that are performed annually, another important aspect that has entered guidelines for selecting donors and the routine of histocompatibility laboratories is to test for the presence of donor specific alloantibodies (Dehn et al. 2019, Little et al. 2021). Their presence has notably been linked to primary graft failure as reviewed by Little et al. (2021, see also the references therein). B cell recognition of non-self HLA molecules and the production of alloantibodies is a major determinant in solid organ transplantation (Tiercy and Claas 2013, Reindl-Schwaighofer et al. 2020, Callemeyn et al. 2022). This is outside the scope of this thesis. However, we think it is important to mention that in a similar and parallel fashion to what has led to the characterization of HLA mismatches and T-cell epitopes in alloHSCT, a lot of efforts have also been invested to analyze B-cell epitopes on HLA molecules and the production of alloantibodies, including the development of predictive algorithms such as PIRCHE already discussed here or HLAMatchmaker (Tambur and Claas 2015, Filippone and Farber 2016, Tambur 2018).

The last important aspects that we want to address in these conclusions and perspectives concern the T-cell repertoire and its reconstitution at the post-transplant stage. As we and other have shown, the alloreactive T-cell repertoire is broad and hardly predictable (Emerson et al. 2014, Arrieta-Bolanos et al. 2018, DeWolf et al. 2018, Bettens et al. 2020). For instance, the tissue distribution and frequency of alloreactive cytotoxic T-cell clones is clearly heterogeneous in patients with GVHD (Koyama et al. 2019). These results can be explained by the sheer number of potential responder T cells that are circulating in any given individual (DeWolf et al. 2018) and also because the TCR is highly flexible and intrinsically crossreactive at recognizing T-cell epitopes (Yin and Mariuzza 2009, Wang et al. 2017). In addition, parameters such as the fitness of clones, cytokine environment and inflammatory events might influence the shape and size of the alloimmune response. Interestingly, very recent data suggest that resident progenitor-like T cells are the main culprit of sustained GVHD in target tissues (Sacirbegovic et al. 2023). In the context of direct T-cell allorecognition which potentially involves up to 10% of T cells both from the memory and naïve compartments, two theories are proposed to explain the strength and frequency of precursor clones eliciting the alloimmune response (Boardman et al. 2016). In the HLA-centric model (or high determinant density hypothesis), some structural differences in the peptide binding region of alloHLA are

recognized as foreign by the docking TCRs, irrespective of the bound peptide. This is facilitated by the high density of alloHLA molecules expressed at the cell surface of APC. In the peptide-centric model (or multiple binary complexes hypothesis), alloHLA are able to present a broad pool of foreign peptides to the TCRs, thus eliciting the activation of a large repertoire of T cells. Both models are supported by functional and structural data and the actual view is that they combine and cooperate *in vivo* (Boardman et al. 2016, Wang et al. 2017). T-cell alloreactivity can also be mediated by the indirect recognition of alloHLA derived into allogeneic peptides presented by self/compatible HLA molecules (Geneugelijk et al. 2014, Marino et al. 2016). This pathway is thought to occur mainly at later stages of alloreactivity in comparison with direct T-cell recognition (Ali et al. 2016) and resembles the usual response of T cells to pathogens, but also to MiHAs. It is elicited by self-HLA-restricted peptide-specific T cells with low precursor frequency (<0.1%) and coming mostly from the naïve repertoire (Fleischhauer and Shaw 2017).

Taken altogether, these data explain why the alloimmune response is characterized by unusual features compared to typical immune responses to pathogens. They also explain why a comprehensive analysis of the alloreactive repertoire and a possible integration to the standard of care and immune monitoring of patients after alloHSCT is very challenging (Fu et al. 2021, Tian et al. 2022). However, this is also a field of investigation that generates a lot of interest and is brimming with potentially breakthrough progresses. For instance, identification followed by longitudinal tracking and monitoring of alloreactive clonotypes (e.g., by combining MLRs and immunosequencing) has shown promise as a biomarker of tolerance and rejection as reviewed in Fu et al (2021) and Tian et al (2022). This could promote the establishment of personalized and tailored immunosuppressive protocols or instruct on the best approaches to induce tolerance. Recent bioinformatics developments and the availability of published TCR databases (Tickotsky et al. 2017, Shugay et al. 2018) should also contribute to a better understanding of the alloimmune response. In addition, the role of Treg in tempering alloimmunity is a complementary axis of investigation that could provide new insights and already offers alternative therapeutic options (Amini et al. 2023). We recently collaborated with the research group of the service of hematology at HUG in a study using a mice model of GVHD combined with immunosequencing and transcriptomic analyses (Lohmeyer et al. 2022). Our group is also currently working on other functional approaches to decipher T-cell alloreactivity using a combination of immunosequencing, MLRs and *in vitro* assays at the single cell level (Li et al. 2018, Cachot et al. 2021, Liu et al. 2022).

How the TCR repertoire is established is another challenging topic of research in humans. In this context, alloHSCT represents a unique model and opportunity for investigating the construction of the repertoire, although with the biases inherent to this therapy such as the conditioning regimen, immunosuppressive environment, post-transplant clinical interventions and disparities in the age of patients and donors. As shown by us and others, several clinical parameters are significantly associated with the diversity of the TCR repertoire after transplantation (see more details in our annexed article and also in the introduction). Notably, cytomegalovirus reactivation/infection can profoundly and durably affect the shape and diversity of the repertoire after alloHSCT, but also in otherwise healthy individuals (Hakki et al. 2003, Sylwester et al. 2005, Suessmuth et al. 2015, Buhler et al. 2020, Calderin Sollet et al. 2023). What is less clear is whether the reconstituting repertoire is predictive of subsequent clinical events, the timeline to which the repertoire acquire its main features (i.e., polyclonal versus oligoclonal profiles) and if it remains stable over time.

Interestingly, in our cohort, the diversity of the TCR repertoire at one year was not associated with the occurrence of clinical events after that time point, despite the significant differences of clonality observed among recipients (Buhler et al. 2020). We just published the follow-up at 5 to 6 years in a subset of patients from this cohort and could show that the diversity of the repertoire as established at one year didn't change much afterward, although we observed some influence of cytomegalovirus. This was further supported by the larger overlap of sequenced clonotypes found between both post-transplant time points by contrast to almost no overlap with the pre-transplant repertoire in donors (Calderin Sollet et al. 2023).

Longitudinal studies have shed some light on the dynamics of the reconstitution and architecture of the repertoire (Pagliuca et al. 2021). Novel factors have also been proposed in order to dissect the composition of the repertoire more precisely such as the notion of private and public clonotypes (Soto et al. 2020, Trofimov et al. 2022) or the so-called neonatal and terminal deoxynucleotidyl transferase (TDT)-dependent T cells (de Greef et al. 2020, Trofimov et al. 2022). Prenatally produced TCRs are characterized by shorter sequences, no N-insertions, a high probability of V(D)J recombination in comparison to TDT-dependent TCRs (Trofimov et al. 2022) and are prevalent in the naïve repertoire (de Greef et al. 2020). Neonatal T-cell clones have been defined by their publicness (i.e., they are highly shared among individuals) and are strongly polyreactive to self and non-self, which could favor their maintenance and replenishment through lifetime. By contrast, the role of TDT-dependent TCRs remains unclear but they could endow each individual with a more private layer of antigen recognition and thus

contribute to the diversity and functionality of the repertoire. They are considered more prone to be clonally deleted by negative selection in the thymus (Trofimov et al. 2022). Although challenged by the very large space of T-cell epitopes and the yet limited availability of TCRs of known specificity, bioinformatics developments offer the possibility to investigate the composition of the repertoire and its correlation with documented infectious exposures, occurrence of GVHD or relapse of cancer (Meysman et al. 2019, Vujovic et al. 2020, Zhang et al. 2021, Jokinen et al. 2023). In addition, the impact of HLA diversity on the shape of the repertoire has been assessed both in cohorts of healthy adults (DeWitt et al. 2018, Krishna et al. 2020, Johnson et al. 2021) and in the context of alloHSCT (Pagliuca et al. 2021). A synthetic approach accounting for the various characteristics of the TCR should allow to disentangle and have a better understanding of the relative contributions of random biological processes like somatic recombination and homeostatic expansions, from antigen-driven contractions of the repertoire following the enrichment of alloreactive or pathogen-specific T cells. On the clinical side, this should improve the reliability and usefulness of post-transplant TCR repertoire monitoring by immunosequencing, in conjunction with better insights on T-cell alloreactivity in the context of GVHD as described above and reviewed lately (Goel et al. 2022).

To conclude this thesis, we have presented some of the latest advances in the field of alloHSCT with a special focus on two important aspects, namely the optimization of histocompatibility between donors and patients and the immune reconstitution of T cells following engraftment. The knowledge and valuable amount of data that have accumulated in recent years and still to be produced in the close future will help clinical decisions for the best donor selection criteria in pre-HSCT situations and should pave the way for novel optimized and personalized post-transplant monitoring approaches with tailored clinical interventions whenever necessary.

References

(1999). "Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium." Nature **401**(6756): 921-923.

Al Malki, M. M., N. C. Tsai, J. Palmer, S. Mokhtari, W. Tsai, T. Cao, . . . R. Nakamura (2021). "Posttransplant cyclophosphamide as GVHD prophylaxis for peripheral blood stem cell HLA-mismatched unrelated donor transplant." Blood Adv **5**(12): 2650-2659.

Ali, J. M., M. C. Negus, T. M. Conlon, I. G. Harper, M. S. Qureshi, R. Motallebzadeh, . . . G. J. Pettigrew (2016). "Diversity of the CD4 T Cell Alloresponse: The Short and the Long of It." Cell Rep **14**(5): 1232-1245.

Allen, S., S. J. Turner, D. Bourges, P. A. Gleeson and I. R. van Driel (2011). "Shaping the T-cell repertoire in the periphery." Immunol Cell Biol **89**(1): 60-69.

Amini, L., J. Kaeda, E. Fritsche, A. Roemhild, D. Kaiser and P. Reinke (2023). "Clinical adoptive regulatory T Cell therapy: State of the art, challenges, and prospective." Frontiers in Cell and Developmental Biology **10**.

Anasetti, C., D. Amos, P. G. Beatty, F. R. Appelbaum, W. Bensinger, C. D. Buckner, . . . et al. (1989). "Effect of HLA compatibility on engraftment of bone marrow transplants in patients with leukemia or lymphoma." N Engl J Med **320**(4): 197-204.

Appelbaum, F. R. (2007). "Hematopoietic-cell transplantation at 50." N Engl J Med **357**(15): 1472-1475.

Arrieta-Bolanos, E., P. Crivello, M. He, T. Wang, S. M. Gadalla, S. Paczesny, . . . K. Fleischhauer (2022). "A core group of structurally similar HLA-DPB1 alleles drives permissiveness after hematopoietic cell transplantation." Blood.

Arrieta-Bolanos, E., P. Crivello, M. Metzing, T. Meurer, M. Ahci, J. Rytlewski, . . . K. Fleischhauer (2018). "Alloreactive T Cell Receptor Diversity against Structurally Similar or Dissimilar HLA-DP Antigens Assessed by Deep Sequencing." Front Immunol **9**: 280.

Attaf, M., M. Legut, D. K. Cole and A. K. Sewell (2015). "The T cell antigen receptor: the Swiss army knife of the immune system." Clin Exp Immunol **181**(1): 1-18.

Ayuk, F., M. Bornhauser, M. Stelljes, T. Zabelina, E. M. Wagner, C. Schmid, . . . W. Bethge (2019). "Predicted Indirectly ReCognizable HLA Epitopes (PIRCHE) Are Associated with Poorer Outcome after Single Mismatch Unrelated Donor Stem Cell Transplantation: A Study of the Cooperative Transplant Study Group (KTS) of the German Group for Bone Marrow and Stem Cell Transplantation (DAG-KBT)." Transfus Med Hemother **46**(5): 370-375.

Bacigalupo, A. (2021). "HLA still matters in allogeneic transplants." Blood **138**(3): 209-211.

Bari, R., P. Rujkijyanont, E. Sullivan, G. Kang, V. Turner, K. Gan and W. Leung (2013). "Effect of donor KIR2DL1 allelic polymorphism on the outcome of pediatric allogeneic hematopoietic stem-cell transplantation." J Clin Oncol **31**(30): 3782-3790.

Baron, F., B. Storer, M. B. Maris, J. Storek, F. Piette, M. Metcalf, . . . M. Boeckh (2006). "Unrelated donor status and high donor age independently affect immunologic recovery after nonmyeloablative conditioning." Biol Blood Marrow Transplant **12**(11): 1176-1187.

Barone, J. C., K. Saito, K. Beutner, M. Campo, W. Dong, C. P. Goswami, . . . S. Hsu (2015). "HLA-genotyping of clinical specimens using Ion Torrent-based NGS." Hum Immunol **76**(12): 903-909.

Battipaglia, G., J. E. Galimard, M. Labopin, A. M. Raiola, D. Blaise, A. Ruggeri, . . . M. Mohty (2022). "Post-transplant cyclophosphamide in one-antigen mismatched unrelated donor transplantation versus haploidentical transplantation in acute myeloid leukemia: a study from the Acute Leukemia Working Party of the EBMT." Bone Marrow Transplant **57**(4): 562-571.

Baumeister, S. H. C., B. Rambaldi, R. M. Shapiro and R. Romee (2020). "Key Aspects of the Immunobiology of Haploidentical Hematopoietic Cell Transplantation." Frontiers in Immunology **11**.

Baxter-Lowe, L. A. (2021). "The changing landscape of HLA typing: Understanding how and when HLA typing data can be used with confidence from bench to bedside." Hum Immunol **82**(7): 466-477.

Benichou, G., B. Gonzalez, J. Marino, K. Ayasoufi and A. Valujskikh (2017). "Role of Memory T Cells in Allograft Rejection and Tolerance." Front Immunol **8**: 170.

Berard, M. and D. F. Tough (2002). "Qualitative differences between naive and memory T cells." Immunology **106**(2): 127-138.

Bertaina, A. and M. Andreani (2018). "Major Histocompatibility Complex and Hematopoietic Stem Cell Transplantation: Beyond the Classical HLA Polymorphism." Int J Mol Sci **19**(2).

Bettens, F., S. Buhler and J. M. Tiercy (2016). "Allorecognition of HLA-C Mismatches by CD8(+) T Cells in Hematopoietic Stem Cell Transplantation Is a Complex Interplay between Mismatched Peptide-Binding Region Residues, HLA-C Expression, and HLA-DPB1 Disparities." Front Immunol **7**: 584.

Bettens, F., Z. Calderin Sollet, S. Buhler and J. Villard (2020). "CD8+ T-Cell Repertoire in Human Leukocyte Antigen Class I-Mismatched Alloreactive Immune Response." Front Immunol **11**: 588741.

Bettens, F., H. Ongen, G. Rey, S. Buhler, Z. Calderin Sollet, E. Dermitzakis and J. Villard (2022). "Regulation of HLA class I expression by non-coding gene variations." PLoS Genet **18**(6): e1010212.

Bettens, F., J. Passweg, U. Schanz, Y. Chalandon, D. Heim, T. Gungor, . . . J. M. Tiercy (2012). "Impact of HLA-DPB1 haplotypes on outcome of 10/10 matched unrelated hematopoietic stem cell donor transplants depends on MHC-linked microsatellite polymorphisms." Biol Blood Marrow Transplant **18**(4): 608-616.

Bettens, F., U. Schanz and J. M. Tiercy (2013). "Lack of recognition of HLA class I mismatches outside alpha1/alpha2 domains by CD8+ alloreactive T lymphocytes: the HLA-B44 paradigm." Tissue Antigens **81**(6): 414-418.

Beziat, V., H. G. Hilton, P. J. Norman and J. A. Traherne (2017). "Deciphering the killer-cell immunoglobulin-like receptor system at super-resolution for natural killer and T-cell biology." Immunology **150**(3): 248-264.

Birnbaum, M. E., J. L. Mendoza, D. K. Sethi, S. Dong, J. Glanville, J. Dobbins, . . . K. C. Garcia (2014). "Deconstructing the peptide-MHC specificity of T cell recognition." Cell **157**(5): 1073-1087.

Bjorkman, P. J. (2006). "Finding the groove." Nat Immunol **7**(8): 787-789.

Blume, K. G. and E. D. Thomas (2000). "A review of autologous hematopoietic cell transplantation." Biol Blood Marrow Transplant **6**(1): 1-12.

Boardman, D. A., J. Jacob, L. A. Smyth, G. Lombardi and R. I. Lechler (2016). "What Is Direct Allorecognition?" Curr Transplant Rep **3**(4): 275-283.

Booth, G. S., E. A. Gehrie, C. D. Bolan and B. N. Savani (2013). "Clinical guide to ABO-incompatible allogeneic stem cell transplantation." Biol Blood Marrow Transplant **19**(8): 1152-1158.

Bosch, M., F. M. Khan and J. Storek (2012). "Immune reconstitution after hematopoietic cell transplantation." Curr Opin Hematol **19**(4): 324-335.

Boudreau, J. E., F. Giglio, T. A. Gooley, P. A. Stevenson, J. B. Le Luduec, B. C. Shaffer, . . . K. C. Hsu (2017). "KIR3DL1/HLA-B Subtypes Govern Acute Myelogenous Leukemia Relapse After Hematopoietic Cell Transplantation." J Clin Oncol **35**(20): 2268-2278.

Bravo-Egana, V., H. Sanders and N. Chitnis (2021). "New challenges, new opportunities: Next generation sequencing and its place in the advancement of HLA typing." Hum Immunol **82**(7): 478-487.

Brown, C. E. and C. L. Mackall (2019). "CAR T cell therapy: inroads to response and resistance." Nat Rev Immunol.

Bugawan, T. L., W. Klitz, A. Blair and H. A. Erlich (2000). "High-resolution HLA class I typing in the CEPH families: analysis of linkage disequilibrium among HLA loci." Tissue Antigens **56**(5): 392-404.

Buhler, S., H. Baldomero, S. Ferrari-Lacraz, A. C. Mamez, S. Masouridi-Levrat, D. Heim, . . . G. Swiss Blood Stem Cell Transplantation (2021). "Analysis of biological models to predict clinical outcomes based on HLA-DPB1 disparities in unrelated transplantation." Blood Adv **5**(17): 3377-3386.

Buhler, S., H. Baldomero, S. Ferrari-Lacraz, J. M. Nunes, A. Sanchez-Mazas, S. Massouridi-Levrat, . . . J. Villard (2019). "High-resolution HLA phased haplotype frequencies to predict the success of unrelated donor searches and clinical outcome following hematopoietic stem cell transplantation." Bone Marrow Transplant **54**(10): 1701-1709.

Buhler, S., F. Bettens, C. Dantin, S. Ferrari-Lacraz, M. Ansari, A. C. Mamez, . . . J. Villard (2020). "Genetic T-cell receptor diversity at 1 year following allogeneic hematopoietic stem cell transplantation." Leukemia **34**(5): 1422-1432.

Buhler, S., J. Di Cristofaro, C. Frassati, A. Basire, V. Galicher, J. Chiaroni and C. Picard (2009). "High levels of molecular polymorphism at the KIR2DL4 locus in French and Congolese populations: impact for anthropology and clinical studies." Hum Immunol **70**(11): 953-959.

Buhler, S., E. Giostra, C. Gbame, C. de Rham, B. Mullhaupt, J. F. Dufour, . . . S. Swiss Transplant Cohort (2016). "A significant effect of the killer cell immunoglobulin-like receptor ligand human leucocyte antigen-C on fibrosis progression in chronic C hepatitis with or without liver transplantation." Liver Int **36**(9): 1331-1339.

Buhler, S. and A. Sanchez-Mazas (2011). "HLA DNA sequence variation among human populations: molecular signatures of demographic and selective events." PLoS One **6**(2): e14643.

Cachot, A., M. Bilous, Y. C. Liu, X. Li, M. Saillard, M. Cenerenti, . . . C. Jandus (2021). "Tumor-specific cytolytic CD4 T cells mediate immunity against human cancer." Sci Adv **7**(9).

Calderin Sollet, Z., A. Schäfer, S. Ferrari-Lacraz, S. Masouridi-Levrat, A.-C. Mamez, A. Pradier, . . . S. Buhler (2023). "CMV serostatus and T-cell repertoire diversity 5 years after allogeneic hematopoietic stem cell transplantation." Leukemia.

Callemeyn, J., B. Lamarthee, A. Koenig, P. Koshy, O. Thaunat and M. Naesens (2022). "Allorecognition and the spectrum of kidney transplant rejection." Kidney Int **101**(4): 692-710.

Carapito, R., N. Jung, M. Kwemou, M. Untrau, S. Michel, A. Pichot, . . . S. Bahram (2016). "Matching for the nonconventional MHC-I MICA gene significantly reduces the incidence of acute and chronic GVHD." Blood **128**(15): 1979-1986.

Carrillo-Bustamante, P., R. J. de Boer and C. Kesmir (2018). "Specificity of inhibitory KIRs enables NK cells to detect changes in an altered peptide environment." Immunogenetics **70**(2): 87-97.

Carrillo-Bustamante, P., C. Kesmir and R. J. de Boer (2016). "The evolution of natural killer cell receptors." Immunogenetics **68**(1): 3-18.

Chabannon, C., J. Kuball, A. Bondanza, F. Dazzi, P. Pedrazzoli, A. Toubert, . . . C. Bonini (2018). "Hematopoietic stem cell transplantation in its 60s: A platform for cellular therapies." Sci Transl Med **10**(436).

Chaudhry, M. S., E. Velardi, F. Malard and M. R. van den Brink (2017). "Immune Reconstitution after Allogeneic Hematopoietic Stem Cell Transplantation: Time To T Up the Thymus." J Immunol **198**(1): 40-46.

Christopher, M. J., A. A. Petti, M. P. Rettig, C. A. Miller, E. Chendamarai, E. J. Duncavage, . . . J. F. DiPersio (2018). "Immune Escape of Relapsed AML Cells after Allogeneic Transplantation." N Engl J Med **379**(24): 2330-2341.

Ciurea, S. O., M. M. Al Malki, P. Kongtim, E. J. Fuchs, L. Luznik, X. J. Huang, . . . A. Nagler (2020). "The European Society for Blood and Marrow Transplantation (EBMT) consensus recommendations for donor selection in haploidentical hematopoietic cell transplantation." Bone Marrow Transplant **55**(1): 12-24.

Ciurea, S. O. and U. D. Bayraktar (2015). ""No donor"? Consider a haploidentical transplant." Blood Rev **29**(2): 63-70.

Ciurea, S. O., M. J. Zhang, A. A. Bacigalupo, A. Bashey, F. R. Appelbaum, O. S. Aljitawi, . . . M. Eapen (2015). "Haploidentical transplant with posttransplant cyclophosphamide vs matched unrelated donor transplant for acute myeloid leukemia." Blood **126**(8): 1033-1040.

Cooley, S., E. Trachtenberg, T. L. Bergemann, K. Saeteurn, J. Klein, C. T. Le, . . . D. J. Weisdorf (2009). "Donors with group B KIR haplotypes improve relapse-free survival after unrelated hematopoietic cell transplantation for acute myelogenous leukemia." Blood **113**(3): 726-732.

Cooley, S., D. J. Weisdorf, L. A. Guethlein, J. P. Klein, T. Wang, C. T. Le, . . . J. S. Miller (2010). "Donor selection for natural killer cell receptor genes leads to superior survival after unrelated transplantation for acute myelogenous leukemia." Blood **116**(14): 2411-2419.

Cosway, E. J., K. D. James, B. Lucas, G. Anderson and A. J. White (2021). "The thymus medulla and its control of alphabetaT cell development." Semin Immunopathol **43**(1): 15-27.

Crivello, P., E. Arrieta-Bolanos, M. He, T. Wang, S. Fingerson, S. M. Gadalla, . . . K. Fleischhauer (2023). "Impact of the HLA Immunopeptidome on Survival of Leukemia Patients After Unrelated Donor Transplantation." J Clin Oncol: JCO2201229.

Crivello, P., A. Heinold, V. Rebmann, H. D. Ottinger, P. A. Horn, D. W. Beelen and K. Fleischhauer (2016). "Functional distance between recipient and donor HLA-DPB1 determines nonpermissive mismatches in unrelated HCT." Blood **128**(1): 120-129.

Crivello, P., L. Zito, F. Sizzano, E. Zino, M. Maiers, A. Mulder, . . . K. Fleischhauer (2015). "The impact of amino acid variability on alloreactivity defines a functional distance predictive of permissive HLA-DPB1 mismatches in hematopoietic stem cell transplantation." Biol Blood Marrow Transplant **21**(2): 233-241.

Crocchiolo, R., E. Zino, L. Vago, R. Oneto, B. Bruno, S. Pollichieni, . . . R. Italian Bone Marrow Donor (2009). "Nonpermissive HLA-DPB1 disparity is a significant independent risk factor for mortality after unrelated hematopoietic stem cell transplantation." Blood **114**(7): 1437-1444.

Cullen, M., J. Noble, H. Erlich, K. Thorpe, S. Beck, W. Klitz, . . . M. Carrington (1997). "Characterization of recombination in the HLA class II region." Am J Hum Genet **60**(2): 397-407.

D'Orsogna, L., H. van den Heuvel, C. van Kooten, S. Heidt and F. H. J. Claas (2017). "Infectious pathogens may trigger specific allo-HLA reactivity via multiple mechanisms." Immunogenetics **69**(8-9): 631-641.

de Greef, P. C., T. Oakes, B. Gerritsen, M. Ismail, J. M. Heather, R. Hermsen, . . . R. J. de Boer (2020). "The naive T-cell receptor repertoire has an extremely broad distribution of clone sizes." Elife **9**.

de Rham, C., Z. Calderin Sollet, P. Burkhard and J. Villard (2020). "Natural Killer Cell Alloreactivity Against Human Induced Pluripotent Stem Cells and Their Neuronal Derivatives into Dopaminergic Neurons." Stem Cells Dev **29**(13): 853-862.

de Rham, C., K. Hadaya, C. Bandelier, S. Ferrari-Lacraz and J. Villard (2014). "Expression of killer cell immunoglobulin-like receptors (KIRs) by natural killer cells during acute CMV infection after kidney transplantation." Transpl Immunol **31**(3): 157-164.

Degauque, N., S. Brouard and J. P. Soulillou (2016). "Cross-Reactivity of TCR Repertoire: Current Concepts, Challenges, and Implication for Allotransplantation." Front Immunol **7**: 89.

Dehn, J., M. Setterholm, K. Buck, J. Kempenich, B. Beduhn, L. Gragert, . . . M. Maiers (2016). "HapLogic: A Predictive Human Leukocyte Antigen-Matching Algorithm to Enhance Rapid Identification of the Optimal Unrelated Hematopoietic Stem Cell Sources for Transplantation." Biol Blood Marrow Transplant **22**(11): 2038-2046.

Dehn, J., S. Spellman, C. K. Hurley, B. E. Shaw, J. N. Barker, L. J. Burns, . . . J. Pidala (2019). "Selection of unrelated donors and cord blood units for hematopoietic cell transplantation: guidelines from the NMDP/CIBMTR." Blood **134**(12): 924-934.

Dekker, L., C. de Koning, C. Lindemans and S. Nierkens (2020). "Reconstitution of T Cell Subsets Following Allogeneic Hematopoietic Cell Transplantation." Cancers (Basel) **12**(7).

Dekker, L., E. Sanders, C. A. Lindemans, C. de Koning and S. Nierkens (2022). "Naive T Cells in Graft Versus Host Disease and Graft Versus Leukemia: Innocent or Guilty?" Front Immunol **13**: 893545.

Detrait, M., S. Morisset, Y. Chalandon, I. Yakoub-Agha, F. Dufosse, M. Labalette, . . . V. Dubois (2015). "Suggestive evidence of a role of HLA-DRB4 mismatches in the outcome of allogeneic hematopoietic stem cell transplantation with HLA-10/10-matched unrelated donors: a French-Swiss retrospective study." Bone Marrow Transplant **50**(10): 1316-1320.

DeWitt, W. S., 3rd, A. Smith, G. Schoch, J. A. Hansen, F. A. t. Matsen and P. Bradley (2018). "Human T cell receptor occurrence patterns encode immune history, genetic background, and receptor specificity." Elife **7**.

DeWolf, S., B. Grinshpun, T. Savage, S. P. Lau, A. Obradovic, B. Shonts, . . . Y. Shen (2018). "Quantifying size and diversity of the human T cell alloresponse." JCI Insight **3**(15).

DeWolf, S. and M. Sykes (2017). "Alloimmune T cells in transplantation." J Clin Invest **127**(7): 2473-2481.

Dickinson, A. M. (2008). "Non-HLA genetics and predicting outcome in HSCT." Int J Immunogenet **35**(4-5): 375-380.

Downing, J. and L. D'Orsogna (2022). "High-resolution human KIR genotyping." Immunogenetics **74**(4): 369-379.

Embgenbroich, M. and S. Burgdorf (2018). "Current Concepts of Antigen Cross-Presentation." Front Immunol **9**: 1643.

Emerson, R. O., J. M. Mathew, I. M. Konieczna, H. S. Robins and J. R. Leventhal (2014). "Defining the alloreactive T cell repertoire using high-throughput sequencing of mixed lymphocyte reaction culture." PLoS One **9**(11): e111943.

Felix, N. J. and P. M. Allen (2007). "Specificity of T-cell alloreactivity." Nat Rev Immunol **7**(12): 942-953.

Fernandez-Vina, M. A., J. P. Klein, M. Haagenson, S. R. Spellman, C. Anasetti, H. Noreen, . . . M. de Lima (2013). "Multiple mismatches at the low expression HLA loci DP, DQ, and DRB3/4/5 associate with adverse outcomes in hematopoietic stem cell transplantation." Blood **121**(22): 4603-4610.

Fernandez-Vina, M. A., T. Wang, S. J. Lee, M. Haagenson, M. Aljurf, M. Askar, . . . M. Setterholm (2014). "Identification of a permissible HLA mismatch in hematopoietic stem cell transplantation." Blood **123**(8): 1270-1278.

Filippone, E. J. and J. L. Farber (2016). "The Humoral Theory of Transplantation: Epitope Analysis and the Pathogenicity of HLA Antibodies." J Immunol Res **2016**: 5197396.

Fleischhauer, K. (2015). "Immunogenetics of HLA-DP--A New View of Permissible Mismatches." N Engl J Med **373**(7): 669-672.

Fleischhauer, K. (2022). "Haplo-PtCy: adjusting the HLA barrier." Blood **139**(10): 1431-1433.

Fleischhauer, K. and D. W. Beelen (2016). "HLA mismatching as a strategy to reduce relapse after alternative donor transplantation." Semin Hematol **53**(2): 57-64.

Fleischhauer, K. and B. E. Shaw (2017). "HLA-DP in unrelated hematopoietic cell transplantation revisited: challenges and opportunities." Blood **130**(9): 1089-1096.

Fleischhauer, K., B. E. Shaw, T. Gooley, M. Malkki, P. Bardy, J. D. Bignon, . . . T. International Histocompatibility Working Group in Hematopoietic Cell (2012). "Effect of T-cell-epitope matching at HLA-DPB1 in recipients of unrelated-donor haemopoietic-cell transplantation: a retrospective study." Lancet Oncol **13**(4): 366-374.

Fu, J., M. Khosravi-Maharlooei and M. Sykes (2021). "High Throughput Human T Cell Receptor Sequencing: A New Window Into Repertoire Establishment and Alloreactivity." Frontiers in Immunology **12**.

Fuchs, E. J. (2017). "Related haploidentical donors are a better choice than matched unrelated donors: Point." Blood Adv **1**(6): 397-400.

Fuchs, E. J., S. R. McCurdy, S. R. Solomon, T. Wang, M. R. Herr, D. Modi, . . . E. W. Petersdorf (2022). "HLA informs risk predictions after haploidentical stem cell transplantation with posttransplantation cyclophosphamide." Blood **139**(10): 1452-1468.

Geneugelijk, K., K. A. Thus and E. Spierings (2014). "Predicting alloreactivity in transplantation." J Immunol Res **2014**: 159479.

Geneugelijk, K., K. A. Thus, H. W. M. van Deutekom, J. J. A. Calis, E. Borst, C. Kesmir, . . . E. Spierings (2019). "Exploratory Study of Predicted Indirectly ReCognizable HLA Epitopes in Mismatched Hematopoietic Cell Transplantations." Front Immunol **10**: 880.

Genomes Project, C., A. Auton, L. D. Brooks, R. M. Durbin, E. P. Garrison, H. M. Kang, . . . G. R. Abecasis (2015). "A global reference for human genetic variation." Nature **526**(7571): 68-74.

Gfeller, D. and M. Bassani-Sternberg (2018). "Predicting Antigen Presentation-What Could We Learn From a Million Peptides?" Front Immunol **9**: 1716.

Giralt, S. and M. R. Bishop (2009). "Principles and overview of allogeneic hematopoietic stem cell transplantation." Cancer Treat Res **144**: 1-21.

Gluckman, E. and V. Rocha (2009). "Cord blood transplantation: state of the art." Haematologica **94**(4): 451-454.

Goel, M., A. Eugster, J. Schetelig, E. Bonifacio, M. Bornhauser and C. S. Link-Rachner (2022). "Potential of TCR sequencing in graft-versus-host disease." Bone Marrow Transplant.

Gooptu, M., R. Romee, A. St Martin, M. Arora, M. Al Malki, J. H. Antin, . . . M. Eapen (2021). "HLA-haploidentical vs matched unrelated donor transplants with posttransplant cyclophosphamide-based prophylaxis." Blood **138**(3): 273-282.

Griffioen, M., C. A. van Bergen and J. H. Falkenburg (2016). "Autosomal Minor Histocompatibility Antigens: How Genetic Variants Create Diversity in Immune Targets." Front Immunol **7**: 100.

Grubic, Z., M. Burek Kamenaric, M. Maskalan, N. Durakovic, R. Vrhovac, K. Stingl Jankovic, . . . R. Zunec (2022). "Various approaches for accessing the influence of human leukocyte antigens disparity in haploidentical stem cell transplantation." Int J Lab Hematol **44**(3): 547-557.

Guethlein, L. A., N. Beyzaie, N. Nemat-Gorgani, T. Wang, V. Ramesh, W. M. Marin, . . . P. Parham (2021). "Following Transplantation for Acute Myelogenous Leukemia, Donor KIR Cen B02 Better Protects against Relapse than KIR Cen B01." J Immunol **206**(12): 3064-3072.

Gyurkocza, B. and B. M. Sandmaier (2014). "Conditioning regimens for hematopoietic cell transplantation: one size does not fit all." Blood **124**(3): 344-353.

Hadaya, K., C. de Rham, C. Bandelier, C. Bandelier, S. Ferrari-Lacraz, S. Jendly, . . . J. Villard (2008). "Natural killer cell receptor repertoire and their ligands, and the risk of CMV infection after kidney transplantation." Am J Transplant **8**(12): 2674-2683.

Hakim, F. T., S. A. Memon, R. Cepeda, E. C. Jones, C. K. Chow, C. Kasten-Sportes, . . . R. E. Gress (2005). "Age-dependent incidence, time course, and consequences of thymic renewal in adults." J Clin Invest **115**(4): 930-939.

Hakki, M., S. R. Riddell, J. Storek, R. A. Carter, T. Stevens-Ayers, P. Sudour, . . . M. Boeckh (2003). "Immune reconstitution to cytomegalovirus after allogeneic hematopoietic stem cell transplantation: impact of host factors, drug therapy, and subclinical reactivation." Blood **102**(8): 3060-3067.

Hamada, S., V. Dubois, A. Koenig and O. Thaunat (2021). "Allograft recognition by recipient's natural killer cells: Molecular mechanisms and role in transplant rejection." HLA **98**(3): 191-199.

Heather, J. M., M. Ismail, T. Oakes and B. Chain (2018). "High-throughput sequencing of the T-cell receptor repertoire: pitfalls and opportunities." Brief Bioinform **19**(4): 554-565.

Hilton, H. G. and P. Parham (2017). "Missing or altered self: human NK cell receptors that recognize HLA-C." Immunogenetics **69**(8-9): 567-579.

Horowitz, A., Z. Djaoud, N. Nemat-Gorgani, J. Blokhuis, H. G. Hilton, V. Beziat, . . . P. Parham (2016). "Class I HLA haplotypes form two schools that educate NK cells in different ways." Sci Immunol **1**(3).

Horowitz, A., L. A. Guethlein, N. Nemat-Gorgani, P. J. Norman, S. Cooley, J. S. Miller and P. Parham (2015). "Regulation of Adaptive NK Cells and CD8 T Cells by HLA-C Correlates with Allogeneic Hematopoietic Cell Transplantation and with Cytomegalovirus Reactivation." J Immunol **195**(9): 4524-4536.

Horowitz, A., D. M. Strauss-Albee, M. Leipold, J. Kubo, N. Nemat-Gorgani, O. C. Dogan, . . . C. A. Blish (2013). "Genetic and environmental determinants of human NK cell diversity revealed by mass cytometry." Sci Transl Med **5**(208): 208ra145.

Horton, R., L. Wilming, V. Rand, R. C. Lovering, E. A. Bruford, V. K. Khodiyar, . . . S. Beck (2004). "Gene map of the extended human MHC." Nat Rev Genet **5**(12): 889-899.

Hurley, C. K. (2021). "Naming HLA diversity: A review of HLA nomenclature." Hum Immunol **82**(7): 457-465.

Hurley, C. K., J. Kempenich, K. Wadsworth, J. Sauter, J. A. Hofmann, D. Schefzyk, . . . J. Dehn (2020). "Common, intermediate and well-documented HLA alleles in world populations: CIWD version 3.0.0." HLA **95**(6): 516-531.

Inman, C. F., S. A. Eldershaw, J. E. Croudace, N. J. Davies, A. Sharma-Oates, T. Rai, . . . P. Moss (2018). "Unique features and clinical importance of acute alloreactive immune responses." JCI Insight **3**(10).

Ishikawa, Y., K. Kashiwase, T. Akaza, Y. Morishima, H. Inoko, T. Sasazuki, . . . T. Juji (2002). "Polymorphisms in TNFA and TNFR2 affect outcome of unrelated bone marrow transplantation." Bone Marrow Transplant **29**(7): 569-575.

Jameson, S. C. (2002). "Maintaining the norm: T-cell homeostasis." Nat Rev Immunol **2**(8): 547-556.

Jenq, R. R. and M. R. van den Brink (2010). "Allogeneic haematopoietic stem cell transplantation: individualized stem cell and immune therapy of cancer." Nat Rev Cancer **10**(3): 213-221.

Johansson, T., J. Partanen and P. Saavalainen (2022). "HLA allele-specific expression: Methods, disease associations, and relevance in hematopoietic stem cell transplantation." Front Immunol **13**: 1007425.

Johnson, S. A., S. L. Seale, R. M. Gittelman, J. A. Rytlewski, H. S. Robins and P. A. Fields (2021). "Impact of HLA type, age and chronic viral infection on peripheral T-cell receptor sharing between unrelated individuals." PLoS One **16**(8): e0249484.

Jokinen, E., A. Dumitrescu, J. Huuhtanen, V. Gligorijevic, S. Mustjoki, R. Bonneau, . . . H. Lahdesmaki (2023). "TCRconv: predicting recognition between T cell receptors and epitopes using contextualized motifs." Bioinformatics **39**(1).

Joris, M. M., A. C. Lankester, P. A. von dem Borne, J. Kuball, M. Bierings, J. J. Cornelissen, . . . M. Oudshoorn (2013). "The impact of frequent HLA haplotypes in high linkage disequilibrium on donor search and clinical outcome after unrelated haematopoietic SCT." Bone Marrow Transplant **48**(4): 483-490.

Joris, M. M., A. C. Lankester, P. A. von dem Borne, J. Kuball, M. Bierings, J. J. Cornelissen, . . . F. H. Claas (2014). "Translating in vitro prediction of cytotoxic T cell alloreactivity to hematopoietic stem cell transplantation outcome." Transpl Immunol **30**(2-3): 59-64.

Juric, M. K., S. Ghimire, J. Ogonek, E. M. Weissinger, E. Holler, J. J. van Rood, . . . H. T. Greinix (2016). "Milestones of Hematopoietic Stem Cell Transplantation - From First Human Studies to Current Developments." Front Immunol **7**: 470.

Kanda, J., R. Brazauskas, Z. H. Hu, Y. Kuwatsuka, K. Nagafuji, H. Kanamori, . . . W. Saber (2016). "Graft-versus-Host Disease after HLA-Matched Sibling Bone Marrow or Peripheral Blood Stem Cell Transplantation: Comparison of North American Caucasian and Japanese Populations." Biol Blood Marrow Transplant **22**(4): 744-751.

Kauppi, L., M. P. Stumpf and A. J. Jeffreys (2005). "Localized breakdown in linkage disequilibrium does not always predict sperm crossover hot spots in the human MHC class II region." Genomics **86**(1): 13-24.

Keen, L. J., T. E. DeFor, J. L. Bidwell, S. M. Davies, B. A. Bradley and J. M. Hows (2004). "Interleukin-10 and tumor necrosis factor alpha region haplotypes predict transplant-related mortality after unrelated donor stem cell transplantation." Blood **103**(9): 3599-3602.

Kim, D. H., J. G. Kim, S. K. Sohn, W. J. Sung, J. S. Suh, K. S. Lee and K. B. Lee (2004). "Clinical impact of early absolute lymphocyte count after allogeneic stem cell transplantation." Br J Haematol **125**(2): 217-224.

Klein, L., B. Kyewski, P. M. Allen and K. A. Hogquist (2014). "Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see)." Nat Rev Immunol **14**(6): 377-391.

Kollman, C., C. W. Howe, C. Anasetti, J. H. Antin, S. M. Davies, A. H. Filipovich, . . . D. L. Confer (2001). "Donor characteristics as risk factors in recipients after transplantation of bone marrow from unrelated donors: the effect of donor age." Blood **98**(7): 2043-2051.

Kollman, C., S. R. Spellman, M. J. Zhang, A. Hassebroek, C. Anasetti, J. H. Antin, . . . M. Eapen (2016). "The effect of donor characteristics on survival after unrelated donor transplantation for hematologic malignancy." Blood **127**(2): 260-267.

Koyama, D., M. Murata, R. Hanajiri, T. Akashi, S. Okuno, S. Kamoshita, . . . H. Kiyoi (2019). "Quantitative Assessment of T Cell Clonotypes in Human Acute Graft-versus-Host Disease Tissues." Biol Blood Marrow Transplant **25**(3): 417-423.

Krishna, C., D. Chowell, M. Gonen, Y. Elhanati and T. A. Chan (2020). "Genetic and environmental determinants of human TCR repertoire diversity." Immun Ageing **17**: 26.

Lakshmikanth, T., A. Olin, Y. Chen, J. Mikes, E. Fredlund, M. Remberger, . . . P. Brodin (2017). "Mass Cytometry and Topological Data Analysis Reveal Immune Parameters Associated with Complications after Allogeneic Stem Cell Transplantation." Cell Rep **20**(9): 2238-2250.

Lalli, P. N. (2023). "The Not-So-Straightforward Path for Selecting Haploidentical Donors." Transplant Cell Ther **29**(3): 141-142.

Le Blanc, K., A. J. Barrett, M. Schaffer, H. Hagglund, P. Ljungman, O. Ringden and M. Remberger (2009). "Lymphocyte recovery is a major determinant of outcome after matched unrelated myeloablative transplantation for myelogenous malignancies." Biol Blood Marrow Transplant **15**(9): 1108-1115.

Lee, S. J., J. Klein, M. Haagenson, L. A. Baxter-Lowe, D. L. Confer, M. Eapen, . . . C. Anasetti (2007). "High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation." Blood **110**(13): 4576-4583.

Li, L., H. Chen, D. Marin, Y. Xi, Q. Miao, J. Lv, . . . K. Rezvani (2019). "A novel immature natural killer cell subpopulation predicts relapse after cord blood transplantation." Blood Adv **3**(23): 4117-4130.

Li, X., M. Soler, C. Szydzik, K. Khoshmanesh, J. Schmidt, G. Coukos, . . . H. Altug (2018). "Label-Free Optofluidic Nanobiosensor Enables Real-Time Analysis of Single-Cell Cytokine Secretion." Small **14**(26): e1800698.

Link-Rachner, C. S., A. Eugster, E. Rucker-Braun, F. Heidenreich, U. Oelschlagel, A. Dahl, . . . J. Schetelig (2018). "T cell receptor alpha repertoire of CD8+ T cells following allogeneic stem cell transplantation using next-generation sequencing." Haematologica.

Little, A. M., A. Akbarzad-Yousefi, A. Anand, N. Diaz Burlinson, P. P. J. Dunn, I. Evseeva, . . . P. A. Wright (2021). "BSHI guideline: HLA matching and donor selection for haematopoietic progenitor cell transplantation." Int J Immunogenet **48**(2): 75-109.

Little, A. M., A. Green, J. Harvey, S. Hemmatpour, K. Latham, S. G. Marsh, . . . D. Sage (2016). "BSHI Guideline: HLA matching and donor selection for haematopoietic progenitor cell transplantation." Int J Immunogenet.

Liu, C. (2021). "A long road/read to rapid high-resolution HLA typing: The nanopore perspective." Hum Immunol **82**(7): 488-495.

Liu, Y. C., S. Ansaryan, X. Li, E. R. Arvelo and H. Altug (2022). "Real-time monitoring of single-cell secretion with a high-throughput nanoplasmonic microarray." Biosens Bioelectron **202**: 113955.

Locatelli, F., D. Pende, M. Falco, M. Della Chiesa, A. Moretta and L. Moretta (2018). "NK Cells Mediate a Crucial Graft-versus-Leukemia Effect in Haploidentical-HSCT to Cure High-Risk Acute Leukemia." Trends Immunol **39**(7): 577-590.

Lohmeyer, J. K., T. Hirai, M. Turkoz, S. Buhler, T. Lopes Ramos, N. Kohler, . . . F. Simonetta (2022). "Analysis of T cell Repertoire and Transcriptome Identifies Mechanisms of Regulatory T cell (Treg) Suppression of GvHD." Blood.

Lorentino, F., N. Sacchi, E. Oldani, V. Miotti, A. Picardi, A. M. Gallina, . . . R. Italian Bone Marrow Donor (2019). "Comparative evaluation of biological HLA-DPB1 mismatch models for survival and graft versus host disease prediction after unrelated donor hematopoietic cell transplantation." Haematologica.

Luznik, L., P. V. O'Donnell, H. J. Symons, A. R. Chen, M. S. Leffell, M. Zahurak, . . . E. J. Fuchs (2008). "HLA-haploidentical bone marrow transplantation for hematologic malignancies using nonmyeloablative conditioning and high-dose, posttransplantation cyclophosphamide." Biol Blood Marrow Transplant **14**(6): 641-650.

Ma, J., Y. Mo, M. Tang, J. Shen, Y. Qi, W. Zhao, . . . C. Qian (2021). "Bispecific Antibodies: From Research to Clinical Application." Front Immunol **12**: 626616.

Macedo, C., E. A. Orkis, I. Popescu, B. D. Elinoff, A. Zeevi, R. Shapiro, . . . D. Metes (2009). "Contribution of naive and memory T-cell populations to the human alloimmune response." Am J Transplant **9**(9): 2057-2066.

Mackall, C. L., C. V. Bare, L. A. Granger, S. O. Sharrow, J. A. Titus and R. E. Gress (1996). "Thymic-independent T cell regeneration occurs via antigen-driven expansion of peripheral T cells resulting in a repertoire that is limited in diversity and prone to skewing." J Immunol **156**(12): 4609-4616.

Marin, W. M., R. Dandekar, D. G. Augusto, T. Yusufali, B. Heyn, J. Hofmann, . . . J. A. Hollenbach (2021). "High-throughput Interpretation of Killer-cell Immunoglobulin-like Receptor Short-read Sequencing Data with PING." PLoS Comput Biol **17**(8): e1008904.

Marino, J., J. Paster and G. Benichou (2016). "Allorecognition by T Lymphocytes and Allograft Rejection." Front Immunol **7**: 582.

Marostica, A. S., K. Nunes, E. C. Castelli, N. S. B. Silva, B. S. Weir, J. Goudet and D. Meyer (2022). "How HLA diversity is apportioned: influence of selection and relevance to transplantation." Philos Trans R Soc Lond B Biol Sci **377**(1852): 20200420.

Martin, P. J., D. M. Levine, B. E. Storer, E. H. Warren, X. Zheng, S. C. Nelson, . . . J. A. Hansen (2017). "Genome-wide minor histocompatibility matching as related to the risk of graft-versus-host disease." Blood **129**(6): 791-798.

Martin, P. S., S. Li, S. Nikiforow, E. P. Alyea, 3rd, J. H. Antin, P. Armand, . . . R. J. Soiffer (2016). "Infused total nucleated cell dose is a better predictor of transplant outcomes than CD34+ cell number in reduced-intensity mobilized peripheral blood allogeneic hematopoietic cell transplantation." Haematologica **101**(4): 499-505.

Mayor, N. P., J. D. Hayhurst, T. R. Turner, R. M. Szydlo, B. E. Shaw, W. P. Bultitude, . . . S. G. E. Marsh (2019). "Recipients Receiving Better HLA-Matched Hematopoietic Cell Transplantation Grafts, Uncovered by a Novel HLA Typing Method, Have Superior Survival: A Retrospective Study." Biol Blood Marrow Transplant **25**(3): 443-450.

Mayor, N. P., J. Robinson, A. J. McWhinnie, S. Ranade, K. Eng, W. Midwinter, . . . S. G. Marsh (2015). "HLA Typing for the Next Generation." PLoS One **10**(5): e0127153.

McCurdy, S. R. and E. J. Fuchs (2016). "Selecting the best haploidentical donor." Semin Hematol **53**(4): 246-251.

Meurer, T., E. Arrieta-Bolanos, M. Metzing, M. M. Langer, P. van Balen, J. H. F. Falkenburg, . . . P. Crivello (2018). "Dissecting Genetic Control of HLA-DPB1 Expression and Its Relation to Structural Mismatch Models in Hematopoietic Stem Cell Transplantation." Front Immunol **9**: 2236.

Meurer, T., P. Crivello, M. Metzing, M. Kester, D. A. Megger, W. Chen, . . . K. Fleischhauer (2021). "Permissive HLA-DPB1 mismatches in HCT depend on immunopeptidome divergence and editing by HLA-DM." Blood **137**(7): 923-928.

Meyer, D. and G. Thomson (2001). "How selection shapes variation of the human major histocompatibility complex: a review." Ann Hum Genet **65**(Pt 1): 1-26.

Meysman, P., N. De Neuter, S. Gielis, D. Bui Thi, B. Ogunjimi and K. Laukens (2019). "On the viability of unsupervised T-cell receptor sequence clustering for epitope preference." Bioinformatics **35**(9): 1461-1468.

Mittelbrunn, M. and G. Kroemer (2021). "Hallmarks of T cell aging." Nat Immunol **22**(6): 687-698.

Monos, D. and M. J. Maiers (2015). "Progressing towards the complete and thorough characterization of the HLA genes by NGS (or single-molecule DNA sequencing): Consequences, opportunities and challenges." Hum Immunol **76**(12): 883-886.

Morishima, S., K. Kashiwase, K. Matsuo, F. Azuma, T. Yabe, A. Sato-Otsubo, . . . P. Japan Marrow Donor (2016). "High-risk HLA alleles for severe acute graft-versus-host disease and mortality in unrelated donor bone marrow transplantation." Haematologica **101**(4): 491-498.

Morishima, S., S. Ogawa, A. Matsubara, T. Kawase, Y. Nannya, K. Kashiwase, . . . P. Japan Marrow Donor (2010). "Impact of highly conserved HLA haplotype on acute graft-versus-host disease." Blood **115**(23): 4664-4670.

Morishima, S., T. Shiina, S. Suzuki, S. Ogawa, A. Sato-Otsubo, K. Kashiwase, . . . P. Japan Marrow Donor (2018). "Evolutionary basis of HLA-DPB1 alleles affects acute GVHD in unrelated donor stem cell transplantation." Blood **131**(7): 808-817.

Mullally, A. and J. Ritz (2007). "Beyond HLA: the significance of genomic variation for allogeneic hematopoietic stem cell transplantation." Blood **109**(4): 1355-1362.

Mussetti, A., A. Paviglianiti, R. Parody and A. Sureda (2021). "Is Post-Transplant Cyclophosphamide the New Methotrexate?" J Clin Med **10**(16).

Mytilineos, D., C. Tsamadou, C. Neuchel, U. Platzbecker, D. Bunjes, N. Schub, . . . J. Mytilineos (2020). "The Human Leukocyte Antigen-DPB1 Degree of Compatibility Is

Determined by Its Expression Level and Mismatch Permissiveness: A German Multicenter Analysis." Front Immunol **11**: 614976.

Nagler, A., M. Labopin, B. Dholaria, E. Angelucci, B. Afanasyev, J. J. Cornelissen, . . . M. Mohty (2021). "Comparison of Haploidentical Bone Marrow versus Matched Unrelated Donor Peripheral Blood Stem Cell Transplantation with Posttransplant Cyclophosphamide in Patients with Acute Leukemia." Clin Cancer Res **27**(3): 843-851.

Nikolich-Zugich, J., M. K. Slifka and I. Messaoudi (2004). "The many important facets of T-cell repertoire diversity." Nat Rev Immunol **4**(2): 123-132.

Norman, P. J., J. A. Hollenbach, N. Nemat-Gorgani, W. M. Marin, S. J. Norberg, E. Ashouri, . . . P. Parham (2016). "Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing." Am J Hum Genet **99**(2): 375-391.

Nunes, E., H. Heslop, M. Fernandez-Vina, C. Taves, D. R. Wagenknecht, A. B. Eisenbrey, . . . N. Sacchi (2011). "Definitions of histocompatibility typing terms." Blood **118**(23): e180-183.

O'Donnell, P. V., L. Luznik, R. J. Jones, G. B. Vogelsang, M. S. Leffell, M. Phelps, . . . E. J. Fuchs (2002). "Nonmyeloablative bone marrow transplantation from partially HLA-mismatched related donors using posttransplantation cyclophosphamide." Biol Blood Marrow Transplant **8**(7): 377-386.

Ogonek, J., M. Kralj Juric, S. Ghimire, P. R. Varanasi, E. Holler, H. Greinix and E. Weissinger (2016). "Immune Reconstitution after Allogeneic Hematopoietic Stem Cell Transplantation." Front Immunol **7**: 507.

Ogonek, J., P. Varanasi, S. Luther, P. Schweier, W. Kuhnau, G. Gohring, . . . L. Hambach (2017). "Possible Impact of Cytomegalovirus-Specific CD8(+) T Cells on Immune Reconstitution and Conversion to Complete Donor Chimerism after Allogeneic Stem Cell Transplantation." Biol Blood Marrow Transplant **23**(7): 1046-1053.

Oostvogels, R., H. M. Lokhorst, M. C. Minnema, M. van Elk, K. van den Oudenalder, E. Spierings, . . . R. M. Spaapen (2014). "Identification of minor histocompatibility antigens based on the 1000 Genomes Project." Haematologica **99**(12): 1854-1859.

Pagliuca, S., C. Gurnari, S. Hong, R. Zhao, S. Kongkiatkamon, L. Terkawi, . . . J. P. Maciejewski (2021). "Clinical and basic implications of dynamic T cell receptor clonotyping in hematopoietic cell transplantation." JCI Insight **6**(13).

Pagliuca, S., C. Gurnari, M. T. Rubio, V. Visconte and T. L. Lenz (2022). "Individual HLA heterogeneity and its implications for cellular immune evasion in cancer and beyond." Frontiers in Immunology **13**.

Parham, P. and L. A. Guethlein (2018). "Genetics of Natural Killer Cells in Human Health, Disease, and Survival." Annu Rev Immunol **36**: 519-548.

Parmar, S., M. Del Lima, Y. Zou, P. A. Patah, P. Liu, P. Cano, . . . M. Fernandez-Vina (2009). "Donor-recipient mismatches in MHC class I chain-related gene A in unrelated donor transplantation lead to increased incidence of acute graft-versus-host disease." Blood **114**(14): 2884-2887.

Pasi, A., R. Crocchiolo, M. Bontempelli, C. Carcassi, G. Carella, L. Crespiatico, . . . M. Martinetti (2011). "The conundrum of HLA-DRB1*14:01/*14:54 and HLA-DRB3*02:01/*02:02 mismatches in unrelated hematopoietic SCT." Bone Marrow Transplant **46**(7): 916-922.

Passweg, J. R., H. Baldomero, M. Ansari, G. M. Baerlocher, M. Bargetzi, Y. Chalandon, . . . S. For The Swiss Blood Stem Cell Transplantation Group (2018). "Haematopoietic cell transplantation in Switzerland, changes and results over 20 years: a report from the Swiss Blood Stem Cell Transplantation Working Group for Blood and Marrow Transplantation registry 1997-2016." Swiss Med Wkly **148**: w14589.

Passweg, J. R., H. Baldomero, P. Bader, C. Bonini, R. F. Duarte, C. Dufour, . . . M. Mohty (2017). "Use of haploidentical stem cell transplantation continues to increase: the 2015 European Society for Blood and Marrow Transplant activity survey report." Bone Marrow Transplant.

Passweg, J. R., H. Baldomero, C. Chabannon, G. W. Basak, R. de la Camara, S. Corbacioglu, . . . T. Marrow (2021). "Hematopoietic cell transplantation and cellular therapy survey of the EBMT: monitoring of activities and trends over 30 years." Bone Marrow Transplant **56**(7): 1651-1664.

Passweg, J. R., U. Schanz, Y. Chalandon, T. Gungor, H. Baldomero, D. Heim, . . . G. Swiss Blood Stem Cell Transplantation (2015). "High-resolution HLA matching in unrelated donor transplantation in Switzerland: differential impact of class I and class II mismatches may reflect selection of nonimmunogenic or weakly immunogenic DRB1/DQB1 disparities." Bone Marrow Transplant **50**(9): 1201-1205.

Pende, D., M. Falco, M. Vitale, C. Cantoni, C. Vitale, E. Munari, . . . L. Moretta (2019). "Killer Ig-Like Receptors (KIRs): Their Role in NK Cell Modulation and Developments Leading to Their Clinical Exploitation." Front Immunol **10**: 1179.

Petersdorf, E. W. (2008). "Optimal HLA matching in hematopoietic cell transplantation." Curr Opin Immunol **20**(5): 588-593.

Petersdorf, E. W. (2013). "Genetics of graft-versus-host disease: the major histocompatibility complex." Blood Rev **27**(1): 1-12.

Petersdorf, E. W. (2016). "Mismatched unrelated donor transplantation." Semin Hematol **53**(4): 230-236.

Petersdorf, E. W. (2017). "Which factors influence the development of GVHD in HLA-matched or mismatched transplants?" Best Pract Res Clin Haematol **30**(4): 333-335.

Petersdorf, E. W., M. Bengtsson, D. De Santis, V. Dubois, K. Fleischhauer, T. Gooley, . . . T. International Histocompatibility Working Group in Hematopoietic-Cell (2020). "Role of HLA-DP Expression in Graft-Versus-Host Disease After Unrelated Donor Transplantation." J Clin Oncol **38**(24): 2712-2718.

Petersdorf, E. W., M. Bengtsson, M. Horowitz, C. McKallor, S. R. Spellman, E. Spierings, . . . T. International Histocompatibility Working Group in Hematopoietic Cell (2022). "HLA-DQ heterodimers in hematopoietic cell transplantation." Blood **139**(20): 3009-3017.

Petersdorf, E. W., M. Carrington, C. O'HUigin, M. Bengtsson, D. De Santis, V. Dubois, . . . T. International Histocompatibility Working Group in Hematopoietic Cell (2020). "Role of HLA-B exon 1 in graft-versus-host disease after unrelated haemopoietic cell transplantation: a retrospective cohort study." Lancet Haematol **7**(1): e50-e60.

Petersdorf, E. W., T. Gooley, M. Malkki, C. Anasetti, P. Martin, A. Woolfrey, . . . J. A. Hansen (2001). "The biological significance of HLA-DP gene variation in haematopoietic cell transplantation." Br J Haematol **112**(4): 988-994.

Petersdorf, E. W., T. A. Gooley, M. Malkki, A. P. Bacigalupo, A. Cesbron, E. Du Toit, . . . T. International Histocompatibility Working Group in Hematopoietic Cell (2014). "HLA-C expression levels define permissible mismatches in hematopoietic cell transplantation." Blood **124**(26): 3996-4003.

Petersdorf, E. W., M. Malkki, T. A. Gooley, P. J. Martin and Z. Guo (2007). "MHC haplotype matching for unrelated hematopoietic cell transplantation." PLoS Med **4**(1): e8.

Petersdorf, E. W., M. Malkki, T. A. Gooley, S. R. Spellman, M. D. Haagenson, M. M. Horowitz and T. Wang (2012). "MHC-resident variation affects risks after unrelated donor hematopoietic cell transplantation." Sci Transl Med **4**(144): 144ra101.

Petersdorf, E. W., M. Malkki, M. M. Horowitz, S. R. Spellman, M. D. Haagenson and T. Wang (2013). "Mapping MHC haplotype effects in unrelated donor hematopoietic cell transplantation." Blood **121**(10): 1896-1905.

Petersdorf, E. W., M. Malkki, K. Hsu, P. Bardy, A. Cesbron, A. Dickinson, . . . T. International Histocompatibility Working Group in Hematopoietic Cell (2013). "16th IHIW: international histocompatibility working group in hematopoietic cell transplantation." Int J Immunogenet **40**(1): 2-10.

Petersdorf, E. W., M. Malkki, C. O'HUigin, M. Carrington, T. Gooley, M. D. Haagenson, . . . P. Stevenson (2015). "High HLA-DP Expression and Graft-versus-Host Disease." N Engl J Med **373**(7): 599-609.

Petersdorf, E. W., C. McKallor, M. Malkki, M. He, S. Spellman, K. C. Hsu, . . . P. A. Stevenson (2023). "Role of NKG2D Ligands and Receptor in Haploidentical Related Donor Hematopoietic Cell Transplantation." Blood Adv.

Petersdorf, E. W., P. Stevenson, M. Bengtsson, D. De Santis, V. Dubois, T. Gooley, . . . M. Carrington (2020). "HLA-B leader and survivorship after HLA-mismatched unrelated donor transplantation." Blood **136**(3): 362-369.

Phan, T. L., J. C. Pritchett, C. Leifer, D. M. Zerr, D. M. Koelle, D. Di Luca and P. Lusso (2018). "HHV-6B infection, T-cell reconstitution, and graft-vs-host disease after hematopoietic stem cell transplantation." Bone Marrow Transplant **53**(12): 1508-1517.

Pidala, J. and C. Anasetti (2023). "Fine-tuning alloreactivity against HLA-DP to control leukemia with tolerable graft-versus-host disease." Haematologica **108**(2): 301-302.

Pidala, J., S. J. Lee, K. W. Ahn, S. Spellman, H. L. Wang, M. Aljurf, . . . C. Anasetti (2014). "Nonpermissive HLA-DPB1 mismatch increases mortality after myeloablative unrelated allogeneic hematopoietic cell transplantation." Blood **124**(16): 2596-2606.

Pidala, J., T. Wang, M. Haagenson, S. R. Spellman, M. Askar, M. Battiwalla, . . . C. Anasetti (2013). "Amino acid substitution at peptide-binding pockets of HLA class I molecules increases risk of severe acute GVHD and mortality." Blood **122**(22): 3651-3658.

Pishesha, N., T. J. Harmand and H. L. Ploegh (2022). "A guide to antigen processing and presentation." Nat Rev Immunol.

Ramadan, A. and S. Paczesny (2015). "Various forms of tissue damage and danger signals following hematopoietic stem-cell transplantation." Front Immunol **6**: 14.

Reindl-Schwaighofer, R., A. Heinzel, G. A. Gualdoni, L. Mesnard, F. H. J. Claas and R. Oberbauer (2020). "Novel insights into non-HLA alloimmunity in kidney transplantation." Transpl Int **33**(1): 5-17.

Rene, C., C. Lozano and J. F. Eliaou (2016). "Expression of classical HLA class I molecules: regulation and clinical impacts: Julia Bodmer Award Review 2015." HLA **87**(5): 338-349.

Rimando, J., M. Slade, J. F. DiPersio, P. Westervelt, F. Gao, C. Liu and R. Romee (2020). "The Predicted Indirectly Recognizable HLA Epitopes (PIRCHE) Score for HLA Class I Graft-versus-Host Disparity Is Associated with Increased Acute Graft-versus-Host Disease in Haploidentical Transplantation with Post-Transplantation Cyclophosphamide." Biol Blood Marrow Transplant **26**(1): 123-131.

Robins, H. (2013). "Immunosequencing: applications of immune repertoire deep sequencing." Curr Opin Immunol **25**(5): 646-652.

Robinson, J., D. J. Barker, X. Georgiou, M. A. Cooper, P. Flicek and S. G. E. Marsh (2020). "IPD-IMGT/HLA Database." Nucleic Acids Res **48**(D1): D948-D955.

Roche, P. A. and K. Furuta (2015). "The ins and outs of MHC class II-mediated antigen processing and presentation." Nat Rev Immunol **15**(4): 203-216.

Rolle, A., D. Jager and F. Momburg (2018). "HLA-E Peptide Repertoire and Dimorphism-Centerpieces in the Adaptive NK Cell Puzzle?" Front Immunol **9**: 2410.

Romee, R., M. Rosario, M. M. Berrien-Elliott, J. A. Wagner, B. A. Jewell, T. Schappe, . . . T. A. Fehniger (2016). "Cytokine-induced memory-like natural killer cells exhibit enhanced responses against myeloid leukemia." Sci Transl Med **8**(357): 357ra123.

Roy, D. C. and C. Perreault (2017). "Major vs minor histocompatibility antigens." Blood **129**(6): 664-666.

Ruggeri, A., L. C. De Wreede, C. R. Muller, P. Crivello, E. F. Bonneville, E. W. Petersdorf, . . . K. Fleischhauer (2023). "Integrating biological HLA-DPB1 mismatch models to predict survival after unrelated hematopoietic cell transplantation." Haematologica **108**(2): 645-652.

Ruggeri, L., M. Capanni, E. Urbani, K. Perruccio, W. D. Shlomchik, A. Tosti, . . . A. Velardi (2002). "Effectiveness of donor natural killer cell alloreactivity in mismatched hematopoietic transplants." Science **295**(5562): 2097-2100.

Sacirbegovic, F., M. Günther, A. Greco, D. Zhao, X. Wang, M. Zhou, . . . W. D. Shlomchik (2023). "Graft-versus-host disease is locally maintained in target tissues by resident progenitor-like T cells." Immunity **56**(2): 369-385.e366.

Sajulga, R., Y. T. Bolon, M. J. Maiers and E. W. Petersdorf (2022). "Assessment of HLA-B genetic variation with an HLA-B leader tool and implications in clinical transplantation." Blood Adv **6**(1): 270-280.

Sanchez-Mazas, A. and D. Meyer (2014). "The relevance of HLA sequencing in population genetics studies." J Immunol Res **2014**: 971818.

Sanchez-Mazas, A. and J. M. Nunes (2018). "Does NGS typing highlight our understanding of HLA population diversity? Some good reasons to say yes and a few to say be careful." Hum Immunol.

Sanchez-Mazas, A., J. M. Nunes, D. Middleton, J. Sauter, S. Buhler, A. McCabe, . . . K. Fleischhauer (2017). "Common and well-documented HLA alleles over all of Europe and within European sub-regions: A catalogue from the European Federation for Immunogenetics." HLA **89**(2): 104-113.

Sanz, J., J. E. Galimard, M. Labopin, B. Afanasyev, E. Angelucci, F. Ciceri, . . . T. Marrow (2020). "Post-transplant cyclophosphamide after matched sibling, unrelated and haploidentical donor transplants in patients with acute myeloid leukemia: a comparative study of the ALWP EBMT." J Hematol Oncol **13**(1): 46.

Sanz, J., J. E. Galimard, M. Labopin, B. Afanasyev, M. I. Sergeevich, E. Angelucci, . . . M. Mohty (2021). "Post-transplant cyclophosphamide containing regimens after matched sibling, matched unrelated and haploidentical donor transplants in patients with acute lymphoblastic leukemia in first complete remission, a comparative study of the ALWP of the EBMT." J Hematol Oncol **14**(1): 84.

Savani, B. N., S. Mielke, K. Rezvani, A. Montero, A. S. Yong, L. Wish, . . . A. J. Barrett (2007). "Absolute lymphocyte count on day 30 is a surrogate for robust hematopoietic recovery and strongly predicts outcome after T cell-depleted allogeneic stem cell transplantation." Biol Blood Marrow Transplant **13**(10): 1216-1223.

Schetelig, J., H. Baldauf, F. Heidenreich, C. Massalski, S. Frank, J. Sauter, . . . M. Bornhauser (2020). "External validation of models for KIR2DS1/KIR3DL1-informed selection of hematopoietic cell donors fails." Blood **135**(16): 1386-1395.

Schone, B., S. Bergmann, K. Lang, I. Wagner, A. H. Schmidt, E. W. Petersdorf and V. Lange (2018). "Predicting an HLA-DPB1 expression marker based on standard DPB1 genotyping: Linkage analysis of over 32,000 samples." Hum Immunol **79**(1): 20-27.

Seggewiss, R. and H. Einsele (2010). "Immune reconstitution after allogeneic transplantation and expanding options for immunomodulation: an update." Blood **115**(19): 3861-3868.

Sewell, A. K. (2012). "Why must T cells be cross-reactive?" Nat Rev Immunol **12**(9): 669-677.

Shaw, B. E. (2017). "Related haploidentical donors are a better choice than matched unrelated donors: Counterpoint." Blood Adv **1**(6): 401-406.

Shaw, B. E., T. Gooley, J. A. Madrigal, M. Malkki, S. G. Marsh and E. W. Petersdorf (2007). "Clinical importance of HLA-DPB1 in haematopoietic cell transplantation." Tissue Antigens **69 Suppl 1**: 36-41.

Shaw, B. E., T. A. Gooley, M. Malkki, J. A. Madrigal, A. B. Begovich, M. M. Horowitz, . . . E. W. Petersdorf (2007). "The importance of HLA-DPB1 in unrelated donor hematopoietic cell transplantation." Blood **110**(13): 4560-4566.

Shaw, B. E., A. M. Jimenez-Jimenez, L. J. Burns, B. R. Logan, F. Khimani, B. C. Shaffer, . . . J. Bolanos-Meade (2021). "National Marrow Donor Program-Sponsored Multicenter, Phase II Trial of HLA-Mismatched Unrelated Donor Bone Marrow Transplantation Using Post-Transplant Cyclophosphamide." J Clin Oncol **39**(18): 1971-1982.

Shimoni, A., M. Labopin, F. Lorentino, M. T. Van Lint, Y. Koc, Z. Gulbas, . . . A. Nagler (2019). "Killer cell immunoglobulin-like receptor ligand mismatching and outcome after haploidentical transplantation with post-transplant cyclophosphamide." Leukemia **33**(1): 230-239.

Shugay, M., D. V. Bagaev, I. V. Zvyagin, R. M. Vroomans, J. C. Crawford, G. Dolton, . . . D. M. Chudakov (2018). "VDJdb: a curated database of T-cell receptor sequences with known antigen specificity." Nucleic Acids Res **46**(D1): D419-D427.

Six, A., M. E. Mariotti-Ferrandiz, W. Chaara, S. Magadan, H. P. Pham, M. P. Lefranc, . . . P. Boudinot (2013). "The past, present, and future of immune repertoire biology - the rise of next-generation repertoire analysis." Front Immunol **4**: 413.

Snowden, J. A., I. Sanchez-Ortega, S. Corbacioglu, G. W. Basak, C. Chabannon, R. de la Camara, . . . T. Marrow (2022). "Indications for haematopoietic cell transplantation for haematological diseases, solid tumours and immune disorders: current practice in Europe, 2022." Bone Marrow Transplant **57**(8): 1217-1239.

Soto, C., R. G. Bombardi, M. Kozhevnikov, R. S. Sinkovits, E. C. Chen, A. Branchizio, . . . J. E. Crowe, Jr. (2020). "High Frequency of Shared Clonotypes in Human T Cell Receptor Repertoires." Cell Rep **32**(2): 107882.

Spellman, S. R. (2022). "Hematology 2022-what is complete HLA match in 2022?" Hematology Am Soc Hematol Educ Program **2022**(1): 83-89.

Spellman, S. R., M. Eapen, B. R. Logan, C. Mueller, P. Rubinstein, M. I. Setterholm, . . . R. Marrow Transplant (2012). "A perspective on the selection of unrelated donors and cord blood units for transplantation." Blood **120**(2): 259-265.

Spierings, E. (2014). "Minor histocompatibility antigens: past, present, and future." Tissue Antigens **84**(4): 374-360.

Stenger, W., A. Kunkele, M. Niemann, K. Todorova, A. Pruss, J. H. Schulte, . . . L. Oevermann (2020). "Donor selection in a pediatric stem cell transplantation cohort using PIRCHE and HLA-DPB1 typing." Pediatr Blood Cancer **67**(3): e28127.

Stern, L., H. McGuire, S. Avdic, S. Rizzetto, B. Fazekas de St Groth, F. Luciani, . . . E. Blyth (2018). "Mass Cytometry for the Assessment of Immune Reconstitution After Hematopoietic Stem Cell Transplantation." Front Immunol **9**: 1672.

Storek, J., M. Geddes, F. Khan, B. Huard, C. Helg, Y. Chalandon, . . . E. Roosnek (2008). "Reconstitution of the immune system after hematopoietic stem cell transplantation in humans." Semin Immunopathol **30**(4): 425-437.

Suessmuth, Y., R. Mukherjee, B. Watkins, D. T. Koura, K. Finstermeier, C. Desmarais, . . . L. S. Kean (2015). "CMV reactivation drives posttransplant T-cell reconstitution and results in defects in the underlying TCRbeta repertoire." Blood **125**(25): 3835-3850.

Sylwester, A. W., B. L. Mitchell, J. B. Edgar, C. Taormina, C. Pelte, F. Ruchti, . . . L. J. Picker (2005). "Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects." J Exp Med **202**(5): 673-685.

Tambur, A. R. (2018). "HLA-Epitope Matching or Eplet Risk Stratification: The Devil Is in the Details." Front Immunol **9**: 2010.

Tambur, A. R. and F. H. Claas (2015). "HLA epitopes as viewed by antibodies: what is it all about?" Am J Transplant **15**(5): 1148-1154.

Tay, G. K., C. S. Witt, F. T. Christiansen, D. Charron, D. Baker, R. Herrmann, . . . et al. (1995). "Matching for MHC haplotypes results in improved survival following unrelated bone marrow transplantation." Bone Marrow Transplant **15**(3): 381-385.

Terasaki, P. I. (2013). "The history of HLA and transplantation." Hirosaki Med. J **64, Supplement**: S45—S52.

Thorsby, E. (2009). "A short history of HLA." Tissue Antigens **74**(2): 101-116.

Thus, K. A., R. A. de Weger, T. A. de Hoop, V. E. Boers Trilles, J. Kuball and E. Spierings (2014). "Complete donor chimerism is a prerequisite for the effect of Predicted Indirectly ReCognizable HLA Epitopes (PIRCHE) on acute graft-versus-host disease." (1938-1964 (Electronic)).

Thus, K. A., M. T. Ruizendaal, T. A. de Hoop, E. Borst, H. W. van Deutekom, L. Te Boome, . . . E. Spierings (2014). "Refinement of the definition of permissible HLA-DPB1 mismatches with predicted indirectly recognizable HLA-DPB1 epitopes." Biol Blood Marrow Transplant **20**(11): 1705-1710.

Thus, K. A., L. Te Boome, J. Kuball and E. Spierings (2014). "Indirectly Recognized HLA-C Mismatches and Their Potential Role in Transplant Outcome." Front Immunol **5**: 210.

Tian, G., M. Li and G. Lv (2022). "Analysis of T-Cell Receptor Repertoire in Transplantation: Fingerprint of T Cell-mediated Alloresponse." Frontiers in Immunology **12**.

Tickotsky, N., T. Sagiv, J. Prilusky, E. Shifrut and N. Friedman (2017). "McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences." Bioinformatics **33**(18): 2924-2929.

Tiercy, J. M. (2014). "HLA-C Incompatibilities in Allogeneic Unrelated Hematopoietic Stem Cell Transplantation." Front Immunol **5**: 216.

Tiercy, J. M. (2016). "How to select the best available related or unrelated donor of hematopoietic stem cells?" Haematologica **101**(6): 680-687.

Tiercy, J. M. and F. Claas (2013). "Impact of HLA diversity on donor selection in organ and stem cell transplantation." Hum Hered **76**(3-4): 178-186.

Ting, J. P. and J. Trowsdale (2002). "Genetic control of MHC class II expression." Cell **109 Suppl**: S21-33.

Toffalori, C., L. Zito, V. Gambacorta, M. Riba, G. Oliveira, G. Bucci, . . . L. Vago (2019). "Immune signature drives leukemia escape and relapse after hematopoietic cell transplantation." Nat Med **25**(4): 603-611.

Trofimov, A., P. Brouillard, J. D. Larouche, J. Seguin, J. P. Laverdure, A. Brasey, . . . C. Perreault (2022). "Two types of human TCR differentially regulate reactivity to self and non-self antigens." iScience **25**(9): 104968.

Vadakekolathu, J. and S. Rutella (2017). "T-Cell Manipulation Strategies to Prevent Graft-Versus-Host Disease in Haploidentical Stem Cell Transplantation." Biomedicines **5**(2).

Velardi, E., E. Clave, L. C. M. Arruda, F. Benini, F. Locatelli and A. Toubert (2021). "The role of the thymus in allogeneic bone marrow transplantation and the recovery of the peripheral T-cell compartment." Semin Immunopathol **43**(1): 101-117.

Venstrom, J. M., G. Pittari, T. A. Gooley, J. H. Chewning, S. Spellman, M. Haagenson, . . . K. C. Hsu (2012). "HLA-C-dependent prevention of leukemia relapse by donor activating KIR2DS1." N Engl J Med **367**(9): 805-816.

Vujovic, M., K. F. Degn, F. I. Marin, A. L. Schaap-Johansen, B. Chain, T. L. Andresen, . . . P. Marcatili (2020). "T cell receptor sequence clustering and antigen specificity." Comput Struct Biotechnol J **18**: 2166-2173.

Wagner, I., D. Schefzyk, J. Pruschke, G. Schofl, B. Schone, N. Gruber, . . . V. Lange (2018). "Allele-Level KIR Genotyping of More Than a Million Samples: Workflow, Algorithm, and Observations." Front Immunol **9**: 2843.

Wang, Y., N. K. Singh, T. T. Spear, L. M. Hellman, K. H. Piepenbrink, R. H. McMahan, . . . B. M. Baker (2017). "How an alloreactive T-cell receptor achieves peptide and MHC specificity." Proc Natl Acad Sci U S A **114**(24): E4792-E4801.

Warren, E. H. and H. J. Deeg (2013). "Dissecting graft-versus-leukemia from graft-versus-host-disease using novel strategies." Tissue Antigens **81**(4): 183-193.

Weinberg, K., B. R. Blazar, J. E. Wagner, E. Agura, B. J. Hill, M. Smogorzewska, . . . D. C. Douek (2001). "Factors affecting thymic function after allogeneic hematopoietic stem cell transplantation." Blood **97**(5): 1458-1466.

Wooldridge, L., J. Ekeruche-Makinde, H. A. van den Berg, A. Skowera, J. J. Miles, M. P. Tan, . . . A. K. Sewell (2012). "A single autoimmune T cell receptor recognizes more than a million different peptides." J Biol Chem **287**(2): 1168-1177.

Wu, Y., N. Zhang, K. Hashimoto, C. Xia and J. M. Dijkstra (2021). "Structural Comparison Between MHC Classes I and II; in Evolution, a Class-II-Like Molecule Probably Came First." Front Immunol **12**: 621153.

Yin, Y. and R. A. Mariuzza (2009). "The multiple mechanisms of T cell receptor cross-reactivity." Immunity **31**(6): 849-851.

Zamora, D., E. R. Duke, H. Xie, B. C. Edmison, B. Akoto, R. Kiener, . . . M. Boeckh (2021). "Cytomegalovirus-specific T-cell reconstitution following letermovir prophylaxis after hematopoietic cell transplantation." Blood **138**(1): 34-43.

Zhang, H., X. Zhan and B. Li (2021). "GIANA allows computationally-efficient TCR clustering and multi-disease repertoire classification by isometric transformation." Nat Commun **12**(1): 4699.

Zhao, X. Y., Y. J. Chang, L. P. Xu, X. H. Zhang, K. Y. Liu, D. Li and X. J. Huang (2014). "HLA and KIR genotyping correlates with relapse after T-cell-replete haploidentical transplantation in chronic myeloid leukaemia patients." Br J Cancer **111**(6): 1080-1088.

Zino, E., G. Frumento, S. Marktel, M. P. Sormani, F. Ficara, S. Di Terlizzi, . . . K. Fleischhauer (2004). "A T-cell epitope encoded by a subset of HLA-DPB1 alleles determines nonpermissive mismatches for hematologic stem cell transplantation." Blood **103**(4): 1417-1424.

Zino, E., L. Vago, S. Di Terlizzi, B. Mazzi, L. Zito, E. Sironi, . . . K. Fleischhauer (2007). "Frequency and targeted detection of HLA-DPB1 T cell epitope disparities relevant in unrelated hematopoietic stem cell transplantation." Biol Blood Marrow Transplant **13**(9): 1031-1040.

Annexes (original articles)

1. **S. Buhler**, H. Baldomero, S. Ferrari-Lacraz, J.M. Nunes, A. Sanchez-Mazas, S. Massouridi-Levrat, D. Heim, J. Halter, G. Nair, Y. Chalandon, U. Schanz, T. Güngör, G. Nicoloso, J.-M. Tiercy, J. Passweg, J. Villard (2019). High resolution HLA phased haplotype frequencies to predict the success of unrelated donor searches and clinical outcome following Hematopoietic Stem Cell Transplantation. Bone Marrow Transplantation 54: 1701–1709.
2. **S. Buhler**, F. Bettens, C. Dantin, S. Ferrari-Lacraz, M. Ansari, A.C. Mamez, S. Masouridi-Levrat, Y. Chalandon, J. Villard (2020). Genetic T-cell receptor diversity at 1 year following allogeneic hematopoietic stem cell transplantation. Leukemia 34: 1422-1432.
3. **S. Buhler**, H. Baldomero, S. Ferrari-Lacraz, A.-C. Mamez, S. Masouridi-Levrat, D. Heim, J. Halter, G. Nair, Y. Chalandon, U. Schanz, T. Güngör, G. Nicoloso, J.R. Passweg, J. Villard (2021). Analysis of biological models to predict clinical outcomes based on HLA-DPB1 disparities in unrelated transplantation. Blood Advances 5: 3377-3386.
4. F. Bettens, Z. Calderin Sollet, **S. Buhler**, J. Villard. CD8+ T-cell repertoire in HLA class I-mismatched alloreactive immune response (2021). Frontiers in Immunology 11, 588741.
5. F. Bettens, H. Ongen, G. Rey, **S. Buhler,** Z. Calderin Sollet, E. Dermitzakis, J. Villard. Regulation of HLA class I expression by non-coding gene variations (2022). PloS genetics 18: e1010212.

Bone Marrow Transplantation (2019) 54:1701–1709
https://doi.org/10.1038/s41409-019-0520-6

ARTICLE

High-resolution HLA phased haplotype frequencies to predict the success of unrelated donor searches and clinical outcome following hematopoietic stem cell transplantation

Stéphane Buhler[1,2] · Helen Baldomero[3] · Sylvie Ferrari-Lacraz[1] · José Manuel Nunes[2,4] · Alicia Sanchez-Mazas[2,4] · Stravroula Massouridi-Levrat[5] · Dominik Heim[3] · Jörg Halter[3] · Gayathri Nair[6] · Yves Chalandon[5] · Urs Schanz[6] · Tayfun Güngör[7] · Grazia Nicoloso[8] · Jean-Marie Tiercy[1,8] · Jakob Passweg[3] · Jean Villard[1]

Received: 30 November 2018 / Accepted: 18 March 2019 / Published online: 5 April 2019

Abstract

HLA matching is a critical factor for successful allogeneic hematopoietic stem cell transplantation. For unrelated donor searches, matching is usually based on high-resolution typing at five HLA loci, looking for a 10/10 match. Some studies have proposed that further matching at the haplotype level could be beneficial for clinical outcome. In this study, we determined the phased haplotypes of 291 patients using family members and segregation analysis. The sum of ranks of the haplotypes carried by patients was used as a surrogate predictor of a successful unrelated donor search. The putative impact of haplotypes was then analyzed in a cohort of 211 recipients transplanted with 10/10 matched unrelated donors. A logistic regression analysis showed a highly significant effect of the haplotypes in the outcome of a search, but we did not find any significant effect on overall survival, graft versus host disease or relapse/progression following HSCT. This study provides useful data for the optimization of unrelated bone marrow donor searches, but does not confirm previous reports that matching at the haplotype level has a clinical impact following HSCT. Due to the extreme polymorphism of HLA genes, further studies are warranted to better understand the many factors at play.

Introduction

Human leukocyte antigen (HLA) matching between recipients and donors is a prerequisite for successful allogeneic hematopoietic stem cell transplantation (HSCT), notably to avoid graft versus host disease (GVHD) as main post-transplant complication. Although new protocols for selecting donors are increasingly sought, even across the histocompatibility barrier [1–3], the gold standard is to look first for an HLA identical sibling. If such a genotypically identical sibling cannot be found, the preferred alternative is to search for a 10/10 or 12/12 phenotypically matched unrelated donor (MUD) [4, 5]. With unrelated donors, matching is based on high-resolution typing at HLA-A, B, C, DRB1, DQB1 and possibly DPB1 and DRB3/4/5 with

Supplementary information The online version of this article (https://doi.org/10.1038/s41409-019-0520-6) contains supplementary material, which is available to authorized users.

✉ Stéphane Buhler
stephane.buhler@hcuge.ch

1 Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Department of Diagnostic, Geneva University Hospitals, Geneva, Switzerland

2 Anthropology Unit, AGP Laboratory, Department of Genetics and Evolution, University of Geneva, Geneva, Switzerland

3 Division of Hematology, Basel University Hospital, Basel, Switzerland

4 Institute of Genetics and Genomics in Geneva (iGE3), Geneva, Switzerland

5 Division of Hematology, Department of Oncology, Geneva University Hospitals, Geneva, Switzerland

6 Division of Hematology, University Hospital of Zurich, Zurich, Switzerland

7 Department of Stem Cell Transplantation, University Children's Hospital Zurich, Zurich, Switzerland

8 Swiss Blood Stem Cells Registry, Swiss Transfusion SRC, Bern, Switzerland

no consideration given to putative haplotype matching between the recipient and his donor. However, even with well-matched unrelated donors, risks of transplant-related mortality are higher as compared to matched sibling donors because of minor histocompatibility antigens (mHA) spread across the whole genome and non-HLA linked polymorphisms (e.g. single nucleotide polymorphisms (SNPs), expression quantitative trait loci (eQTL), microsatellites) within the extended HLA region [6–12]. Thus, in an attempt to leverage such a hurdle, several studies have suggested a possible beneficial impact of HLA haplotype matching at reducing post-transplantation complications in patients transplanted with 10/10 MUD [13–16]. A haplotype defines which allele belongs to which copy of the two chromosomes, or alternatively, which alleles segregate together on a single chromosome. In practice, HLA haplotypes can either be phased unambiguously by family segregation analysis or imputed statistically from genotype data and HLA frequencies in populations. The underlying hypothesis is that if a recipient and his unrelated donor are matched for the same common haplotypes, such haplotypes could carry more conserved DNA segments shared by descent (including at nearby favorable SNPs) compared to rare haplotypes present in the population.

Besides the extremely high level of polymorphism, complex patterns of association define classical HLA genes. Some pairs like B~C and DRB1~DQB1 are found in tight association on chromosome 6 [17], whereas other loci are defined by weaker (HLA-A) or non-significant linkage (HLA–DPB1), due to more distant location or recombination hotspots [18, 19]. These characteristics represent a significant hindrance to determine HLA multi-locus haplotypes and their corresponding frequencies [20]. In consequence, powerful methodologies have been developed to assess haplotype frequencies in various populations [21, 22] and in large cohorts of unrelated donors [23], but such approaches need to rely on representative sample sizes and on assumptions that are not always met in practice [22]. In this context, the availability of families typed at several HLA loci for the purpose of HSCT-related donor searches has provided informative data to characterize HLA haplotypes by segregation analysis [24–26], contributing to define the probability of finding suitable unrelated donors [27–29] or to study clinical outcome of unrelated HSCT [13]. However, studies using phased haplotypes remain scarce in the literature and phased HLA haplotypes are needed in more populations because HLA frequencies vary significantly according to geography [30–32].

The first aim of this study was to investigate haplotype segregation in a large cohort of patients and their family living in Switzerland. This would allow us to constitute a reference panel to help in optimizing future unrelated donor searches for the significant proportion of patients in need of a transplantation with no HLA identical sibling [5]. The second aim was to use these phased haplotypes for predicting the outcome of unrelated donor searches for patients waiting for a HSCT. Within the last aim, we analyzed HSCT outcomes in a group of recipients transplanted with 10/10 MUD and we tested their HLA haplotypes frequencies as a potential relevant parameter in the clinical follow-up.

Material and methods

Patients

The phased haplotypes of individuals living in Switzerland were determined from the HLA-A, B and DRB1 typing of 843 patients and 2132 family members (Figure S1), allowing to constitute a cohort of 291 patients with high-resolution phased haplotypes. High-resolution typing was also performed at HLA-C and DQB1 for 290 of these patients as potential candidates for an unrelated donor search. Allogeneic HSCT was performed in 140 patients, including 101 with 10/10 MUD and 39 with mismatched donors. To enlarge the clinical cohort, 111 recipients of 10/10 MUD with family segregation data were obtained from the other Swiss transplant centers (Table S1). This study was approved by the ethical committee of the institution (CER 06-208 and 08-208R), and patients' informed consents were obtained.

Statistical analyses

HLA haplotypes

Haplotype frequencies were either estimated by direct counting on 291 patients based on segregating haplotypes or by using an implementation of the expectation–maximization (EM) algorithm on 6114 unrelated Swiss donors based on multi-locus unphased genotypes. Hardy–Weinberg (HW) equilibrium assumptions were assessed using a nested likelihood procedure. Global linkage disequilibrium between pairs of loci was tested using a resampling procedure and linkage disequilibrium for individual haplotypes was determined using standardized residuals. All these analyses were performed with the hla-net.eu Gene[RATE] tools [21, 33].

Unrelated donor searches

The sum of ranks of the phased haplotypes carried by each patient was considered as a surrogate predictor of a successful search (i.e. finding at least one 10/10 MUD). Haplotypes were ranked based on high-resolution haplotype

frequencies estimated on 6114 donors from the Swiss registry (SBSC). The choice of the sum of haplotype ranks is analogous to a non-parametric approach with the goal of not relying directly and too heavily on estimated haplotype frequencies as search outcome determinants. Confusion matrices, logistic regression and receiver operating characteristic (ROC) curve were generated in R (version 3.5.0) using the packages ggplot2, reshape, caret and ROCR.

Clinical outcome

We followed the reasoning applied by Joris et al. [13] to consider that a low haplotype ranking in a recipient (i.e. carrying one or two frequent haplotypes) was a good proxy for haplotype matching with his 10/10 MUD. As unrelated donors are selected at a worldwide scale, SBSC frequencies were contrasted with frequencies estimated on donors from the United States [34]. Recipients were subdivided into categories based on the ranking of their haplotypes, using rank 50 and rank 20 as two distinct cut-offs to classify haplotypes as common or rare, and were subsequently analyzed in separate models for survival (see Tables S1 and S2, Table 2 and Fig. 2 for the categories considered). Obviously, the chosen cut-offs and categories are arbitrary to a degree, but this provided a compromise to the very heterogeneous values proposed in previous studies for defining common haplotypes [13, 15]. It also allowed to consider alternative groups while keeping sufficient and meaningful numbers of patients within each one. Furthermore, we would expect that a strong effect of haplotypes on clinical outcome should be robust and consistent across different cut-offs to be really considered as a relevant parameter. Secondary outcomes (relapse/progression, acute GVHD, chronic GVHD and survival status) were tested by univariate analyses and by estimating cumulative incidence. Cox proportional-hazards models were used to evaluate the effect of potential confounding variables, in addition to haplotypes, on overall survival, progression-free survival, relapse/progression and chronic GVHD. Because of missing dates for the onset of acute GVHD, a logistic regression was performed instead. The parameters considered were DPB1 matching, source of stem cells, year of treatment, type and stage of disease, patient age at transplantation, transplantation center, T-cell depletion, conditioning, cytomegalovirus (CMV) serological status and recipient/donor gender combination. Donor age was not analyzed because of missing data. HLA–DPB1 matching was also investigated as an explanatory variable for overall survival (OS), for occurrence of acute GVHD and for relapse/progression. These analyses were generated with SPSS on a total of 211 recipients (Table S1) with parameters equally distributed across groups except for recipient/donor gender (Table S2). The median waiting time was 116 days between donor search and transplantation with no significant difference among recipients according to haplotype groups.

Results

HLA haplotypes determination

A total of 420 distinct high-resolution HLA-A~B~DRB1 haplotypes were phased by segregation analysis in the 291 patients and HW equilibrium was not rejected at any locus. The most frequent haplotypes are listed in Table 1. None of them reached a frequency of 5% and only seven had a frequency >1%, with most haplotypes observed just twice or once in the cohort (Table S3). The three loci were not significantly associated to each other (p-value of 1 according to the likelihood-ratio test, no extreme value according to parametric resampling for global linkage disequilibrium). This was in agreement with the observation that only few haplotypes were in complete linkage across the three loci (Tables 1 and S3).

As family data are seldom available to confirm haplotype frequencies when samples are typed for HLA, we also estimated haplotype frequencies on our data by using the EM algorithm without accounting for phase information and we compared the results with those of our segregation analysis. It showed us that 265 haplotypes were simultaneously assigned by both approaches, while 155 were assigned by segregation analysis only and 258 were found by EM only. This represented a low concordance (39%) between the two approaches. In addition, the haplotypes assigned by both approaches sometimes exhibited frequency differences. As an example, haplotype A*02:01~B*08:01~DRB1*03:01 exhibits a frequency of 0.52% through segregation analysis, whereas it reaches an overestimated frequency of 1.53% through EM; this is because its alleles frequently occur simultaneously at the genotype level but are most often not linked on the same chromosome (Figure S2).

The HLA-A~B~DRB1 phased haplotypes in the 291 patients were ranked according to their frequencies estimated on the 6114 Swiss volunteer donors (Tables 1 and S3). Both sets of data were cross-tabulated in order to predict the most probable extended HLA haplotypes in Switzerland as listed in Table S4.

Predicting 10/10 matched unrelated donors searches

Using a logistic regression model with the sum of ranks as an explanatory variable (Fig. 1a), we could show a significant association to search outcome ($p = 2.25e{-}14$). We then investigated the best rank cut-off to predict

Table 1 Most frequent HLA-A~B~DRB1 phased haplotypes and linkage disequilibrium among allele pairs in the cohort of 291 patients

Haplotype	Freq.	Count	LD A-B	LD A-DRB1	LD B-DRB1	Rank SBSC	Rank NMDP EUR	Rank NMDP AFA	Rank NMDP API	Rank NMDP HIS
A*01:01~B*08:01~DRB1*03:01	0.0326	19	8.38	4.08	11.67	1	1	2	40	2
A*03:01~B*07:02~DRB1*15:01	0.0206	12	3.93	2.11	9.71	2	2	7	NA	3
A*02:01~B*07:02~DRB1*15:01	0.0172	10	0.31	0.73	9.71	5	4	41	615	9
A*03:01~B*35:01~DRB1*01:01	0.0137	8	4.31	2.96	8.38	4	8	149	75	23
A*29:02~B*44:03~DRB1*07:01	0.012	7	11.31	3.50	9.25	6	5	9	1257	1
A*01:01~B*57:01~DRB1*07:01	0.012	7	6.78	0.36	5.85	8	7	58	8	22
A*26:01~B*38:01~DRB1*13:01	0.0103	6	9.71	2.26	4.56	42	53	NA	NA	162
A*02:01~B*44:02~DRB1*04:01	0.0086	5	3.37	2.53	4.24	3	3	10	1292	46
A*30:01~B*13:02~DRB1*07:01	0.0086	5	11.98	3.98	7.17	10	10	178	4	16
A*24:02~B*07:02~DRB1*15:01	0.0086	5	1.14	0.67	9.71	13	13	337	183	72
A*02:01~B*51:01~DRB1*11:01	0.0086	5	2.38	0.08	3.47	15	28	122	496	27
A*01:01~B*08:01~DRB1*15:01	0.0086	5	8.38	−0.23	−0.17	25	25	544	NA	NA
A*24:02~B*08:01~DRB1*03:01	0.0086	5	0.34	0.63	11.67	47	34	1379	268	42
A*02:01~B*15:01~DRB1*04:01	0.0069	4	2.22	2.53	7.06	9	6	21	1419	39
A*02:01~B*18:01~DRB1*03:01	0.0069	4	0.22	0.17	2.82	68	71	80	NA	40
A*02:01~B*40:01~DRB1*13:02	0.0052	3	1.11	−0.35	3.16	7	9	81	NA	554
A*02:01~B*08:01~DRB1*03:01	0.0052	3	−1.86	0.17	11.67	17	11	20	NA	18
A*02:01~B*18:01~DRB1*11:04	0.0052	3	0.22	0.08	5.69	24	33	NA	NA	70
A*31:01~B*40:01~DRB1*04:04	0.0052	3	5.40	4.09	6.07	27	20	98	1264	137
A*02:01~B*51:01~DRB1*08:01	0.0052	3	2.38	0.55	1.78	29	73	245	NA	203
A*32:01~B*44:03~DRB1*07:01	0.0052	3	2.06	1.81	9.25	120	291	NA	236	234
A*01:01~B*15:17~DRB1*13:02	0.0052	3	4.04	2.70	8.41	210	327	NA	55	284
A*02:01~B*13:02~DRB1*13:01	0.0052	3	−0.04	0.09	1.15	333	NA	1131	NA	NA

LD: pairwise linkage disequilibrium as defined by standardized residuals; values ≥2 correspond to a significant association. Haplotypes in complete linkage (i.e. across the three loci) are shown in bold.

Rank of haplotypes estimated in 6114 volunteer donors from the Swiss registry (SBSC) and in four large groups (EUR: donors of European descent, AFA: donors of African descent, API: donors of Asian descent, HIS: donors of South American descent) of volunteer donors from the National Marrow Donor Program (NMDP) (ref. [34])

NA not available

search outcome with good sensibility and specificity, and the inspection of boxplots (Fig. 1b) suggested a sum of 1000, i.e. corresponding to patients carrying at least one very rare haplotype not seen in SBSC or carrying two infrequent haplotypes. To confirm this preliminary assessment, we ran an ROC curve analysis, which showed that a sum of 1000 was a good trade-off between true and false-positive rates (Figure S3). Using this cut-off value of 1000, we achieved a sensitivity of 0.71 and a specificity of 0.72. Most false positives (i.e. no 10/10 MUD found despite a sum <1000) were due to patients carrying rare allele(s) or unusual B~C or DRB1~DQB1 association(s), which significantly impaired the chances of finding a donor. By contrast, false negatives (i.e. at least one 10/10 MUD found despite a sum >1000) were often observed in patients carrying one very frequent haplotype besides the rare one, explaining why a donor could still be found in this specific constellation.

HLA haplotypes and clinical outcome following HSCT

The univariate analyses did not reveal a significant effect of haplotypes on overall survival (Table 2 and Fig. 2) or on other outcomes such as GVHD (Table 2). Moreover, cumulative incidence for relapse and GVHD was not different across haplotype groups (Figure S4). By contrast, better DPB1 matching was slightly, although significantly, associated to less acute GVHD ($p = 0.02$) and higher relapse/progression rate ($p = 0.03$), but not to better overall survival (Table 2).

Haplotypes were never significant in multivariate analyses. Recipients' age, stage of disease and transplantation center were significantly associated to survival (Tables 3 and S5), progression-free survival and relapse/progression (results not shown). In models inspected, older recipients and/or recipients with an advanced disease had a lower

chance of survival and progression-free survival but a higher risk of relapse/progression. Furthermore, in agreement with the results obtained in the univariate setting, DPB1 matching was not associated to better survival (Tables 3 and S5), but was a significant risk factor for relapse/progression and progression-free survival when not accounting for conditioning (results not shown). The other variables considered were never found to be significant, except type of disease in the models for relapse/progression because of a higher risk in recipients suffering from acute leukemia. Regarding chronic GVHD, no variable was significant. Logistic regression for acute GVHD revealed small but significant effects of T-cell depletion and source of cells and minor differences between acute leukemia and the other diseases (results not shown).

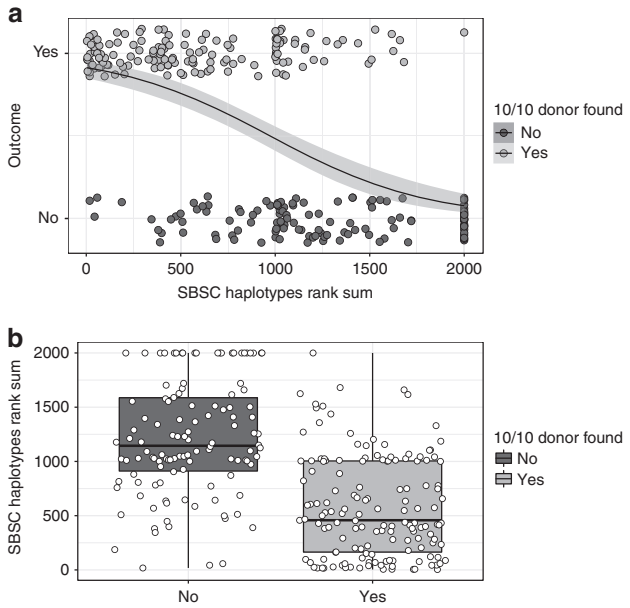

Fig. 1 Unrelated search outcome with the sum of haplotype ranks used as an explanatory variable, **a** logistic regression on the data represented by the black line with confidence interval in light gray, **b** box-and-whisker plots

Discussion

HLA haplotype determination is usually not based on family segregation, but relies on estimations performed with state-of-the-art EM algorithm implementations [35, 36]. Interestingly, the comparative analysis undertaken in this study showed that only a mere 39% assignation concordance was achieved between the real phased haplotypes in our cohort and a "blind" estimation with the EM algorithm. The discrepancies were mostly due to rare haplotypes and this problem has recently been discussed [20], but this also concerned the frequent haplotype

Table 2 Clinical outcome following HSCT: summary of univariate analyses

Outcome	Explanatory variable	Test	Statistic	df	*p*-Value
Overall survival (OS)	geno50	Log-Rank	1.72	2	0.42
	geno50.bis	Log-Rank	1.54	1	0.22
	geno20	Log-Rank	1.21	1	0.27
Relapse/progression	geno50	χ^2	2.32	2	0.31
	geno50.bis	Fisher	–	–	0.68 (two-sided)
	geno20	Fisher	–	–	>0.99 (two-sided)
aGVHD	geno50	χ^2	0.69	4	0.95
	geno50.bis	χ^2	0.5	2	0.78
	geno20	χ^2	0.63	2	0.73
cGVHD	geno50	χ^2	1.38	2	0.5
	geno50.bis	Fisher	–	–	0.83 (two-sided)
	geno20	Fisher	–	–	0.27 (two-sided)
Survival status at this date	geno50	χ^2	7.55	4	0.11
	geno50.bis	χ^2	6.59	2	0.04
	geno20	χ^2	1.08	2	0.58
Overall survival (OS)	DPB1 matching	Log-Rank	0.52	2	0.77
aGVHD	DPB1 matching	χ^2	7.64	2	0.02
Relapse/progression	DPB1 matching	χ^2	7.01	2	0.03

Geno50: recipients carrying 2, 1 or 0 common haplotypes with a frequency ≤ rank 50; geno50.bis: recipients carrying 2 common haplotypes with a frequency ≤ rank 50 versus recipients carrying any rare haplotypes with a frequency > rank 50; geno20: recipients carrying 0 or 1 rare haplotype versus recipients carrying 2 rare haplotypes with a frequency > rank 20

aGVHD acute graft vesus host disease, *cGVHD* chronic graft versus host disease, *df* degrees of freedom

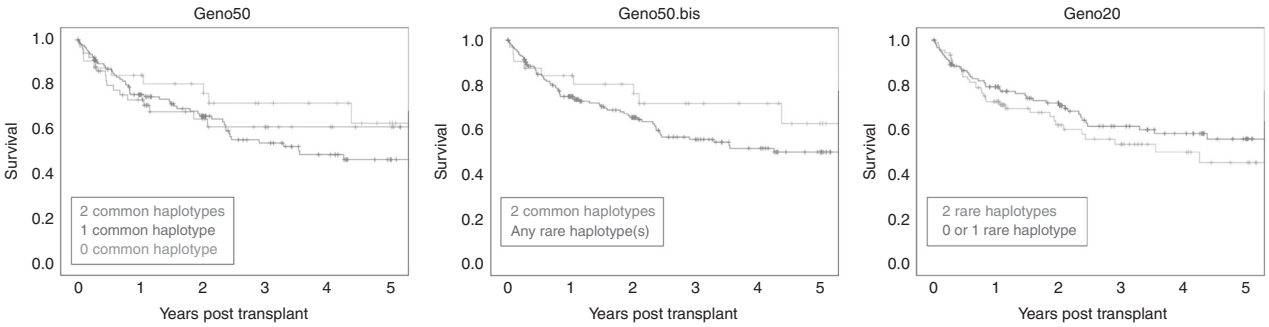

Fig. 2 Kaplan–Meier plots for the different genotype categories considered regarding haplotype frequency and HSCT outcome. Geno50: recipients carrying 2, 1 or 0 common haplotypes with a frequency ≤ rank 50; geno50.bis: recipients carrying 2 common haplotypes with a frequency ≤ rank 50 versus recipients carrying any rare haplotypes with a frequency > rank 50; geno20: recipients carrying 0 or 1 rare haplotype versus recipients carrying 2 rare haplotypes with a frequency > rank 20

Table 3 Cox regression model for overall survival with geno50

							95.0% CI for Exp(B)	
Explanatory variable	*B*	SE	Wald	df	Sig.	Exp(B)	Lower	Upper
TX center			22.715	3	0			
2	−0.592	0.344	2.965	1	0.085	0.553	0.282	1.085
3	1.017	0.31	10.754	1	0.001	2.765	1.506	5.078
4	0.789	0.69	1.309	1	0.253	2.201	0.57	8.509
Age			10.77	3	0.013			
Age (20–40)	1.02	0.543	3.528	1	0.06	2.773	0.957	8.041
Age (40–60)	1.03	0.473	4.735	1	0.03	2.801	1.108	7.081
Age (>60)	1.688	0.528	10.206	1	0.001	5.409	1.92	15.236
Disease stage			11.286	2	0.004			
Disease stage (intermediate)	0.268	0.285	0.888	1	0.346	1.308	0.748	2.285
Disease stage (advanced)	1.042	0.317	10.828	1	0.001	2.835	1.524	5.272
DPB1 MM			0.628	2	0.731			
DPB1 MM (1 MM)	0.084	0.319	0.07	1	0.791	1.088	0.583	2.032
DPB1 MM (2 MM)	−0.145	0.353	0.169	1	0.681	0.865	0.433	1.729
geno50			0.805	2	0.669			
geno50 (1 common haplotype)	0.291	0.377	0.595	1	0.44	1.337	0.639	2.799
geno50 (0 common haplotype)	0.376	0.428	0.771	1	0.38	1.456	0.63	3.368

Baseline for TX center = 1, for age = <20, for disease stage = early, for DPB1 MM = 0 MM, for geno50 = 2 common haplotypes. *MM* mismatch, *TX* transplant

A*02:01~B*08:01~DRB1*03:01. This illustrates the usefulness of family data for characterizing high-resolution multi-locus haplotypes [26] when sample sizes are not huge (meaning hundred thousand, or even millions of individuals). In addition, the use of next-generation sequencing (NGS) technologies is expected to increase the variability of high-resolution haplotypes. For instance, our data at third-field-level resolution includes A*02:01:01~B*08:01:01 ~DRB1*03:01:01 and A*02:01:01~B*08:01:02~DRB1* 03:01:01 haplotypes.

In the unrelated setting, the probability of finding a 10/10 matched donor is largely determined by haplotype frequencies, our analyses thus agree with previous publications [13, 27–29, 37–39]. Accurate prediction allows to define the optimal strategy to find the best suitable donor, whether a matched unrelated, mismatched unrelated or a haploidentical donor [40, 41].

By exploring several different models, we show that the presence of frequent haplotype(s) in the patients had no impact on HSCT outcome and thus partly differ from previous studies. Notably, Petersdorf et al. [15] found that haplotype matching between donors and recipients was associated with less grade 3–4 acute GVHD and with a higher risk of disease recurrence, but not with overall survival. In their publication, however, the haplotypes were determined in both the recipients and their MUDs with a

DNA microarray method that defines the physical linkage between HLA-A, B and DRB1 alleles, whereas our study determined them from family segregation analysis in recipients with indirect imputation of haplotype matching with their MUDs. A second difference between the two studies is the over representation in the former of the frequent haplotype A*01:01~B*08:01~DRB1*03:01, which is found in about 11% of the 246 donor–recipient pairs analyzed and which is always matched except in three cases, thus strongly contributing statistically to the results. In our cohort, this haplotype is present in only 6% of the 211 recipients. A second publication based on a large Japanese cohort [14] showed that among the three major conserved extended haplotypes (HP-P1, 2 and 3) found in Japanese, HP-P2 significantly reduced the risk of grade 2–4 acute GVHD, while HP-P3 tended to increase this risk, suggesting that conserved haplotypes may be beneficial or deleterious for the clinical outcome. In a third publication [13], an effect of haplotypes was seen only on the incidence of ≥grade 2 acute GVHD (survival or relapse were not significant) and only in the category "1 or 2 frequent haplotype(s) (FH)" (but not in the categories 1 FH and 2 FH taken individually). Moreover, the "1 of 2 FH" category was not associated with GVHD in univariate analysis and was only significant ($p = 0.026$) when adjusting for other factors in a multivariate model.

The controversial effect of haplotype matching might be related to methodological heterogeneity at defining common haplotypes, limitations to achieve sufficient statistical power and perhaps more likely to different impacts of individual haplotypes.

In 10/10 matched unrelated transplantation, the role of HLA and non-HLA genes is critical in the pathophysiology of GVHD and other outcomes [7, 10, 42]. HLA–DPA1 and DPB1, which are usually not in linkage disequilibrium with the other HLA genes [18, 19], are often not considered to be part of haplotypes and are not systematically characterized during donor selection [5], although the role of DPB1 has been well documented [43–45]. In our study, DPB1 mismatching is significantly associated with acute GHVD as an independent factor, albeit only in the univariate setting (Table 2). The lack of linkage between DPB1 and other loci observed even in high-frequency haplotypes might account for the difficulty in demonstrating an impact of these haplotypes on clinical outcome.

Furthermore, the presence of polymorphisms in the so called "identical HLA haplotypes" is demonstrated by routine typing with NGS, because this methodology reveals many new polymorphisms in exons not encoding the peptide binding region, as well as in non-coding regions. There is no reason to believe that the situation is different outside HLA genes across chromosome 6. Indeed, the impact of non-HLA genes (e.g. cytokines, cytokine receptors) and polymorphisms such as microsatellites and SNPs has been reported to influence clinical outcome in unrelated HSCT [6, 8, 9, 11, 12, 46]. Although SNPs can directly affect the sequences of immunogenic peptides leading to mHA disparities, they may also modify genes encoding proteins involved in the pathophysiology of GVHD, such as TNF alpha, complement, TAP1/2, LMP1/7.

One important limitation of our study resides in the small number of transplanted patients, although it is close to the number of patients included in the seminal study of Petersdorf et al., which did not allow testing the putative effect(s) of individual haplotypes. Moreover, because the Swiss population is highly heterogeneous, we may lack statistical power to detect modest effects of haplotype matching. Therefore, our results can perhaps not be generalized to countries characterized by lower levels of diversity.

In conclusion, our study establishes the list of haplotypes observed in Switzerland, which shows a high population diversity despite its small size [30]. As expected, we observe that frequent haplotypes are strongly associated with a high probability to find a 10/10 matched unrelated donor. On the other hand, our results support the hypothesis that haplotype matching does not impact the clinical outcome and that more evidences are needed to better understand the numerous factors involved.

Acknowledgements This study was supported by the Swiss National Science Foundation (grant #310030_173237/1), the Academic Society of the University of Geneva, the Ernest Boninchi Foundation (grant 2016 to AS-M), IRGHET (International Research Group on unrelated Hematopoietic stem cell Transplantation), the Dr Henri Dubois-Ferrière Dinu Lippatti foundation and the Philantropy Settlement.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Ciurea SO, Bayraktar UD. "No donor"? Consider a haploidentical transplant. Blood Rev. 2015;29:63–70.
2. McCurdy SR, Fuchs EJ. Selecting the best haploidentical donor. Semin Hematol. 2016;53:246–51.
3. Petersdorf EW. Mismatched unrelated donor transplantation. Semin Hematol. 2016;53:230–6.
4. Petersdorf EW. Optimal HLA matching in hematopoietic cell transplantation. Curr Opin Immunol. 2008;20:588–93.
5. Tiercy JM. How to select the best available related or unrelated donor of hematopoietic stem cells? Haematologica. 2016;101: 680–7.
6. Bettens F, Passweg J, Schanz U, Chalandon Y, Heim D, Gungor T, et al. Impact of HLA-DPB1 haplotypes on outcome of 10/10 matched unrelated hematopoietic stem cell donor transplants depends on MHC-linked microsatellite polymorphisms. Biol Blood Marrow Transplant. 2012;18:608–16.
7. Dickinson AM. Non-HLA genetics and predicting outcome in HSCT. Int J Immunogenet. 2008;35:375–80.
8. Ishikawa Y, Kashiwase K, Akaza T, Morishima Y, Inoko H, Sasazuki T, et al. Polymorphisms in TNFA and TNFR2 affect outcome of unrelated bone marrow transplantation. Bone Marrow Transplant. 2002;29:569–75.
9. Keen LJ, DeFor TE, Bidwell JL, Davies SM, Bradley BA, Hows JM. Interleukin-10 and tumor necrosis factor alpha region haplotypes predict transplant-related mortality after unrelated donor stem cell transplantation. Blood. 2004;103:3599–602.
10. Mullally A, Ritz J. Beyond HLA: the significance of genomic variation for allogeneic hematopoietic stem cell transplantation. Blood. 2007;109:1355–62.
11. Parmar S, Del Lima M, Zou Y, Patah PA, Liu P, Cano P, et al. Donor-recipient mismatches in MHC class I chain-related gene A in unrelated donor transplantation lead to increased incidence of acute graft-versus-host disease. Blood. 2009;114:2884–7.
12. Petersdorf EW, Malkki M, Horowitz MM, Spellman SR, Haagenson MD, Wang T. Mapping MHC haplotype effects in unrelated donor hematopoietic cell transplantation. Blood. 2013;121:1896–905.
13. Joris MM, Lankester AC, von dem Borne PA, Kuball J, Bierings M, Cornelissen JJ, et al. The impact of frequent HLA haplotypes in high linkage disequilibrium on donor search and clinical outcome after unrelated haematopoietic SCT. Bone Marrow Transplant. 2013;48:483–90.
14. Morishima S, Ogawa S, Matsubara A, Kawase T, Nannya Y, Kashiwase K, et al. Impact of highly conserved HLA haplotype on acute graft-versus-host disease. Blood. 2010;115:4664–70.
15. Petersdorf EW, Malkki M, Gooley TA, Martin PJ, Guo Z. MHC haplotype matching for unrelated hematopoietic cell transplantation. PLoS Med. 2007;4:e8.
16. Tay GK, Witt CS, Christiansen FT, Charron D, Baker D, Herrmann R, et al. Matching for MHC haplotypes results in improved survival following unrelated bone marrow transplantation. Bone Marrow Transplant. 1995;15:381–5.
17. Bugawan TL, Klitz W, Blair A, Erlich HA. High-resolution HLA class I typing in the CEPH families: analysis of linkage disequilibrium among HLA loci. Tissue Antigens. 2000;56:392–404.
18. Cullen M, Noble J, Erlich H, Thorpe K, Beck S, Klitz W, et al. Characterization of recombination in the HLA class II region. Am J Hum Genet. 1997;60:397–407.
19. Kauppi L, Stumpf MP, Jeffreys AJ. Localized breakdown in linkage disequilibrium does not always predict sperm crossover hot spots in the human MHC class II region. Genomics. 2005;86:13–24.
20. Louzoun Y, Alter I, Gragert L, Albrecht M, Maiers M. Modeling coverage gaps in haplotype frequencies via Bayesian inference to improve stem cell donor selection. Immunogenetics. 2018;70:279–92.
21. Nunes JM, Buhler S, Roessli D, Sanchez-Mazas A, collaboration H-n. The HLA-net GENE[RATE] pipeline for effective HLA data analysis and its application to 145 population samples from Europe and neighbouring areas. Tissue Antigens. 2014;83:307–23.
22. Riccio ME, Buhler S, Nunes JM, Vangenot C, Cuenod M, Currat M, et al. 16(th) IHIW: analysis of HLA population data, with updated results for 1996 to 2012 workshop data (AHPD project report). Int J Immunogenet. 2013;40:21–30.
23. Eberhard HP, Feldmann U, Bochtler W, Baier D, Rutt C, Schmidt AH, et al. Estimating unbiased haplotype frequencies from stem cell donor samples typed at heterogeneous resolutions: a practical study based on over 1 million German donors. Tissue Antigens. 2010;76:352–61.
24. Ikeda N, Kojima H, Nishikawa M, Hayashi K, Futagami T, Tsujino T, et al. Determination of HLA-A, -C, -B, -DRB1 allele and haplotype frequency in Japanese population based on family study. Tissue Antigens. 2015;85:252–9.
25. Pedron B, Yakouben K, Adjaoud D, Auvrignon A, Landman J, Guerin V, et al. Listing of common HLA alleles and haplotypes based on the study of 356 families residing in the Paris, France, area: implications for unrelated hematopoietic stem cell donor selection. Hum Immunol. 2005;66:721–31.
26. Testi M, Battarra M, Lucarelli G, Isgro A, Morrone A, Akinyanju O, et al. HLA-A-B-C-DRB1-DQB1 phased haplotypes in 124 Nigerian families indicate extreme HLA diversity and low linkage disequilibrium in Central-West Africa. Tissue Antigens. 2015;86:285–92.
27. Pedron B, Duval M, Elbou OM, Moskwa M, Jambou M, Vilmer E, et al. Common genomic HLA haplotypes contributing to successful donor search in unrelated hematopoietic transplantation. Bone Marrow Transplant. 2003;31:423–7.
28. Pedron B, Guerin-El Khourouj V, Dalle JH, Ouachee-Chardin M, Yakouben K, Corroyez F, et al. Contribution of HLA-A/B/C/DRB1/DQB1 common haplotypes to donor search outcome in unrelated hematopoietic stem cell transplantation. Biol Blood Marrow Transplant. 2011;17:1612–8.
29. Testi M, Andreani M, Locatelli F, Arcese W, Troiano M, Battarra M, et al. Influence of the HLA characteristics of Italian patients on donor search outcome in unrelated hematopoietic stem cell transplantation. Tissue Antigens. 2014;84:198–205.
30. Buhler S, Nunes JM, Nicoloso G, Tiercy JM, Sanchez-Mazas A. The heterogeneous HLA genetic makeup of the Swiss population. PLoS ONE. 2012;7:e41400.
31. Romon I, Montes C, Ligeiro D, Trindade H, Sanchez-Mazas A, Nunes JM, et al. Mapping the HLA diversity of the Iberian Peninsula. Hum Immunol. 2016;77:832–40.
32. Sanchez-Mazas A, Buhler S, Nunes JM. A new HLA map of Europe: regional genetic variation and its implication for peopling history, disease-association studies and tissue transplantation. Hum Hered. 2013;76:162–77.
33. Nunes JM. Using uniformat and gene[rate] to analyze data with ambiguities in population genetics. Evol Bioinform Online. 2016;11(Suppl 2):19–26.
34. Maiers M, Gragert L, Klitz W. High-resolution HLA alleles and haplotypes in the United States population. Hum Immunol. 2007;68:779–88.
35. Eberhard HP, Madbouly AS, Gourraud PA, Balere ML, Feldmann U, Gragert L, et al. Comparative validation of computer programs for haplotype frequency estimation from donor registry data. Tissue Antigens. 2013;82:93–105.
36. Sanchez-Mazas A, Vidan-Jeras B, Nunes JM, Fischer G, Little AM, Bekmane U, et al. Strategies to work with HLA data in human populations for histocompatibility, clinical transplantation, epidemiology and population genetics: HLA-NET methodological recommendations. Int J Immunogenet. 2012;39:459–72. quiz 73-6

37. Olson JA, Gibbens Y, Tram K, Kempenich J, Novakovich J, Buck K, et al. Identification of a 10/10 matched donor for patients with an uncommon haplotype is unlikely. HLA. 2017;89:77–81.
38. Tiercy JM. Unrelated hematopoietic stem cell donor matching probability and search algorithm. Bone Marrow Res. 2012;2012:695018.
39. Tiercy JM, Nicoloso G, Passweg J, Schanz U, Seger R, Chalandon Y, et al. The probability of identifying a 10/10 HLA allele-matched unrelated donor is highly predictable. Bone Marrow Transplant. 2007;40:515–22.
40. Bashey A, Zhang X, Sizemore CA, Manion K, Brown S, Holland HK, et al. T-cell-replete HLA-haploidentical hematopoietic transplantation for hematologic malignancies using post-transplantation cyclophosphamide results in outcomes equivalent to those of contemporaneous HLA-matched related and unrelated donor transplantation. J Clin Oncol. 2013;31:1310–6.
41. Luznik L, O'Donnell PV, Symons HJ, Chen AR, Leffell MS, Zahurak M, et al. HLA-haploidentical bone marrow transplantation for hematologic malignancies using nonmyeloablative conditioning and high-dose, posttransplantation cyclophosphamide. Biol Blood Marrow Transplant. 2008;14:641–50.
42. Petersdorf EW. Which factors influence the development of GVHD in HLA-matched or mismatched transplants? Best Pract Res Clin Haematol. 2017;30:333–5.
43. Fernandez-Vina MA, Klein JP, Haagenson M, Spellman SR, Anasetti C, Noreen H, et al. Multiple mismatches at the low expression HLA loci DP, DQ, and DRB3/4/5 associate with adverse outcomes in hematopoietic stem cell transplantation. Blood. 2013;121:4603–10.
44. Fleischhauer K, Shaw BE. HLA-DP in unrelated hematopoietic cell transplantation revisited: challenges and opportunities. Blood. 2017;130:1089–96.
45. Shaw BE, Gooley TA, Malkki M, Madrigal JA, Begovich AB, Horowitz MM, et al. The importance of HLA-DPB1 in unrelated donor hematopoietic cell transplantation. Blood. 2007;110:4560–6.
46. Nowak J, Nestorowicz K, Graczyk-Pol E, Mika-Witkowska R, Rogatko-Koros M, Jaskula E, et al. HLA-inferred extended haplotype disparity level is more relevant than the level of HLA mismatch alone for the patients survival and GvHD in T cell-replate hematopoietic stem cell transplantation from unrelated donor. Hum Immunol. 2018;79:403–12.

Leukemia (2020) 34:1422–1432
https://doi.org/10.1038/s41375-019-0654-y

ARTICLE

Stem cell transplantation

Genetic T-cell receptor diversity at 1 year following allogeneic hematopoietic stem cell transplantation

Stéphane Buhler[1] · Florence Bettens[1] · Carole Dantin[2] · Sylvie Ferrari-Lacraz[1] · Marc Ansari[3] · Anne-Claire Mamez[2] · Stavroula Masouridi-Levrat[2] · Yves Chalandon[2] · Jean Villard[1]

Received: 2 July 2019 / Revised: 23 October 2019 / Accepted: 13 November 2019 / Published online: 26 November 2019

Abstract

After allogeneic hematopoietic stem cell transplantation (HSCT), immune reconstitution leads to the development of a new T-cell repertoire. Immune reconstitution could be influenced by events such as conditioning, infections, and graft versus host disease (GVHD). Factors influencing the TCR diversity are of great interest to fine-tune the strategy for donor selection and to optimize standard of care. In this work, immunosequencing of the TCR CDR3β region was carried out in a large cohort of 116 full chimeric recipients at 1 year post-HSCT and their respective donors prior to transplantation. The repertoire overlap before and after HSCT was minimal, supporting de novo reconstitution as a primary pathway at any age. Among the parameters investigated, increased patient and/or donor age as well as positive CMV serologic status reinforced by CMV infection/reactivation were the ones significantly associated with a reduced diversity at 1 year post-HSCT. CMV-specific T-cell clones were shown to influence the clonality of the repertoire alongside the expansion of limited numbers of non-CMV T-cell populations. Interestingly, at the exception of CMV infection/reactivation, TCR diversity was not predictive of GVHD, relapse, death, or infections post-HSCT.

Introduction

Allogeneic hematopoietic stem cell transplantation (alloHSCT) is a standard treatment of hematologic disorders such as leukemia and primary immunodeficiencies. The polymorphism of human leukocyte antigen (HLA) genes is a major factor for the global outcome, notably for avoiding graft rejection and for minimizing the risk of relapse and development of severe graft versus host disease (GVHD) [1–3]. HLA is also instrumental in the processes leading to immune reconstitution, especially the T-cell repertoire diversity, considered a key factor for prognosis and long term survival. Homeostatic proliferation, which controls the size of the T-cell pool during the reconstitution of the T-cell compartment [4], is mediated by the presentation of antigens by HLA molecules to T cells. Direct presentation in case of HLA mismatch, indirect presentation of alloantigens (minor histocompatibility antigens), or pathogen peptides, (e.g., derived from latent viruses present in the recipient) drive a cytokine-mediated expansion of the mature donor T cells infused with the graft [5, 6]. This pathway, which is independent of the thymus, result in a skewed repertoire closely associated with infections and GVHD [7–9]. By contrast, de novo maturation of naive T cells derived from lymphoid precursors of the donor and selected in the thymus by the self-HLA molecules presenting self-peptides will restore a broad and fully responsive repertoire. HLA has been suggested to bias the T-cell receptor (TCR) V gene usage of maturing thymocytes [10]. The thymus is highly sensitive to conditioning, immunosuppression and GVHD

Supplementary information The online version of this article (https://doi.org/10.1038/s41375-019-0654-y) contains supplementary material, which is available to authorized users.

✉ Stéphane Buhler
stephane.buhler@hcuge.ch

✉ Jean Villard
jean.villard@hcuge.ch

[1] Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Department of Diagnostic, Geneva University Hospitals, Geneva, Switzerland

[2] Service of Haematology, Department of Oncology, Geneva University Hospitals and Faculty of Medicine, University of Geneva, Geneva, Switzerland

[3] Pediatric Department, Onco-Hematology Unit, Geneva University Hospitals and Cansearch research laboratory, Geneva Medical School, Geneva, Switzerland

and is significantly influenced by age. Age is associated with thymic involution, impairing the renewed thymopoiesis [11–15]. It has been suggested that thymus-dependent reconstitution takes months following HSCT, leaving the patients at risk of infections and other complications. However, it is not well understood whether T-cell diversity is directly predictive of these clinical events.

Pre-, peri-, and post-transplant factors influencing the T-cell repertoire have been investigated leading to an improvement of the standard of care [16–18]. Recent technological developments have also provided the capacity to analyse more precisely and in-depth the complex processes of reconstitution, offering new perspectives for the post-HSCT follow-up. On the one hand, system-level profiling is now possible with mass cytometry, allowing comprehensive monitoring of immune reconstitution [19, 20]. On the other hand, the central role and dynamics of T cells reconstitution can be investigated by high throughput immunosequencing [21–23], providing a more detailed snapshot of the TCR diversity compared with methodologies like spectratyping [24, 25]. Current protocols target the most variable complementary-determining region of the somatically rearranged TCR alpha and beta chains (i.e., CDR3 at the V-J or V-D-J junctions, respectively), allowing to track down single clones among thousand T cells sequenced in parallel. A limited number of studies have been published using this technology after alloHSCT. These studies are characterized by the inclusion of a small number of patients and a variety of factors which could influence the T-cell repertoire reconstitution. One study reported a reduction of the TCR diversity following GVHD or disease relapse [26]. In another study, a better recovery of the T-cell diversity was reported in recipients of double cord blood units compared with conventional or T depleted peripheral blood stem cells (PBSC) [27]. Acute GVHD and steroid treatment were associated with a high diversity in contrast to cytomegalovirus (CMV) and Epstein Barr virus infections [27]. A role of CMV reactivation generating holes in the underlying repertoire was recently proposed with a massive expansion of effector memory (T_{EM}) cells and contraction of naive subsets post-HSCT [28]. Antithymocyte therapy has been reported by some authors to lead to lower TCR diversity in recipients of CMV-positive donors, while other variables (e.g., GVHD and CMV reactivation) were not predictive of the diversity [29]. By contrast, a reduction of diversity was observed with GVHD but not with conditioning by another group [30].

In this work, immunosequencing of the TCR CDR3β region was carried out in a large cohort of 116 full chimeric recipients at 1 year post-HSCT and their respective HSC donors. The first aim was to investigate repertoire overlap before and after HSCT. Within the second aim, parameters possibly associated with the reconstitution of the repertoire at 1 year were analyzed. The third aim was to assess whether T-cell diversity could be used as a marker to predict posttransplant complications.

Material and methods

Patients and donors

Patients receiving an allogeneic HSC graft in Geneva between 2000 and 2016, who had full donor chimerism without sign of relapse at 1 year were selected. The cohort consisted of 116 donor/recipient pairs with characteristics described in Table 1. Post-HSCT complications were recorded during the first year and also after that period (Table 2). A standard immunosuppressive treatment consisting of methotrexate or mycophenolate combined with cyclosporine A or tacrolimus was provided to all patients. Partial T-cell depletion is sometimes included in the institution protocol and was considered as an explanatory variable in the statistical analyses [31]. This study was approved by the ethical committee of the institution (CER 06–208 and 08–208R).

Table 1 Patient's and transplant's characteristics

Parameter	$N = 116$ donor/recipient pairs
Recipient age in years	Median: 45.5, minimum: 0 (4,5 months), maximum: 66
Recipient sex	Female: 49, male: 67
Primary disease	AA: 6, ALL: 14, AML: 36, CLL: 3, CML: 16, inborn error: 1, lymphoma: 13, MDPS: 5, MDS: 13, MPS: 2, myeloma: 5, solid tumor: 2
Donor age in years	Median: 39.5, minimum: 1, maximum 65
Donor sex	Female: 50, male: 66
Type of donor	Unrelated: 42, related: 70, haploidentical: 4
Source of stem cells	BM: 16, CB: 2, PBSC: 98
Conditioning	MAC: 78 (including 33 with ATG), RIC: 38 (including 31 with ATG)
Ex vivo T-cell depletion	Fraction depleted: 0%: 48, 50%: 62, 100%: 6
Number of DLI	0: 90, 1: 6, 2: 6, 3: 7, 4: 4, 5: 2, 7: 1
HLA matching	10/10: 101, <10/10: 15
Sex matching	F/F: 23, F/M: 27, M/F: 26, M/M: 40
CMV serologic status	D−/R−: 31, D−/R+: 20, D+/R−: 16, D+/R+: 49

ATG antithymocyte globulin, *AA* aplastic anemia, *ALL* acute lymphoblastic leukemia, *AML* acute myeloid leukemia, *BM* bone marrow, *CB* cord blood, *CLL* chronic lymphocytic leukemia, *CML* chronic myeloid leukemia, *D* donor, *DLI* donor lymphocyte infusion, *F* female, *M* male, *MAC* myeloablative conditioning, *MDPS* combined myelodysplastic syndrome/myeloproliferative neoplasm, *MDS* myelodysplastic syndrome, *MPS* myeloproliferative syndrome, *PBSC* peripheral blood stem cells, *R* recipient, *RIC* reduced intensity conditioning

Table 2 Posttransplantation complications among 116 alloHSCT recipients

Type of complication	Within 1 year, *n* (%)	After 1 year, *n* (%)
Acute GVHD	68 (58.6%)	16 (13.8%)
	Grade 1: 23 (19.8%)	Grade 1: 7 (6%)
	Grade ≥ 2: 43 (37.1%)	Grade ≥ 2: 9 (7.8%)
	Nonavailable: 2 (1.7%)	–
Chronic GVHD	28 (24.1%)	20 (17.2%)
Relapse	24 (20.7%)	19 (16.4%)
Death	–	24 (20.7%)
Any infection	84 (72.4%)	34 (29.3%)
Viral infection (CMV and/or others)	63 (54.3%)	22 (20%)
Cytomegalovirus infection/reactivation	43 (37.1%)	7 (6%)
	CMV alone: 23 (19.8%)	CMV alone: 3 (2.6%)
	CMV and other virus: 20 (17.2%)	CMV and other virus: 4 (3.4%)
Bacterial infection	48 (41.4%)	19 (16.4%)
Fungal infection	10 (8.6%)	6 (5.2%)
Parasitic infection	2 (1.7%)	2 (1.7%)

HLA typing and chimerism

DNA extracted on an automatic system (QIAGEN GmbH, Hilden, Germany) from Ficoll purified peripheral blood mononuclear cells was obtained from donors shortly before transplantation (time point 1) and from their full chimeric recipients at 1 year post-HSCT (time point 2). HLA typing was performed by reverse PCR-sequence-specific oligonucleotide microbead arrays and high throughput sequencing (One Lambda, Canoga Park, CA, USA) or PCR-sequence-specific primers (Genovision, Milan Analytika AG, Switzerland). Chimerism was performed by STR analysis (AmpFlSTR® Identifiler, Invitrogen-Thermofisher, Waltham MA, USA), the detection sensitivity is <3% (i.e., patients were selected for the study if the donor chimerism was ≥97%).

Immunosequencing

High throughput sequencing of the TCR CDR3β region was carried out on Illumina MiSeq and HiSeq systems following a multiplex PCR (Adaptive Biotechnologies ImmunoSEQ© assay) [32, 33]. Donor/recipient pairs were analyzed at survey resolution targeting 120,000 T cells. Reproducibility and sampling performance was assessed in five selected pairs using triplicates analyzed at deep resolution (i.e., targeting 400,000 T cells). Productive rearrangements were retrieved from the ImmunoSEQ© analyzer platform and formatted for the analyses to be carried out in R with the help of GNU/Linux scripts. The counts of clones with CDR3 rearrangements sharing synonymous nucleotide substitutions (i.e., an identical amino acid sequence) were pooled.

Peptide binding predictions

The FASTA sequences of the 190 canonical proteins of CMV strain AD169 were downloaded from https://www.uniprot.org/. This is one reference proteome for CMV which includes manual annotations and Swiss-Prot reviewing. These data were then submitted to the NetMHCpan 4.0 server available at http://www.cbs.dtu.dk/services/NetMHCpan/ to perform HLA class I binding predictions [34]. The predictions were performed on all possible AD169 nonamers (i.e., 9mer represent the preferred length of peptides bound by HLA class I molecules) and with each of the 91 alleles observed in our patients at high resolution (i.e., second field [35]). The binding predictions performed on individual alleles were combined according to the HLA types in order to estimate the theoretical capacity of each patient to present CMV derived peptides. We considered the total number of strong and weak binders using affinity ranks <0.5% or >0.5% and <2%, respectively. Mean peptide binding affinity in nanoMolar units was also estimated in each patient over retained 9mer binders. Association with CMV infection/reactivation within the first year following transplantation was analysed through logistic regression.

CMV-specific clones

T cells were characterized as potentially specific for CMV if their CDR3β sequence matched one of the 164 clones identified by Emerson et al. as CMV-associated (i.e., clones with significant enrichment in CMV-positive subjects) and/or one of the 919 clones reported as CMV-reactive (i.e., clones able to recognize CMV antigens) [36]. Both sets are partially overlapping, the second one being particularly biased toward reactivity to 65 kDa phosphoprotein (pp65) and 55 kDa immediate-early protein 1. HLA restriction at low resolution (i.e., first field level [35]) is proposed for about half of these clones, we thus checked if we could find a concordance with HLA types in patients.

Statistical analyses

Clone's frequencies were estimated from the number of sequenced templates and used to describe the commonness/rareness of given T cells or of groups of T cells (e.g., CMV-specific clones). Overlap of the TCR repertoire before and after HSCT was estimated by Jaccard and Morisita's

similarity indices [37, 38]. These indices vary between 0 (no overlap) and 1 (complete overlap). The standard Jaccard index gauges similarity using the ratio of shared clones at both time points on the total number of clones, the standardized index additionally weights the similarity using clone's frequencies and the Morisita's index is based on statistical dispersion of clones assuming that diversity increases with sample size. The diversity of the TCR repertoire at both time points (i.e., pre and post-HSCT) was estimated by productive clonality. This index is robust to sampling and accounts both for the richness and eveness of the repertoire and ranges between 0 (polyclonal/diverse) and 1 (monoclonal/invariant). The difference of clonality between both time points was considered as a surrogate for repertoire reconstitution during the first year. Graphical inspection of the data and univariate linear modeling were used to identify independent variables significantly associated with reconstitution. Post hoc tests (Tukey HSD) were applied to determine group(s) with a different mean. Multivariable models were subsequently explored and validated with diagnostic plots of the residuals. We also assessed whether donor parameters were associated with the clonality of the repertoire infused with the graft. Furthermore, logistic regressions were performed considering clonality as a possible explanatory variable for posttransplant complications within or after the first year, based on the repertoire of the donor or recipient, respectively. Additional analyses are detailed in supplementary figure legends where appropriate. All the analyses were performed in R (version 3.5.0) using the packages ggplot2, reshape2, tcR, GGally, scales, and party.

Results

Repertoire overlap

Immunosequencing yielded a total of 3,582,584 private and public CDR3β clones with variable frequencies pre and post-HSCT, including 2,894,321 unique rearrangements (i.e., meaning that almost 81% of the rearrangements were private to a given donor/recipient pair). The repertoire overlap at both time points was low according to the three similarity indices, but slightly less in some pairs for Morisita's index (Fig. 1a, b and S1). Mean (±SD) values obtained were 0.017 (±0.014) for standard Jaccard, 0.026 (±0.025) for generalized Jaccard and 0.13 (±0.16) for Morisita's index, respectively, with significantly correlated distributions of values (Fig. 1S). Deep resolution sequencing provided a very good concordance with survey resolution and high reproductibility between replicates (data not shown). We also verified that the number of productive templates sequenced before and after alloHSCT was not correlated with the indices of similarity.

Parameters influencing repertoire reconstitution

Productive clonality exhibited a skewed pattern with a shift from a mostly polyclonal repertoire pre-HSCT to more predominant oligoclonal profiles post-HSCT (paired *t*-test $p < 2.2e{-}16$). Several parameters were significantly associated with this change of clonality according to the univariate analyses (Table S1). The reduction of diversity (i.e., equal to an increased clonality), was lower in young recipients (≤20 years old, Fig. 2a) or when the graft was infused from a young donor (≤30 years old, Fig. 2b), but only according to these discrete categories (Table S1). The CMV serologic status and CMV infection/reactivation (defined as CMV DNA in plasma above the limit of detection, currently 2.1E + 1 UI/ml, in patients with or without clinical symptoms) were both significantly associated with a reduced diversity post-HSCT (Fig. 2c). Specifically, the CMV-positive donor/recipient (D+/R+) group differed from the other groups according to Tukey HSD. Regarding CMV infection/reactivation, it significantly reduced the repertoire diversity in all groups (no observation in the D−/R− group). Moreover, the significant association was mainly driven by CMV, although a trend was observed for other viruses (Fig. S2). Conditioning, T-cell depletion, source of stem cells, donor lymphocyte infusions (DLI), acute and chronic GVHD, relapse and other infections (e.g., bacterial) were not significant. Age and CMV-related variables were included in multivariable models and were significant taken two-by-two (Table 3), but without interaction (data not shown). The strong influence of CMV serologic status and CMV infection/reactivation was also observed using conditional inference framework analysis (Fig. S3). Looking at the repertoire infused with the graft, a significant effect of donor's age and CMV status was found. Both variables were interacting with a shift toward oligoclonality detected in the group of CMV-positive donors aged >30 years (Table S2 and Fig. S7a).

Impact of cytomegalovirus on repertoire reconstitution

CMV-specific clones

A total of 1978 CDR3β clones were defined as specific for CMV in different donor/recipient pairs, including 299 with a unique rearrangement. Thus, many of these rearrangements were public, including one observed in 59 pairs. In addition, 10.5% of the CMV-specific clones were shared in the cohort before and after alloHSCT compared with only 1.6% of the non-CMV clones, representing a drastic increase of overlap (Fisher's exact test $p < 2.2e{-}16$). Among the 299 rearrangements, 167 were described with an HLA restriction [36]. Cross-tabulating this information with HLA types in patients,

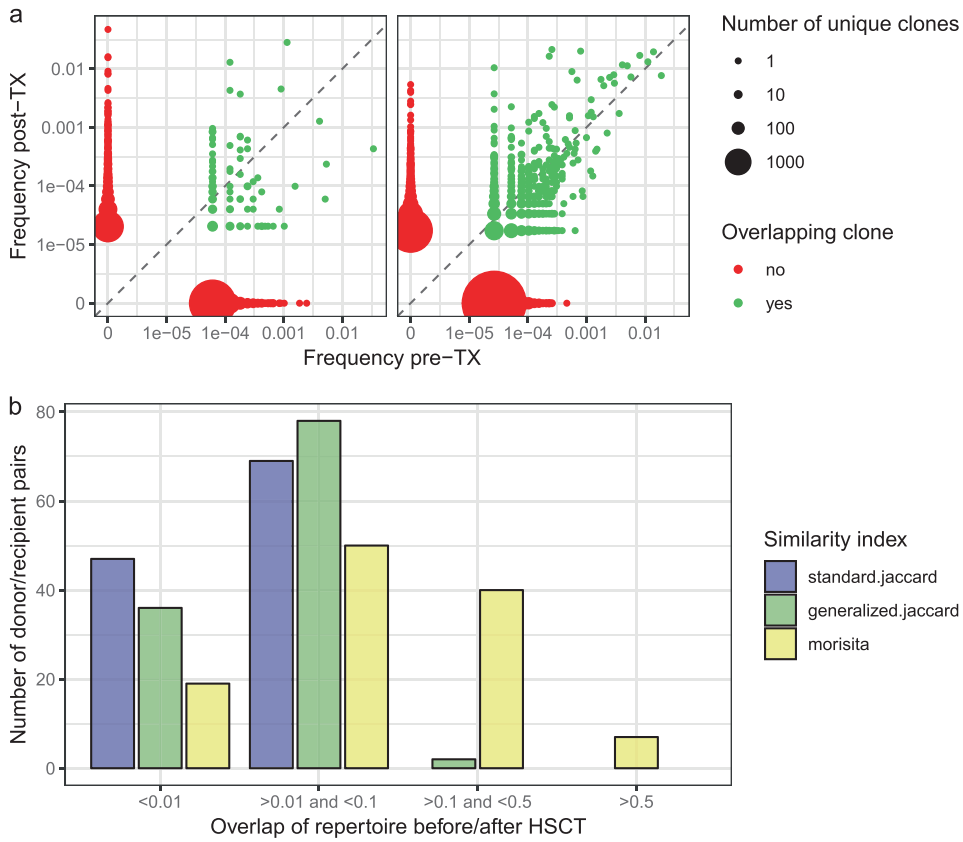

Fig. 1 Overlap of the TCR CDR3β repertoire in donors prior to transplantation and in their full chimeric recipients at 1 year post-HSCT. **a** Prototypical examples of overlap in two selected donor/recipient pairs. The clone's frequencies pre and post-TX (along *x* and *y* axis, respectively) and number of unique clones (dot size) are represented by scatter plots. Clones that are only observed at one time point are colored in red, while clones observed at both time points (i.e., overlapping clones) are colored in green. The values for standard Jaccard, standardized Jaccard, and Morisita's index are 0.021, 0.014, and 0.013 for pair #1 plotted on the left and 0.082, 0.117, and 0.264 for pair #2 plotted on the right, respectively. **b** Repertoire overlap according to the indices of similarity among the 116 donor/recipient pairs at several cutoff values. These indices vary between 0 (no overlap) and 1 (complete overlap) and are represented along the *x* axis. The number of donor/recipient pairs at each cutoff is plotted along the *y* axis. TX: allogeneic HSCT

we could observe the allele corresponding to the proposed restriction in 43% of the cases supporting the assignation of these clones as specific for CMV. Indeed, the number of HLA alleles, the flexibility of possible interactions at the TCR-peptide-HLA interface and the randomness of TCR rearrangements in distinct individuals makes this concordance highly improbable just by chance [39–41]. The number and cumulated frequency of CMV-specific clones are reported in Fig. 3, according to CMV serologic status and infection/reactivation post-HSCT. An increase above a frequency threshold of 1/1000 post-HSCT was only observed in the D+/R+ group when no infection/reactivation occurred. By contrast, all groups (except D−/R−, no observation) exhibited increased frequencies in case of infection/reactivation. We also investigated the relationship between the cumulative frequency of CMV-specific clones and clonality and found a significant correlation (Pearson's $r = 0.249$, $p = 0.007$; Spearman's rho $= 0.318$, $p = 0.0005$, Fig. S4a). This was confirmed by a powerful resampling approach (Fig. S4b). Finally, we demonstrated that although the frequency of CMV-specific clones increases according to CMV status, such clones were never dominant in the whole repertoire of the donors and patients (Fig. 4).

CMV peptide binding predictions and infection/reactivation

A total of 64,054 9mer were derived from the proteins of AD169 and submitted to NetMHCpan. Among this large number of possible 9mer, close to half of them (29,054, 45.4%) were predicted as binders to one or several HLA class I alleles (5.2 and 6.6 alleles on average for strong and weak 9mer binders, respectively). On average, HLA-A, B, and C alleles were predicted to bind strongly (or weakly) 809 (1587), 784 (1693), and 1358 (3134) peptides, respectively. Thus, considering HLA class I conjointly, every patient could theoretically cover a broad spectrum of CMV derived peptides (Fig. S5). Furthermore, no significant association with CMV infection/reactivation was observed (Fig. S6).

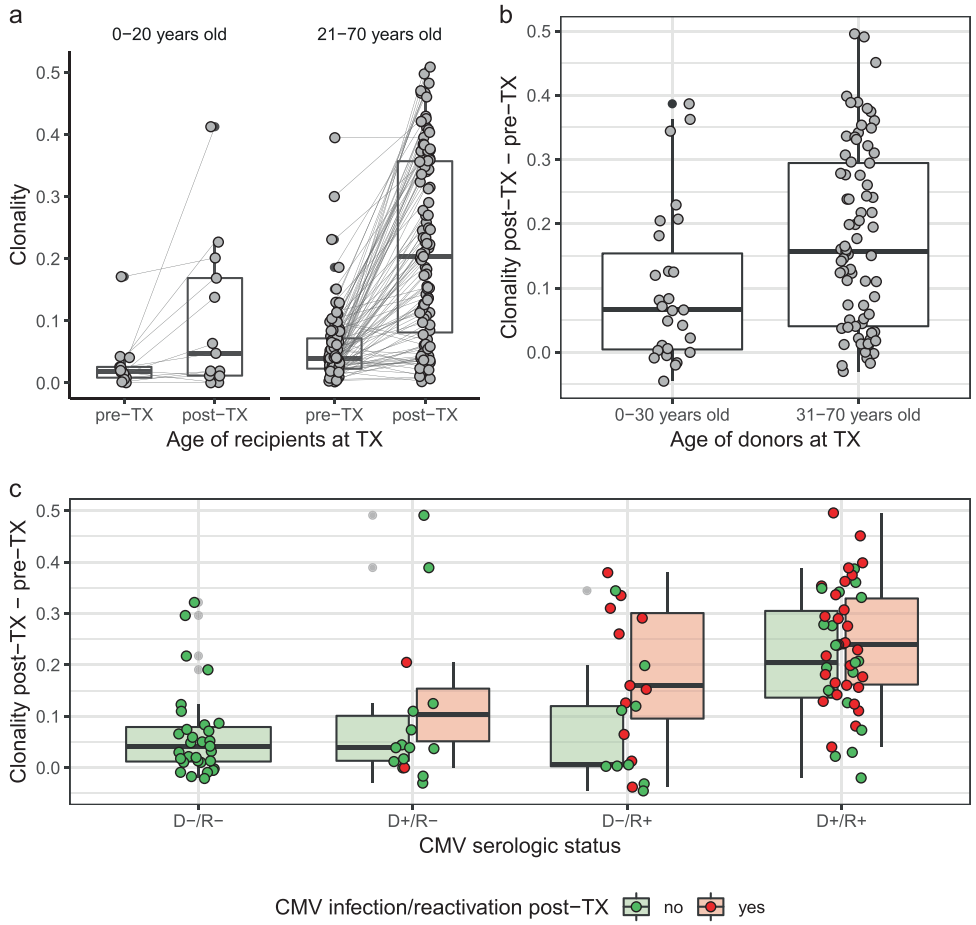

Fig. 2 Productive clonality of the TCR CDR3β repertoire in donors prior to transplantation and in their full chimeric recipients at 1 year post-HSCT. Clonality is shown along the *y* axis according to (**a**) age of recipients at transplantation ($n = 13$ and 103 for recipients ≤ or >20 years old, respectively), (**b**) age of donors at transplantation ($n = 27$ and 79 for donors ≤ or > 30 years old, respectively, the age of ten donors is unknown) and (**c**) cytomegalovirus (CMV) serologic status and occurrence of CMV infection/reactivation within the first year post-HSCT (see Table 1 for the numbers included in each category). **a** Clonality is plotted separately for both time points (i.e., pre and post-TX) with gray lines connecting donor/recipient pairs. Clonality varies between 0 (polyclonal/diverse repertoire) and 1 (monoclonal/invariant repertoire). **b**, **c** The variation of the repertoire diversity is plotted as clonality post-TX minus clonality pre-TX. Thus, a negative/positive value indicates an increased/decreased repertoire diversity at 1 year post-HSCT, respectively. D: donor CMV negative (−) or positive (+), R: recipient CMV negative (−) or positive (+), TX: allogeneic HSCT

Is T-cell diversity predictive of posttransplant complications?

Overall, clonality was not associated with an increased risk of clinical events within and after the first year post-alloHSCT (Tables S2 and S3), excepted that the risk of CMV infection/reactivation slightly increased with grafts from donors with a lower TCR diversity (Table S2 and Fig. S7b).

Discussion

Using immunosequencing we could describe the reconstitution of the T-cell repertoire diversity at 1 year after alloHSCT in a cohort of 116 full donor chimeric patients. Although the transplant infused usually includes a large amount of T cells, our data show that the repertoire 1 year after the procedure is very different, with only a few overlaps. This strongly suggests that a new repertoire can be reconstituted at any age through thymic dependent or independent pathways [9, 17, 18]. TCR monitoring by immunosequencing is very powerful, but it only provides a snapshot of the repertoire and the technology can be challenging in terms of analyses [37, 38]. However, with three indices of similarity going in the same direction and with a serie of controls performed at deep resolution to exclude sample size issues, we are pretty confident that our data are robust.

The diversity of the repertoire has been correlated to many clinical factors, however in this study, only three of them are significantly associated with clonality: age of the

Table 3 Multivariable analyses by linear regression for TCR reconstitution at 1 year

Response variable	Explanatory variables	Categories tested	Baseline	Coefficient estimate	*p* value
Clonality post-TX - pre-TX	(1) recipient age group	0–20, 21–70 years old	0–20 years old	0.08 (21–70 years old)	*0.026*
		D−/R−, D−/R+, D+/R−, D+/R+	D−/R−	0.04 (D+/R−)	0.287
	(2) CMV serologic status			0.08 (D−/R+)	*0.022*
				0.16 (D+/R+)	*1.34E-08*
	(1) donor age group	0–30, 31–70 years old	0–30 years old	0.06 (31–70 years old)	*0.023*
		D−/R−, D−/R+, D+/R−, D+/R+	D−/R−	0.04 (D+/R−)	0.266
	(2) CMV serologic status			0.1 (D−/R+)	*0.009*
				0.17 (D+/R+)	*4.61E-08*
	(1) CMV infection/reactivation	No, yes	No	0.06 (yes)	*0.043*
		D−/R−, D−/R+, D+/R−, D+/R+	D−/R−	0.02 (D+/R−)	0.5
	(2) CMV serologic status			0.05 (D−/R+)	0.209
				0.13 (D+/R+)	*4.82E-05*

The intercepts of the regressions are not shown. Significant *p*-values are shown in italic. There was no significant interaction between the above variables (i.e., models testing for interaction are not shown)

CMV cytomegalovirus, *D* donor, *R* recipient

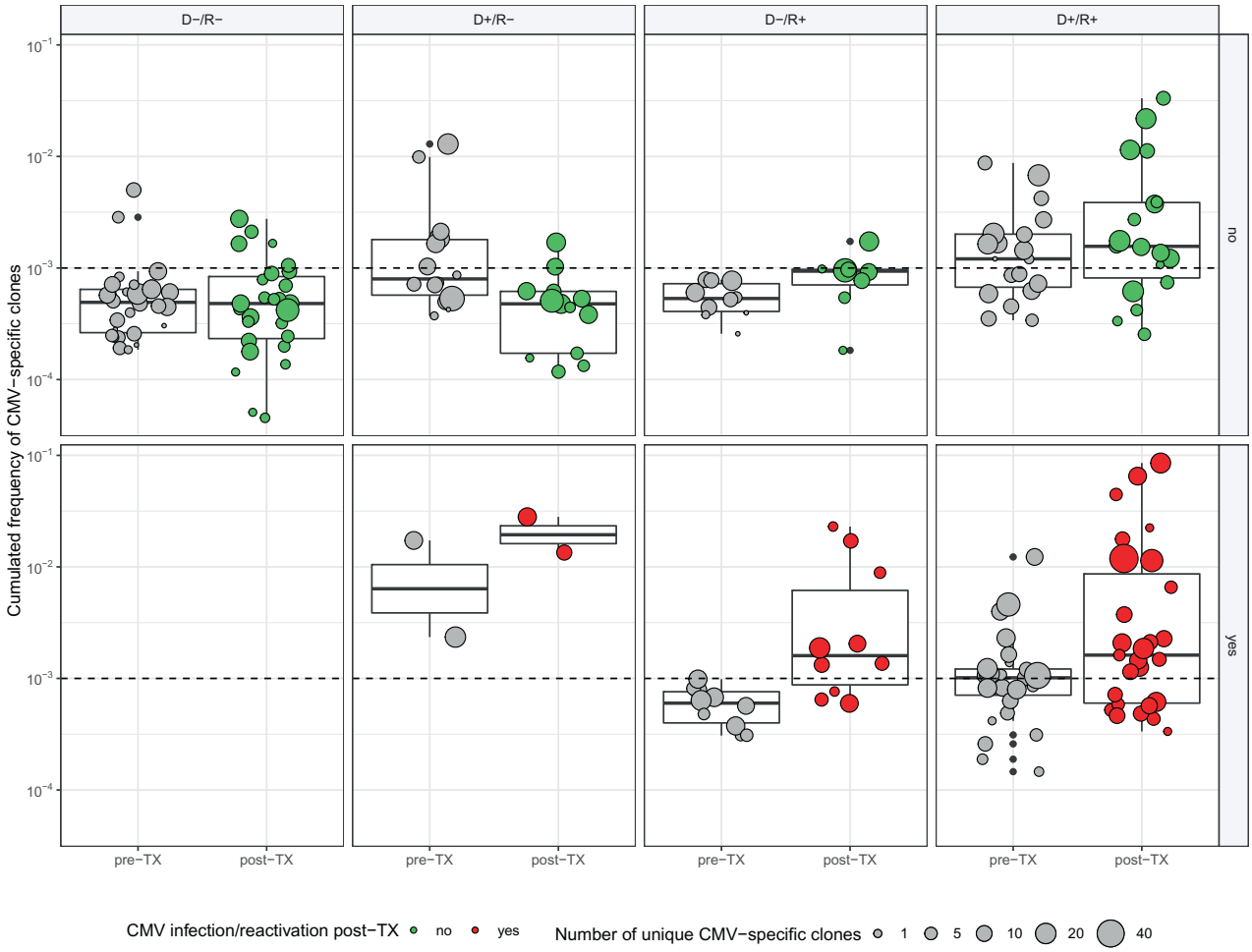

Fig. 3 Cumulative frequency (represented along the log scaled *y* axis) and number (represented by the dot size) of CMV-specific T-cell clones in donors prior to transplantation and in their full chimeric recipients at 1 year post-HSCT according to CMV serologic status (shown in the four panels along *x* axis) and CMV infection/reactivation within the first year post-HSCT (shown in the upper and lower panels). A frequency threshold of 1/1000 is indicated by the dotted line. D: donor CMV negative (−) or positive (+), R: recipient CMV negative (−) or positive (+), TX: allogeneic HSCT

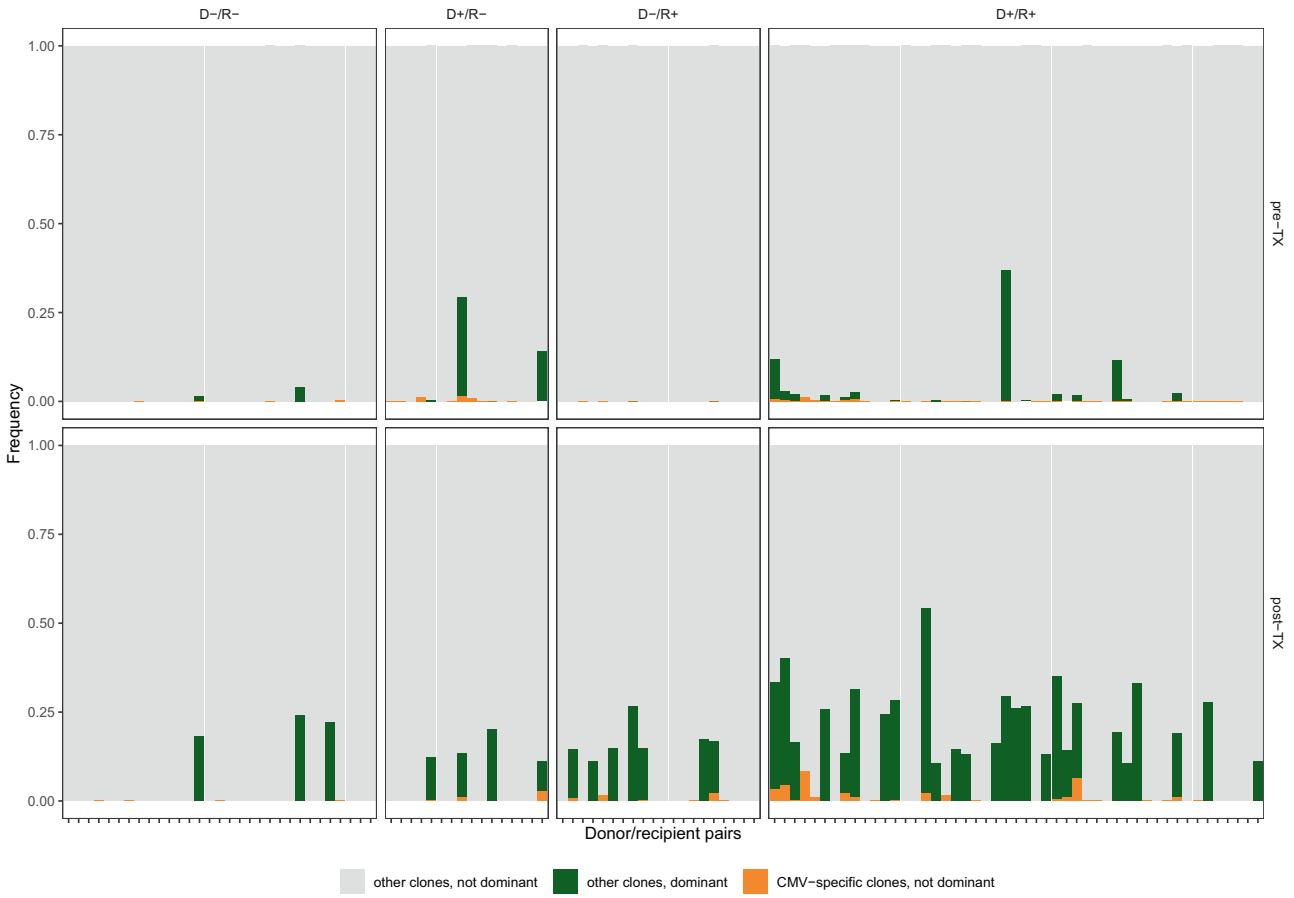

Fig. 4 Distribution of T-cell clones in donors prior to transplantation (upper panels) and in recipients at 1 year post-HSCT (lower panels). Clones are categorized according to two parameters, 1) their CMV specificity, nonspecific clones being classified as other and 2) their dominance (i.e., clones exhibiting a frequency of 10% or more in a given pair, either in the donor or in the recipient, are considered as dominant). This allowed to classify all sequenced clones into three subclasses shown in different colors on the plot. Of note, a fourth subclass consisting in dominant CMV-specific clones was not observed in any donor or patient. The *y* axis represents the cumulative frequencies of clones comprised within each subclass while donor/recipient pairs are listed along the *x* axis according to CMV serologic status (indicated on the top). The sizes of the faceted plots are proportional to the number of pairs comprised within each group

patient, age of the donor, and CMV. Interestingly, GVHD which has previously been associated with repertoire diversity in some reports [26, 30] was not significant in our cohort. Yet, a tendency was observed between acute GVHD and CMV infection/reactivation, although not significant, indicating that a more robust immunosuppressive treatment in patients with higher grade acute GVHD led to a higher prevalence of CMV infection/reactivation ($p = 0.063$). The association between acute GVHD and post-HSCT infections, especially CMV, is well described [8, 42, 43], the risk being dependent on the dose of steroid administration [44] and the GVHD grade [45].

An impact of the age of recipients on clonality was expected. Previous reports have already demonstrated similar findings which reflect the capacity of the thymus of the recipient to generate a more diverse repertoire at a younger age [12, 13, 15, 46, 47]. The contribution of young donors is also significant to explain a more diverse repertoire post-HSCT in our cohort. It has been previously proposed that advanced donor age could delay immune recovery [48], possibly because of the decreased frequency of prethymic T-cell progenitors within the graft. In addition, the repertoire diversity is higher with lymphoid progenitors from cord blood compared with bone marrow donors [49]. Similarly, the proliferation of peripheral T cells from young donors could be more efficient than from older donors. However we would expect a better overlap of the repertoire before and after T-cell reconstitution. We thus speculate that T-cell precursors coming from young donors associated with the thymic independent pathway [50, 51] best explain these results. The importance of donor age is of great interest in the strategy of finding the best donor for alloHSCT patients, e.g., what is the best option between an old haploidentical donor or an HLA matched or mismatched young unrelated donor? T-cell repertoire diversity at 1 year could be a parameter to take into consideration in this

context. The third factor significantly influencing the repertoire reconstitution is the CMV serologic status of the recipient and the donor reinforced by CMV infection/reactivation. In the D−/R− group, no primo infection was detected and the T-cell diversity is higher (less oligoclonal) compared with the other groups. The D+/R− status is also less associated with oligoclonality. In this later group, the rate of infection/reactivation is low, as expected, with only two reported cases. The reservoir of CMV is coming with the donor cells and could be controlled by the donor-specific T cells infused with the graft (i.e., only 6 of the 116 alloHSCT analyzed in this study were with 100% depleted T cells). The D−/R+ group is more prone to develop oligoclonality in case of infection/reactivation. This could be explained either by (a) residual recipient CMV-specific T cells (like tissue resident memory cells) which have escaped the conditioning regimen and proliferate, (b) by the primary response of donor-specific T cells, (c) by other crossreactive donor T cells or (d) by the proliferation of nonspecific T cells induced as a bystander effect of inflammation. The D+/R+ group, with and without CMV infection/reactivation, exhibits the strongest association with oligoclonality. In this group, the reservoir of latent CMV of the recipient and donor can stimulate donor CMV-specific T cells infused with the graft. Recipient CMV-positive serostatus has been proposed as the main factor determining CMV infection post-HSCT [52] and CMV-specific cytotoxic T cells were observed more frequently in D+/R+ grafts [53]. To validate these explanations, we analysed the presence of CMV-specific T cells and confirmed that their frequencies were mirroring the change of clonality observed in the different groups of our cohort depending on infection/reactivation. Interestingly, all patients have a good capacity to present CMV derived peptides according to their HLA class I restriction. Therefore, the presence of T cells able to respond broadly to CMV is expected in every individual. In agreement with our data, T-cell response to CMV is substantial and directed toward multiple antigens [54, 55]. CMV is the largest among known human viruses against which up to 10% of the CD4 and CD8 memory compartments are committed in the blood of seropositive subjects [55–57]. By comparing CMV-specific clones and non-CMV clones in our cohort, we could also show that the repertoire overlap is significantly driven by CMV but otherwise remains very limited before and after alloHSCT. Interestingly, in our cohort the combined frequencies of CMV-specific clones were well below 10% and none of them were found among clones described as dominant (Fig. 4). Dominant clones were observed more often in recipients with severe (i.e., ≥grade 2) acute GVHD ($p = 0.01$) and were especially overrepresented in the D+/R+ group ($p < 2.5e{-}10$), but were not significantly associated with CMV infection/reactivation. Some of these clones can be truly CMV specific and are not yet described in the current database. Alternatively these clones could be the signature of concomitant viral infections, indeed we found a trend, although not significant, of the impact of other infections on clonality. Finally, the presence of oligoclonality due to the presence of non-CMV dominant clones could be explained by the homeostatic proliferation of nonspecific T cells triggered by cytokines secreted during the symptomatic or asymptomatic anti-CMV immune response, especially in the D+/R+ group. We speculate that these nonspecific clones could be recent thymic emigrant or mature naive T cells which proliferate very efficiently [58], much more than memory anti-CMV-specific T-cell clones. Moreover, severe GVHD could also compound the reduction of diversity alongside CMV by promoting the expansion of alloreactive clones. A similar observation was made in a recent study based on TCRα diversity where dominant clones not observed before transplantation were found in patients suffering from CMV reactivation or extensive chronic GVHD [29]. In our study, in noninfected patients the repertoire is more diverse and a weak overlap is observed before and after transplantation. This suggests that at 1 year post-HSCT the repertoire is mainly composed of new T cells emerging from thymic dependent or independent pathways. This is also true but to a lesser degree in patients with CMV infection/reactivation as discussed above. Unfortunately, we do not have information about the naive or memory phenotype of these T cells, but previous reports have already established that at one year the T-cell populations are mainly naive (CD45RA+) [17, 18].

It is important to stress some limitations of our study. One drawback is that we could not differentiate CD4 and CD8 subpopulations because we did not have enough cells to perform cell sorting. For instance, a durable and significant imprint of CMV on T-cell reconstitution specifically in shaping the CD8+ memory T-cell compartment has been described [28, 59, 60].

The diversity of the T-cell repertoire at 1 year was interestingly not predictive of any subsequent clinical event such as infection (including CMV), GVHD, relapse, or mortality. This strongly suggests that clinical management remains the key factor to prevent and treat any event post transplantation.

The analysis of the donor T-cell repertoire was also instructive as it revealed that diversity was not significantly associated with GVHD, relapse or any infection except CMV infection/reactivation. This suggests that either grafts with reduced T cell diversity may lack CMV protective T cells and thus increase the risk of infection, or more probably that CMV-positive donors (i.e., with a more oligoclonal repertoire) are overrepresented in CMV-positive recipients, the more at risk of developing an

infection/reactivation. Our data also suggest that despite being significantly shifted in some donors and in many recipients, the repertoire infused with the graft, although transient, was diverse enough to afford protection in the early reconstitution phase and was then reconstituted to a sufficient extent at 1 year.

In conclusion, our study demonstrates the weak repertoire overlap before and at 1 year after alloHSCT. Age of the patient and the donor play a significant role. Reduced diversity at 1 year is mainly associated with CMV serostatus and infection/reactivation. It is to note that although CMV-specific clones are central to the observed shift of clonality they never predominate in the repertoire.

Acknowledgements This study was supported by the Swiss National Science Foundation (grant #310030_173237/1), the Academic Society of the University of Geneva, IRGHET (International Research Group on unrelated Hematopoietic stem cell Transplantation), the Dr Henri Dubois-Ferrière Dinu Lippatti foundation and the Philantropy Settlement. The authors are grateful to Lydie Brunet for her technical expertize and to the technicians of the LNRH for their most efficient support for HLA typing. We would also like to thanks Dr José Manuel Nunes (Biosc-iGE3) for his valuable suggestions on data analysis using R.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Petersdorf EW. Genetics of graft-versus-host disease: the major histocompatibility complex. Blood Rev. 2013;27:1–12.
2. Fleischhauer K, Beelen DW. HLA mismatching as a strategy to reduce relapse after alternative donor transplantation. Semin Hematol. 2016;53:57–64.
3. Fleischhauer K, Shaw BE. HLA-DP in unrelated hematopoietic cell transplantation revisited: challenges and opportunities. Blood. 2017;130:1089–96.
4. Jameson SC. Maintaining the norm: T-cell homeostasis. Nat Rev Immunol. 2002;2:547–56.
5. Degauque N, Brouard S, Soulillou JP. Cross-reactivity of TCR repertoire: current concepts, challenges, and implication for allotransplantation. Front Immunol. 2016;7:89.
6. DeWolf S, Sykes M. Alloimmune T cells in transplantation. J Clin Investig. 2017;127:2473–81.
7. Inman CF, Eldershaw SA, Croudace JE, Davies NJ, Sharma-Oates A, Rai T, et al. Unique features and clinical importance of acute alloreactive immune responses. JCI Insight 2018;3:10.
8. Phan TL, Pritchett JC, Leifer C, Zerr DM, Koelle DM, Di Luca D, et al. HHV-6B infection, T-cell reconstitution, and graft-vs-host disease after hematopoietic stem cell transplantation. Bone marrow Transplant. 2018;53:1508–17.
9. Mackall CL, Bare CV, Granger LA, Sharrow SO, Titus JA, Gress RE. Thymic-independent T cell regeneration occurs via antigen-driven expansion of peripheral T cells resulting in a repertoire that is limited in diversity and prone to skewing. J Immunol. 1996;156:4609–16.
10. Sharon E, Sibener LV, Battle A, Fraser HB, Garcia KC, Pritchard JK. Genetic variation in MHC proteins is associated with T cell receptor expression biases. Nature genetics. 2016;48:995–1002.
11. Chaudhry MS, Velardi E, Malard F, van den Brink MR. Immune reconstitution after allogeneic hematopoietic stem cell transplantation: Time To T Up the Thymus. J Immunol. 2017;198:40–6.
12. Clave E, Busson M, Douay C, Peffault de Latour R, Berrou J, Rabian C, et al. Acute graft-versus-host disease transiently impairs thymic output in young patients after allogeneic hematopoietic stem cell transplantation. Blood. 2009;113:6477–84.
13. Hakim FT, Memon SA, Cepeda R, Jones EC, Chow CK, Kasten-Sportes C, et al. Age-dependent incidence, time course, and consequences of thymic renewal in adults. J Clin Investig. 2005; 115:930–9.
14. Lynch HE, Goldberg GL, Chidgey A, Van den Brink MR, Boyd R, Sempowski GD. Thymic involution and immune reconstitution. Trends Immunol. 2009;30:366–73.
15. Weinberg K, Blazar BR, Wagner JE, Agura E, Hill BJ, Smogorzewska M, et al. Factors affecting thymic function after allogeneic hematopoietic stem cell transplantation. Blood. 2001; 97:1458–66.
16. Bosch M, Khan FM, Storek J. Immune reconstitution after hematopoietic cell transplantation. Curr Opin Hematol. 2012;19:324–35.
17. Ogonek J, Kralj Juric M, Ghimire S, Varanasi PR, Holler E, Greinix H, et al. Immune reconstitution after allogeneic hematopoietic stem cell transplantation. Front Immunol. 2016;7:507.
18. Storek J, Geddes M, Khan F, Huard B, Helg C, Chalandon Y, et al. Reconstitution of the immune system after hematopoietic stem cell transplantation in humans. Semin Immunopathol 2008; 30:425–37.
19. Stern L, McGuire H, Avdic S, Rizzetto S, Fazekas de St Groth B, Luciani F, et al. Mass cytometry for the assessment of immune reconstitution after hematopoietic stem cell transplantation. Front Immunol. 2018;9:1672.
20. Lakshmikanth T, Olin A, Chen Y, Mikes J, Fredlund E, Remberger M, et al. Mass cytometry and topological data analysis reveal immune parameters associated with complications after allogeneic stem cell transplantation. Cell Rep. 2017;20:2238–50.
21. Heather JM, Ismail M, Oakes T, Chain B. High-throughput sequencing of the T-cell receptor repertoire: pitfalls and opportunities. Brief Bioinform 2018;19:554–65.
22. Robins H. Immunosequencing: applications of immune repertoire deep sequencing. Curr Opin Immunol. 2013;25:646–52.
23. Six A, Mariotti-Ferrandiz ME, Chaara W, Magadan S, Pham HP, Lefranc MP, et al. The past, present, and future of immune repertoire biology—the rise of next-generation repertoire analysis. Front Immunol. 2013;4:413.
24. Cochet M, Pannetier C, Regnault A, Darche S, Leclerc C, Kourilsky P. Molecular detection and in vivo analysis of the specific T cell response to a protein antigen. Eur J Immunol. 1992;22:2639–47.
25. Gorski J, Yassai M, Zhu X, Kissela B, Kissella B, Keever C, et al. Circulating T cell repertoire complexity in normal individuals and bone marrow recipients analyzed by CDR3 size spectratyping. Correlation with immune status. J Immunol. 1994;152:5109–19.
26. Yew PY, Alachkar H, Yamaguchi R, Kiyotani K, Fang H, Yap KL, et al. Quantitative characterization of T-cell repertoire in allogeneic hematopoietic stem cell transplant recipients. Bone marrow Transplant. 2015;50:1227–34.
27. van Heijst JW, Ceberio I, Lipuma LB, Samilo DW, Wasilewski GD, Gonzales AM, et al. Quantitative assessment of T cell repertoire recovery after hematopoietic stem cell transplantation. Nat Med. 2013;19:372–7.
28. Suessmuth Y, Mukherjee R, Watkins B, Koura DT, Finstermeier K, Desmarais C, et al. CMV reactivation drives posttransplant T-

cell reconstitution and results in defects in the underlying TCRbeta repertoire. Blood. 2015;125:3835–50.
29. Link-Rachner CS, Eugster A, Rucker-Braun E, Heidenreich F, Oelschlagel U, Dahl A, et al. T cell receptor alpha repertoire of CD8+ T cells following allogeneic stem cell transplantation using next-generation sequencing. Haematologica. 2019;104:622–31.
30. Gkazi AS, Margetts BK, Attenborough T, Mhaldien L, Standing JF, Oakes T, et al. Clinical T cell receptor repertoire deep sequencing and analysis: an application to monitor immune reconstitution following cord blood transplantation. Front Immunol. 2018;9:2547.
31. Chalandon Y, Roosnek E, Mermillod B, Waelchli L, Helg C, Chapuis B. Can only partial T-cell depletion of the graft before hematopoietic stem cell transplantation mitigate graft-versus-host disease while preserving a graft-versus-leukemia reaction? A prospective phase II study. Biol Blood Marrow Transplant. 2006; 12:102–10.
32. Carlson CS, Emerson RO, Sherwood AM, Desmarais C, Chung MW, Parsons JM, et al. Using synthetic templates to design an unbiased multiplex PCR assay. Nat Commun. 2013;4:2680.
33. Robins HS, Campregher PV, Srivastava SK, Wacher A, Turtle CJ, Kahsai O, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. Blood. 2009;114:4099–107.
34. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: improved peptide-MHC Class I interaction predictions integrating eluted ligand and peptide binding affinity data. J Immunol. 2017;199:3360–8.
35. Nunes E, Heslop H, Fernandez-Vina M, Taves C, Wagenknecht DR, Eisenbrey AB, et al. Definitions of histocompatibility typing terms. Blood. 2011;118:e180–3.
36. Emerson RO, DeWitt WS, Vignali M, Gravley J, Hu JK, Osborne EJ, et al. Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire. Nat Genet 2017;49:659–65.
37. Laydon DJ, Bangham CR, Asquith B. Estimating T-cell repertoire diversity: limitations of classical estimators and a new approach. Philos Trans R Soc Lond B Biol Sci. 2015;370.
38. Rempala GA, Seweryn M. Methods for diversity and overlap analysis in T-cell receptor populations. J Math Biol. 2013;67:1339–68.
39. Bentzen AK, Such L, Jensen KK, Marquard AM, Jessen LE, Miller NJ, et al. T cell receptor fingerprinting enables in-depth characterization of the interactions governing recognition of peptide-MHC complexes. Nat Biotechnol. 2018;36:1191–6.
40. Birnbaum ME, Mendoza JL, Sethi DK, Dong S, Glanville J, Dobbins J, et al. Deconstructing the peptide-MHC specificity of T cell recognition. Cell 2014;157:1073–87.
41. Illing PT, Pymm P, Croft NP, Hilton HG, Jojic V, Han AS, et al. HLA-B57 micropolymorphism defines the sequence and conformational breadth of the immunopeptidome. Nat Commun. 2018;9:4693.
42. Ljungman P, Perez-Bercoff L, Jonsson J, Avetisyan G, Sparrelid E, Aschan J, et al. Risk factors for the development of cytomegalovirus disease after allogeneic stem cell transplantation. Haematologica. 2006;91:78–83.
43. Miller HK, Braun TM, Stillwell T, Harris AC, Choi S, Connelly J, et al. Infectious Risk after Allogeneic Hematopoietic Cell Transplantation Complicated by Acute Graft-versus-Host Disease. Biol Blood Marrow Transplant. 2017;23:522–8.
44. Hakki M, Riddell SR, Storek J, Carter RA, Stevens-Ayers T, Sudour P, et al. Immune reconstitution to cytomegalovirus after allogeneic hematopoietic stem cell transplantation: impact of host factors, drug therapy, and subclinical reactivation. Blood 2003;102:3060–7.
45. Cantoni N, Hirsch HH, Khanna N, Gerull S, Buser A, Bucher C, et al. Evidence for a bidirectional relationship between cytomegalovirus replication and acute graft-versus-host disease. Biol Blood Marrow Transplant. 2010;16:1309–14.
46. Storek J, Joseph A, Dawson MA, Douek DC, Storer B, Maloney DG. Factors influencing T-lymphopoiesis after allogeneic hematopoietic cell transplantation. Transplantation. 2002;73:1154–8.
47. Toubert A, Glauzy S, Douay C, Clave E. Thymus and immune reconstitution after allogeneic hematopoietic stem cell transplantation in humans: never say never again. Tissue Antigens. 2012;79:83–9.
48. Baron F, Storer B, Maris MB, Storek J, Piette F, Metcalf M, et al. Unrelated donor status and high donor age independently affect immunologic recovery after nonmyeloablative conditioning. Biol Blood Marrow Transplant. 2006;12:1176–87.
49. Talvensaari K, Clave E, Douay C, Rabian C, Garderet L, Busson M, et al. A broad T-cell repertoire diversity and an efficient thymic function indicate a favorable long-term immune reconstitution after cord blood stem cell transplantation. Blood. 2002;99:1458–64.
50. Fusco A, Panico L, Gorrese M, Bianchino G, Barone MV, Grieco V, et al. Molecular evidence for a thymus-independent partial T cell development in a FOXN1-/- athymic human fetus. PLoS ONE. 2013;8:e81786.
51. McClory S, Hughes T, Freud AG, Briercheck EL, Martin C, Trimboli AJ, et al. Evidence for a stepwise program of extrathymic T cell development within the human tonsil. J Clin Investig. 2012;122:1403–15.
52. Melendez-Munoz R, Marchalik R, Jerussi T, Dimitrova D, Nussenblatt V, Beri A, et al. Cytomegalovirus infection incidence and risk factors across diverse hematopoietic cell transplantation platforms using a standardized monitoring and treatment approach: a comprehensive evaluation from a single institution. Biol Blood Marrow Transplant. 2019;25:577–86.
53. Ogonek J, Varanasi P, Luther S, Schweier P, Kuhnau W, Gohring G, et al. Possible impact of cytomegalovirus-specific CD8(+) T cells on immune reconstitution and conversion to complete donor chimerism after allogeneic stem cell transplantation. Biol Blood Marrow Transplant. 2017;23:1046–53.
54. Elkington R, Walker S, Crough T, Menzies M, Tellam J, Bharadwaj M, et al. Ex vivo profiling of CD8+-T-cell responses to human cytomegalovirus reveals broad and multispecific reactivities in healthy virus carriers. J Virol. 2003;77:5226–40.
55. Sylwester AW, Mitchell BL, Edgar JB, Taormina C, Pelte C, Ruchti F, et al. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. J Exp Med. 2005;202:673–85.
56. Camargo JF, Komanduri KV. Emerging concepts in cytomegalovirus infection following hematopoietic stem cell transplantation. Hematol Oncol Stem Cell Ther. 2017;10:233–8.
57. Boeckh M, Geballe AP. Cytomegalovirus: pathogen, paradigm, and puzzle. J Clin Investig. 2011;121:1673–80.
58. Fink PJ, Hendricks DW. Post-thymic maturation: young T cells assert their individuality. Nat Rev Immunol. 2011;11:544–9.
59. Itzykson R, Robin M, Moins-Teisserenc H, Delord M, Busson M, Xhaard A, et al. Cytomegalovirus shapes long-term immune reconstitution after allogeneic stem cell transplantation. Haematologica 2015;100:114–23.
60. Lugthart G, van Ostaijen-Ten Dam MM, Jol-van der Zijde CM, van Holten TC, Kester MG, Heemskerk MH, et al. Early cytomegalovirus reactivation leaves a specific and dynamic imprint on the reconstituting T cell compartment long-term after hematopoietic stem cell transplantation. Biol Blood Marrow Transplant. 2014;20:655–61.

REGULAR ARTICLE

blood advances

Analysis of biological models to predict clinical outcomes based on HLA-DPB1 disparities in unrelated transplantation

Stéphane Buhler,[1] Helen Baldomero,[2] Sylvie Ferrari-Lacraz,[1] Anne-Claire Mamez,[3] Stavroula Masouridi-Levrat,[3] Dominik Heim,[2] Jörg Halter,[2] Gayathri Nair,[4] Yves Chalandon,[3] Urs Schanz,[4] Tayfun Güngör,[5] Grazia Nicoloso,[6] Jakob R. Passweg,[2] and Jean Villard,[1] on behalf of the Swiss Blood Stem Cell Transplantation Group

[1]Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Department of Diagnostic, Geneva University Hospitals, Geneva, Switzerland; [2]Division of Hematology, Basel University Hospital, Basel, Switzerland; [3]Division of Hematology, Department of Oncology, Geneva University Hospitals and Faculty of Medicine, University of Geneva, Geneva, Switzerland; [4]Division of Medical Oncology and Hematology, University Hospital of Zurich, Zurich, Switzerland; [5]Department of Stem Cell Transplantation, University Children's Hospital Zurich, Zurich, Switzerland; and [6]Swiss Blood Stem Cells Registry, Swiss Transfusion SRC, Bern, Switzerland

Key Points

- Risk of acute GVHD after unrelated HSCT is the highest when single HLA-DPB1 mismatches in the patient have a high cell surface expression.
- TCE nonpermissiveness and predicted indirectly recognizable HLA-II epitopes (PIRCHE II) are also predictive of acute GVHD.

HLA compatibility is a key factor for survival after unrelated hematopoietic stem cell transplantation (HSCT). HLA-A, -B, -C, -DRB1, and -DQB1 are usually matched between donor and recipient. By contrast, HLA-DPB1 mismatches are frequent, although it is feasible to optimize donor selection and DPB1 matching with prospective typing. Because classical DPB1 allele mismatches are often unavoidable, however, several biological models have been developed to predict the optimal DPB1 mismatch combination for less graft-versus-host disease (GVHD) and better overall survival. In 909 recipient/donor pairs, we analyzed the role of 3 biological models: T-cell epitopes (TCEs) based on the immunogenicity of DPB1, cell surface expression of DPB1 molecules based on a single-nucleotide polymorphism located in the 3′ untranslated region, and the Predicted Indirectly ReCognizable HLA Epitopes (PIRCHE) model based on the presentation of allogeneic peptides derived from mismatched HLA, compared with the classical allele mismatch. Matching for both DPB1 alleles remains the best option to prevent acute GVHD. In the situation of one DPB1 allele mismatch, the donor associated with the lowest acute GVHD risks is mismatched for an allele with a low expression profile in the recipient, followed by a permissive TCE3/4 mismatch and/or the absence of PIRCHE II potential against the recipient. In the context of 2 DPB1 mismatches, the same considerations apply for a permissive TCE3/4 mismatch and no PIRCHE II. By combining the biological models, the most favorable DPB1 constellation can be defined. This approach will help optimize donor selection and improve post-HSCT complications and patient prognosis.

Introduction

The significant role of HLA-DPB1 allele mismatches in hematopoietic stem cell transplantation (HSCT) has been well described.[1-3] Historically, HLA-DPB1 matching was not considered in the selection of unrelated donors, and mismatches were expected in up to 80% to 85% of otherwise matched unrelated transplant pairs (ie, recipient transplanted with 10/10 matched unrelated donors [MUDs]).[2,4] Nowadays, with the introduction of routine HLA-DPB1 typing of patients and upfront typing at donor recruitment, it is feasible

Submitted 9 December 2020; accepted 23 April 2021; published online 3 September 2021. DOI 10.1182/bloodadvances.2020003998.

Requests for access to deidentified HLA matching data for the 909 recipient/donor pairs can be submitted to the corresponding author. Clinical data for the 909 patients are recorded within the European Society for Blood and Marrow Transplantation ProMISe database. The data are not public but are searchable with a granted authorization to access the database.

The full-text version of this article contains a data supplement.

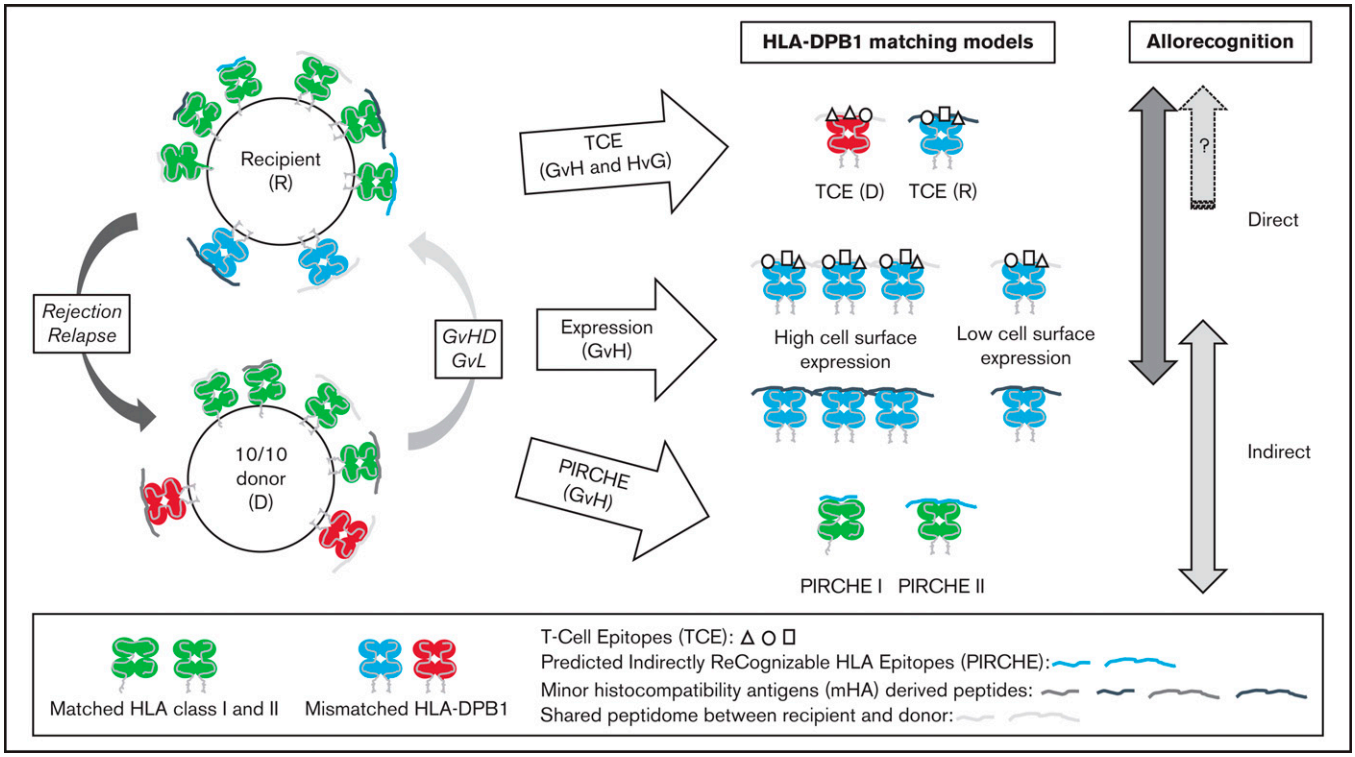

Figure 1. The biological models and their theoretical and relative contributions to T-cell alloreactivity against HLA-DPB1 incompatibilities. Schematic view of HSCT involving a recipient and a 10/10 MUD carrying at least one HLA-DPB1 allele mismatch. Matched HLA class I and II molecules are shown in green; the mismatched HLA-DPB1 molecules in donor and recipient are shown in red and blue, respectively. The donor and recipient differ genetically at genes encoding minor histocompatibility antigens that can be derived into antigenic peptides presented in the peptide-binding groove of HLA molecules (represented by different shades of gray). They can also present some peptides in common (ie, shared peptidome shown in the same gray tone). The 3 biological models for predicting permissive HLA-DPB1 mismatches are illustrated, each one differing in the type of T-cell allorecognition possibly involved (ie, direct and/or indirect) and in the vector of incompatibility (ie, in both directions for TCEs or GvH for expression and PIRCHE). Furthermore, the expression model is limited to situations in which the donor and recipient differ by only one HLA-DPB1 mismatch with an incompatibility in the GvH direction (ie, also including bidirectional mismatches). The expression model was not designed for double mismatches, which were considered unacceptable given the high risk of severe grade 3 to 4 aGVHD. The TCE model is determined by immunogenic variations called TCEs that are located within the peptide-binding region of HLA molecules. These epitopes can be directly recognized by alloreactive T cells. TCE permissiveness and differential immunopeptidome presentation have been proposed to be mediated by HLA-DM peptide editing, thus also potentially involving the indirect pathway of allorecognition in the TCE model. In the expression model, high cell surface expression of the mismatched HLA-DPB1 molecule in the recipient can favor two types of allorecognition, either direct, notably through the TCEs, or indirect, in cases of allopeptides derived from minor histocompatibility antigens and presented in the peptide-binding groove of the mismatched (but also of the matched) HLA-DPB1 molecule. Finally, indirect allorecognition is also expected in cases of allopeptides derived from mismatched HLA-DPB1 molecule(s) that can be presented in the peptide-binding groove of shared HLA class I or II molecules. This type of recognition is described by the PIRCHE model. D, donor; R, recipient.

to identify HLA 12/12 matched donors for many patients. However, HLA 12/12 matched donors cannot be identified for a substantial number of patients. In this context in which mismatches are often unavoidable, several alternative matching strategies have been sought to define some level of biological permissiveness and to improve clinical outcomes.

The first biological model is based on the immunogenicity of HLA-DPB1 molecules inferred from T-cell epitopes (TCEs) localized in the peptide-binding region. Three (TCE3)[5] or four (TCE4)[6] functional groups of alleles were defined, respectively, allowing classification of mismatches as permissive or nonpermissive. More attention has been given to TCE3 compared with TCE4,[7,8] until recently.[9]

In a different conceptual model, the risks of acute graft-versus-host disease (aGVHD) in transplants with a single HLA-DPB1 mismatch were associated with a single-nucleotide polymorphism (rs9277534) located in the 3′ untranslated region of the HLA-DPB1 regulatory region shown to significantly influence the quantity of cell surface expression of DPB1 molecules mediating allorecognition (ie, for proof-of-principle that expression is a functional determinant).[10] The expression model was not designed for double mismatches, which were considered unacceptable given the high risk of severe grade 3 to 4 aGVHD. The presence of a high expression allele in HLA-DPB1–matched transplantations was also linked to an increased risk of GVHD, probably because of enhanced donor recognition of minor histocompatibility antigens presented by the recipient. Interestingly, a strong correlation between the two rs9277534 variants and TCE grouping has been observed, suggesting that the immunogenicity of HLA-DPB1 molecules could be related, at least to some extent, to their expression levels.[11,12]

A third model relies on the indirect component of allorecognition and the presentation of allogeneic peptides derived from mismatched HLA molecules to T cells by a shared HLA molecule between the

Table 1. Patient, donor, and transplant characteristics

Characteristic	Value	Characteristic	Value	Characteristic	Value
Age of patients		**EBMT risk score**		**Source of stem cells**	
<20 y	123 (14%)	1-2	56 (6%)	Bone marrow	139 (15%)
20-40 y	153 (17%)	3-4	560 (62%)	Peripheral blood stem cells	769 (85%)
40-60 y	357 (39%)	5	293 (32%)	Cord blood	1 (0.1%)
60-70 y	248 (27%)	**Comorbid conditions***		**Total body irradiation**	
>70 y	28 (3%)	No	342 (38%)	No	628 (69%)
Year of treatment		Yes	382 (42%)	Yes	279 (31%)
2008	50 (5.5%)	Missing	185 (20%)	NA	2 (0.2%)
2009	58 (6%)	**Karnofsky performance scale index**		**Conditioning**	
2010	51 (6%)	90-100	688 (76%)	Myeloablative	465 (51%)
2011	76 (8.5%)	≤80	214 (23%)	Reduced intensity	443 (49%)
2012	80 (9%)	Missing	7 (1%)	NA	1 (0.1%)
2013	77 (8%)	**No. of allograft**		**Graft manipulation**	
2014	93 (10%)	First	873 (96%)	None	191 (21%)
2015	100 (11%)	Not first	36 (4%)	Serotherapy/other	641 (71%)
2016	96 (11%)	**Sex matching (D/R)**		In vitro T-cell depletion	77 (8%)
2017	115 (13%)	Male/male	436 (48%)	**HLA-DRB3/4/5 matching**	
2018	113 (12%)	Female/male	118 (13%)	Matched	845 (93%)
Type of diagnosis		Male/female	187 (21%)	1 mismatch DRB3	35 (4%)
Acute leukemia	506 (56%)	Female/female	168 (18%)	1 mismatch DRB4	28 (3%)
MDS/MPN	181 (20%)	**CMV serostatus matching (D/R)**		1 mismatch DRB3 and DRB4	1 (0.1%)
Lymphoid malignancy†	84 (9%)	Negative/negative	338 (37%)	**Transplant center‡**	
NMD	72 (8%)	Positive/negative	95 (11%)	202	327 (36%)
PCD	42 (5%)	Negative/positive	180 (20%)	208	237 (26%)
CML	23 (2%)	Positive/positive	287 (32%)	261	263 (29%)
ST	1 (0.1%)	**Age of donors, y**		334	82 (9%)
Status of disease		Median	31.3		
Early	445 (49%)	IQR	25.2-40.0		
Intermediate	279 (31%)				
Late	185 (20%)				

CML, chronic myeloid leukemia; CMV, cytomegalovirus; D, donor; IQR, interquartile range; MDS/MPN, myelodysplastic/myeloproliferative syndromes; NA, nonavailable; NMD, all nonmalignant disorders; PCD, plasma cell disorders; R, recipient; ST, solid tumor.

*Based on the hematopoietic cell transplantation–specific comorbidity index.

†Lymphoid malignancy regroups non-Hodgkin lymphoma, Hodgkin disease, and chronic lymphatic leukemia/prolymphocytic leukemia.

‡The transplant center code according to European Society for Blood and Marrow Transplantation (EBMT) is listed for the 4 allogeneic centers of Switzerland. All covariables tested in univariate analyses are shown here.

donor and recipient. An in silico approach was developed to predict the number of such peptides labeled PIRCHE (Predicted Indirectly ReCognizable HLA Epitopes).[13] The presence of PIRCHE was shown to correlate with clinical outcomes after HSCT.[14-16] The biological models and their theoretical and relative contributions to T-cell alloreactivity against HLA-DPB1 incompatibilities are schematically presented in Figure 1. Each biological model considers different aspects of allorecognition (ie, direct and/or indirect), although some information is shared across models. In addition, each model is not applicable to every situation of matching (ie, transplants with 1 or 2 mismatches and vector of incompatibility, as shown in Figure 1).

Based on direct, indirect, or both pathways of antigen recognition, it is unclear if the models can act synergistically or if each model is independent. In the current study, we analyzed clinical outcome in 909 recipient/donor pairs with a focus on the DPB1 matched/mismatched allele(s) stratified according to the classical model compared with the TCE, expression, or PIRCHE models.

Based on the results, we propose an algorithm for the selection of unrelated donors with lowest aGVHD risks that includes every model depending on the DPB1 matched/mismatched allele(s) constellation. Our data could be relevant to further refine the donor search and also to help in the strategy of exploiting the HLA-DPB1 mismatch permissiveness in cellular immunotherapy.[17]

Materials and methods

Study design, patients, and HLA-DPB1 typing

The role of HLA-DPB1 matching in the Swiss cohort was analyzed retrospectively by considering all 10/10 matched allografts performed from 2008 to 2018. This comprised a total of 909 patients

from 4 transplant centers. It mainly consisted of first allografts performed with peripheral blood stem cells as treatment of hematologic malignancies. GVHD prophylaxis was by drugs, mainly cyclosporine with methotrexate or mycophenolate; 71% used serotherapy, mostly antithymocyte globulin added to the drug regimen. Eight percent used in vitro T-cell depletion by Alemtuzumab. No posttransplant cyclophosphamide was used in these patients. Table 1 provides details on patient, donor, and transplant characteristics.

In Switzerland, prospective HLA-DPB1 typing in transplant candidates and selected unrelated donors was introduced at the end of 2016 with the development of high-throughput sequencing. Prospective typing was also performed from 2012 onward for each patient, with several potential 10/10 MUDs identified in the Bone Marrow Donors Worldwide/World Marrow Donor Association database. At the time of the search request, 7.8% of patients in the cohort had only one potential 10/10 MUD identified in the Bone Marrow Donors Worldwide/World Marrow Donor Association database; 28.8% had between 1 and 5 MUDs; and 63.4% had >5 MUDs. For the purpose of the current study, retrospective typing was performed for the donor/recipient pairs not yet fully characterized by using reverse polymerase chain reaction sequence–specific oligonucleotide microbead arrays (One Lambda, Canoga Park, CA) and polymerase chain reaction sequence–specific primers (Genovision, Milan Analytika AG, Rheinfelden, Switzerland). Complementary matching at HLA-DRB3/4/5 was also available and was included as a covariable in the analyses.

This study was approved by the ethical committee of the canton of Geneva and the Geneva University Hospital (CER 06-208 and 08-208R) and was conducted in accordance with the Declaration of Helsinki.

HLA-DPB1 matching models

Several models were considered in this study. The classical approach of counting the number of HLA-DPB1 allele mismatches was first examined, either with or without taking into account the direction of the vector of incompatibility in case of transplants mismatched for one allele. In terms of models inferring biological permissiveness, permissive and nonpermissive TCE mismatches were defined for each pair in the cohort, as previously described for the TCE3 and TCE4 algorithms using Linux/Bash scripts.[5,6,18] The nonpermissive mismatches were then split into 2 subgroups (ie, graft-versus-host [GvH] and host-versus-graft [HvG] incompatibilities) or kept together for the analyses. The cell surface expression of HLA-DPB1 alleles in donors and recipients was inferred from the described linkage between exonic variation and the 3′ untranslated region rs9277534-G/A polymorphism.[10,19] This allowed the classification of the expression level (ie, respectively high for G-linked alleles and low for A-linked alleles) of all single HLA-DPB1 mismatches with a vector of incompatibility in the GvH direction. In this model, pairs defined by 2 mismatches or by 1 mismatch in the HvG direction only could not be classified and were excluded from the analyses (one-third of the cohort, n = 305). The numbers of PIRCHE derived from the recipient's mismatched HLA-DPB1 allele(s) and potentially presented in the GvH direction on shared HLA class I (PIRCHE I) or class II (PIRCHE II) molecules between donor and recipient were identified by using the PIRCHE Web tool (www.pirche.com, version 3.1.147) as described elsewhere.[13] Of note, only potential binders to HLA-DRB1/3/4/5 molecules were considered in this study because typing for HLA-DQA1 and HLA-DPA1 was not available to predict peptide-binding affinities to HLA-DQ and HLA-DP heterodimers. The distribution of HLA-DPB1 mismatches according to the different biological models is detailed in supplemental Table 1 and is presented for PIRCHE I and II in supplemental Figure 1.

Statistical analysis

The clinical end points considered were overall survival, transplant-related mortality (TRM), grade 2 to 4 aGVHD, grade 3 to 4 aGVHD, chronic GVHD, and relapse/progression (Rel/prog). Initial exploration of the data (ie, single and combined HLA-DPB1 matching models and all covariables listed in Table 1) was done by using univariate analysis (ie, Kaplan-Meier, log-rank test, and cumulative incidence with competing risks). The data were then fitted into Cox multivariable regression models to compare the hazard ratios (HRs) between appropriate HLA-DPB1 matching groups and for each outcome adjusted for relevant covariables. A customary model-building strategy was used in which all covariables somehow associated with outcome and significant in univariate analysis were entered into the model and nonsignificant covariables were eliminated in a backward stepwise model-building procedure. A P value <.05 was considered as significant.

Each of the nine HLA-DPB1 matching models considered in this study was analyzed individually for the different outcomes. The most relevant biological models (ie, with a significant HR retrieved for at least 1 subgroup of patients) were then combined two-by two into 5 models. The subgroups considered at this stage consisted of all possible pairwise combinations from both selected models (eg, TCE3.1 and expression) and were either the same as for the individual analyses (eg, permissive or nonpermissive for TCEs) or pooled subgroups (eg, recipient with a highly expressed mismatch, thus not accounting for the expression level of the mismatch in the donor in contrast to the analyses performed for expression alone). The choice of pooling categories was made firsthand on the basis of the results obtained in the individual matching models (ie, keeping the relevant information for further testing) but also to keep a meaningful number of patients within each subgroup.

Results

HLA-DPB1 matching for predicting HSCT outcomes

Univariate analyses revealed higher risks of aGVHD and lower incidence of relapse in several HLA-DPB1 mismatched groups compared with fully matched allografts. The other clinical end points were not associated with HLA-DPB1 (selected Kaplan-Meier plots are shown for grade 2-4 aGVHD and relapse in supplemental Figures 2-5). Multivariable analyses confirmed this profile (Table 2; supplemental Table 2). Compared with 12/12 transplants (or alternatively to the absence of PIRCHE), the risks of grade 2 to 4 aGVHD were significantly increased: (1) with the presence of one (if bidirectional) or two allele mismatches; (2) with the presence of at least one PIRCHE II; (3) with the presence of a highly expressed mismatched allele in recipient (not the donor); (4) with the presence of nonpermissive TCE3/TCE4 mismatches; and (5) with the presence of permissive TCE3 mismatches, the latest group with a P value of .05. The risks of grade 3 to 4 aGVHD were increased significantly for expression (ie, in pairs with a highly expressed mismatched allele in both the recipient and donor; P = .01), and similar, although not significant, results to grade 2 to 4 aGVHD were observed for most models. Interestingly, HLA-DRB3/4/5 matching

Table 2. Multivariable analyses for aGVHD grade 2 to 4 and Rel/prog and association with each HLA-DPB1 matching model

		aGVHD grade 2 to 4 (N = 860, events = 297)*					Rel/prog (N = 837, events = 302)*†				
HLA-DPB1 matching	**Categories**	**Events/n**	**HR**	**95% CI**		***P***	**Events/n**	**HR**	**95% CI**		***P***
Classical matching.1	Matched	67/250	1.00		–		96/237	1.00		–	
	1 mismatch	145/398	1.40	1.05	1.87	***.02***	145/393	0.81	0.62	1.05	.11
	2 mismatch	85/212	1.52	1.10	2.10	***.01***	61/207	0.62	0.45	0.86	***.004***
Classical matching.2	matched	67/250	1.00		–		96/237	1.00		–	
	1 mismatch bidirectional	101/254	1.60	1.17	2.19	***.003***	96/248	0.84	0.63	1.11	.22
	1 mismatch GvH	23/79	1.01	0.63	1.63	.96	23/76	0.66	0.42	1.05	.08
	1 mismatch HvG	21/65	1.20	0.73	1.95	.48	26/69	0.88	0.57	1.36	.56
	2 mismatch	85/212	1.52	1.10	2.10	***.01***	61/207	0.62	0.45	0.86	***.004***
TCE3.1	Matched	67/250	1.00		–		96/237	1.00		–	
	Nonpermissive	115/288	1.54	1.13	2.09	***.006***	91/284	0.68	0.51	0.91	***.01***
	Permissive	115/322	1.36	1.00	1.84	***.05***	115/316	0.80	0.61	1.05	.10
TCE3.2	Matched	67/250	1.00		–		96/237	1.00		–	
	Nonpermissive GvH	71/165	1.76	1.25	2.46	***.001***	50/163	0.67	0.48	0.95	***.02***
	Nonpermissive HvG	44/123	1.28	0.87	1.88	.21	41/121	0.70	0.48	1.01	.06
	Permissive	115/322	1.36	1.00	1.84	***.05***	115/316	0.80	0.61	1.05	.10
TCE4.1	Matched	67/250	1.00		–		96/237	1.00		–	
	Nonpermissive	156/403	1.48	1.11	1.97	***.008***	127/403	0.66	0.51	0.86	***.002***
	Permissive	74/207	1.37	0.99	1.91	.06	79/197	0.92	0.68	1.24	.58
TCE4.2	Matched	67/250	1.00		–		96/237	1.00		–	
	Nonpermissive GvH	87/228	1.45	1.05	1.99	***.02***	69/230	0.63	0.46	0.86	***.003***
	Nonpermissive HvG	69/175	1.52	1.08	2.13	***.02***	58/173	0.71	0.51	0.98	***.04***
	Permissive	74/207	1.37	0.99	1.91	.06	79/197	0.92	0.68	1.24	.58
Expression	Matched	67/250	1.00		–		96/237	1.00		–	
	R-high, D-high	24/47	2.84	1.76	4.58	***<.001***	13/44	0.74	0.41	1.33	.31
	R-high, D-low	47/121	1.56	1.07	2.27	***.02***	31/116	0.61	0.41	0.92	***.02***
	R-low, D-high	13/57	0.81	0.45	1.48	.50	33/59	1.11	0.75	1.66	.60
	R-low, D-low	40/106	1.38	0.93	2.05	.11	39/102	0.76	0.52	1.10	.15
PIRCHE I	0	137/438	1.00		–		161/424	1.00		–	
	1-3	116/297	1.26	0.98	1.62	.07	99/284	0.83	0.64	1.06	.13
	>3	44/125	1.17	0.83	1.66	.37	42/129	0.84	0.60	1.18	.32
PIRCHE II	0	110/389	1.00		–		151/378	1.00		–	
	1-10	97/250	1.51	1.15	1.99	***.003***	83/236	0.76	0.58	0.99	***.04***
	>10	90/221	1.46	1.09	1.95	***.01***	68/223	0.73	0.55	0.97	***.03***

.05 > $P \geq .01$ are shown in bold and italic; $P < 0.01$ are shown in bold, italic, and underlined. Significant covariables retained for aGVHD: HLA-DRB3/4/5 matching, graft manipulation, and transplant center.

Significant covariables retained for relapse/progression (Rel/prog): European Society for Blood and Marrow Transplantation risk score, graft manipulation, and transplant center. CI, confidence interval; D, donor; Events, number of events in the risk category for the specified outcome; n, number of patients in the risk category for the specified outcome; R, recipient.

*The number of patients/events for the regressions with expression is N = 579/191 and 558/212 for aGVHD ≥2 and Rel/prog, respectively.

†Patients with nonmalignant disorder are excluded from analyses on relapse.

was also a risk factor for aGVHD, and the risks were mainly driven by HLA-DRB3 mismatches. However, this observation relies on a very small number of patients (n = 35 and 28 for DRB3 and DRB4 mismatches, respectively, results not shown).

Mirroring closely the increased risk of aGVHD, a lower incidence of relapse/progression was observed for the same groups compared with 12/12 allografts, except for one allele mismatches, R-high/D-high mismatched pairs, and TCE3 permissive mismatches, which were not statistically different (Table 2).

Less convincing results were observed for the other outcomes, as defined by *P* values very close to .05. These results are presented in supplemental Table 2 and are not discussed further here.

Combined biological models of permissiveness

The most relevant biological models analyzed in the previous section (ie, PIRCHE II, expression, TCE3.1 and TCE4.1) were combined two-by-two to investigate potential additive or synergistic effects on

Table 3. Multivariable analyses for acute GVHD grade 2 to 4 and Rel/prog and association with combined biological models

		aGVHD grade 2 to 4 (N = 860, events = 297)*					Rel/prog (N = 837, events = 302)*†				
HLA-DPB1 matching	**Categories**	**Events/n**	**HR**	**95% CI**		**P**	**Events/n**	**HR**	**95% CI**		**P**
TCE3.1 and PIRCHE II	Matched	67/250	1.00		–		96/237	1.00		–	
	Nonpermissive, no PIRCHE II	13/51	0.85	0.47	1.53	.58	20/49	0.87	0.53	1.42	.57
	Nonpermissive, PIRCHE II	102/237	1.73	1.27	2.37	***<.001***	71/235	0.65	0.47	0.88	***.005***
	Permissive, no PIRCHE II	31/91	1.28	0.84	1.97	.25	36/95	0.88	0.59	1.29	.50
	Permissive, PIRCHE II	84/231	1.40	1.01	1.93	***.04***	79/221	0.77	0.57	1.03	.08
TCE4.1 and PIRCHE II	Matched	67/250	1.00		–		96/237	1.00		–	
	Nonpermissive, no PIRCHE II	25/91	0.93	0.59	1.48	.77	33/92	0.75	0.50	1.12	.15
	Nonpermissive, PIRCHE II	131/312	1.67	1.24	2.25	***.001***	94/311	0.64	0.48	0.85	***.002***
	Permissive, no PIRCHE II	19/51	1.48	0.89	2.48	.13	23/52	1.15	0.72	1.82	.56
	Permissive, PIRCHE II	55/156	1.35	0.94	1.93	.10	56/145	0.85	0.61	1.19	.34
Expression and PIRCHE II	Matched	67/250	1.00		–		96/237	1.00		–	
	R-high, no PIRCHE II‡	9/13	NI	NI	NI	NI	4/11	NI	NI	NI	NI
	R-high, PIRCHE II	62/155	1.72	1.21	2.44	***.002***	40/149	0.61	0.42	0.88	***.008***
	R-low, no PIRCHE II	10/42	0.86	0.44	1.67	.66	16/42	0.72	0.42	1.23	.23
	R-low, PIRCHE II	43/121	1.30	0.88	1.92	.19	56/119	0.95	0.68	1.32	.75
Expression and TCE3.1	Matched	67/250	1.00		–		96/237	1.00		–	
	R-high, nonpermissive	39/89	1.87	1.26	2.80	***.002***	26/86	0.73	0.47	1.13	.15
	R-high, permissive	32/79	1.79	1.17	2.73	***.008***	18/74	0.55	0.33	0.92	***.02***
	R-low, nonpermissive	11/36	1.10	0.58	2.10	.77	19/35	1.00	0.61	1.65	1.00
	R-low, permissive	42/127	1.21	0.82	1.79	.34	53/126	0.85	0.61	1.20	.36
Expression and TCE4.1	Matched	67/250	1.00		–		96/237	1.00		–	
	R-high, nonpermissive	43/98	1.88	1.28	2.77	***<.001***	29/95	0.71	0.47	1.09	.12
	R-high, permissive	28/70	1.76	1.13	2.75	***.01***	15/65	0.54	0.31	0.93	.03
	R-low, nonpermissive	27/89	1.11	0.71	1.74	.65	35/91	0.71	0.48	1.05	.08
	R-low, permissive	26/74	1.28	0.81	2.02	.29	37/70	1.16	0.79	1.70	.45

.05 > $P \geq .01$ are shown in bold and italic; $P < .01$ are shown in bold, italic, and underlined. Significant covariables retained for relapse/progression (Rel/prog): European Society for Blood and Marrow Transplantation risk score, graft manipulation, and transplant center. Significant covariables retained for aGVHD: HLA-DRB3/4/5 matching, graft manipulation, and transplant center. CI, confidence interval; D, donor; Events, number of events in the risk category for the specified outcome; n, number of patients in the risk category for the specified outcome; R, recipient.

*The number of patients/events for the regressions with expression is N = 579/191 and 558/212 for aGVHD ≥2 and Rel/prog, respectively.

†Patients with nonmalignant disorder are excluded from analyses on relapse.

‡Not interpretable (NI) because of the very small number of patients in this group.

the primary outcomes. Again, highly significant HRs were observed for grade 2 to 4 aGVHD and relapse but not for the other outcomes (Table 3; supplemental Table 3). Foremost, the presence of a highly expressed mismatched allele in the recipient was associated with an increased risk of aGVHD compared with 12/12 allografts, whereas no difference was observed for lower expressed HLA-DPB1 mismatches. This observation was consistent across all subgroups; that is, combined with permissive or nonpermissive TCE3/TCE4 mismatches or with the absence or presence of PIRCHE II. In other words, considering TCE or PIRCHE II was not informative for stratifying the risks of aGVHD within groups with different levels of expression. By contrast, the results for TCEs and PIRCHE II were less straightforward. The presence of PIRCHE II was associated with a 67% or 73% risk increase of aGVHD compared with 12/12 transplants when combined with nonpermissive TCE4 or TCE3 mismatches, respectively, and a slightly less increase in combination with TCE3 permissive mismatches (40%; $P = .04$) but not with TCE4 ($P = .1$). Moreover, the absence of PIRCHE II in the context of either permissive or nonpermissive TCE3/TCE4 mismatches was not significantly different from 12/12 transplants (pairwise log-rank tests, data not shown; supplemental Figure 4). Taken together, our results suggest that PIRCHE II could add information to the TCE model, at least in case of nonpermissiveness. The results for grade 3 to 4 aGVHD were very similar to those for grade 2 to 4 aGVHD, but significant differences were not retrieved because these events were much rarer in our cohort (supplemental Table 3).

Regarding relapse/progression, the risks were decreased compared with 12/12 transplants in the following 3 groups: nonpermissive TCE3/TCE4 mismatches combined with the presence of PIRCHE II, highly expressed mismatched allele in recipient and presence of PIRCHE II, or when combined with permissive TCE3/TCE4 mismatches.

Discussion

This retrospective study is in line with the growing amount of evidence pointing to HLA-DPB1 matching and its related biological models of permissiveness as important parameters to consider for

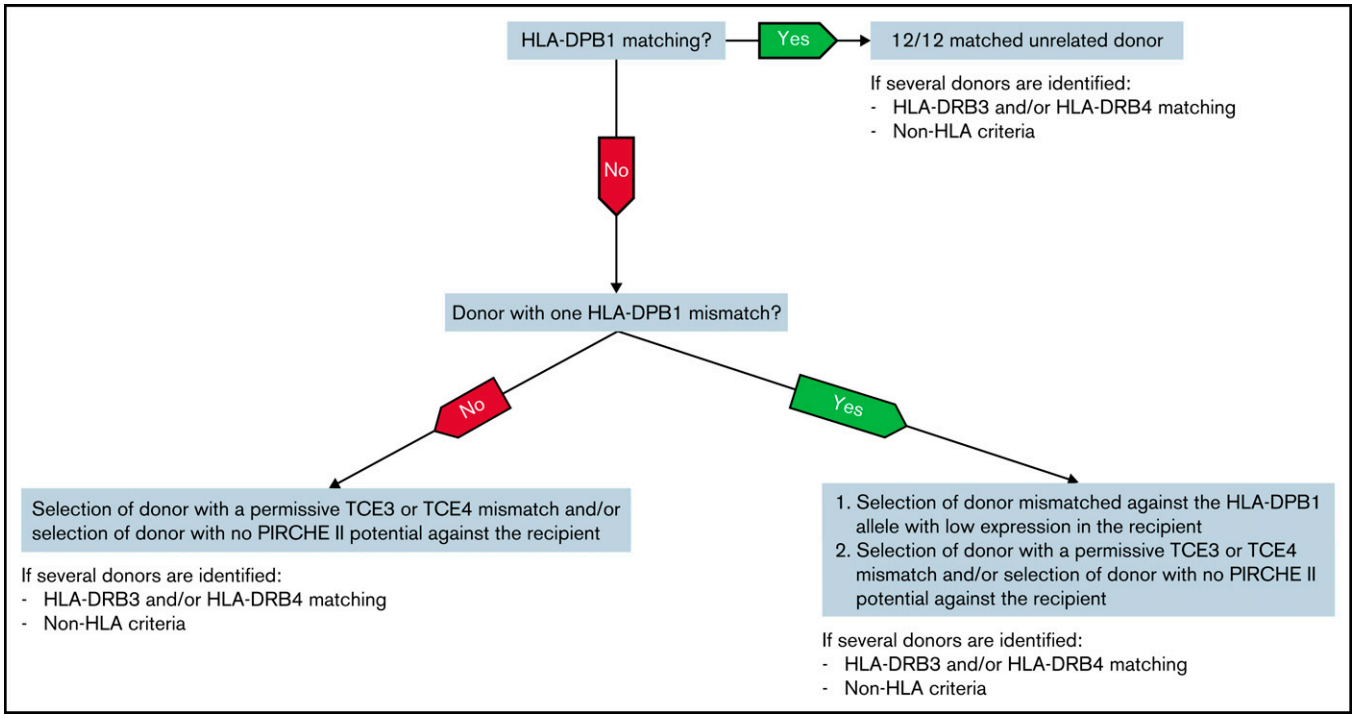

Figure 2. Decision tree for optimizing donor selection for HSCT candidates based on the results of this study and developed to reduce aGVHD risks in patients with a low risk of relapse. Classical HLA-DPB1 matching is the first parameter considered if a 10/10 donor can be identified. In case a 12/12 matched donor is not identified, the feasibility of selecting a donor with one or two biological permissive mismatches is explored. In cases when one or several donors carrying only one HLA-DPB1 allele mismatch are available, a low expression mismatched allele should be preferred in the recipient, followed by the feasibility to select a permissive TCE3 or TCE4 mismatch and/or to avoid PIRCHE II. The prioritization of expression over the 2 other biological models is based on the combined analyses in Table 3 showing that expression adds new information beyond TCEs and PIRCHE II. Of course, TCEs and PIRCHE II in addition to expression or as an alternative strategy are also informative. When allele matching cannot be achieved (ie, for selecting among unrelated donors with two HLA-DPB1 mismatches), the feasibility to select a permissive TCE3 or TCE4 mismatch and/or to avoid PIRCHE II applies. According to our data, TCEs and PIRCHE II should be considered equally during donor selection. The role of HLA-DRB3/4 matching and additional non-HLA factors should also be considered.

optimizing donor selection to improve prognosis. Among the several primary clinical end points examined here, aGVHD and relapse were shown to be significantly influenced by HLA-DPB1 matching. A balance between deleterious and protective effects (ie, higher risk of aGVHD vs lower incidence of relapse) has been proposed to explain why DPB1 allele mismatches have usually not been associated with a significant difference in survival,[3,9] with few exceptions.[4,20]

Allowing for a more fine-tuned approach than just counting the number of mismatches, the biological models examined here were also informative regarding the risks of aGVHD and relapse, except for PIRCHE I. Interestingly, we found no significant differences for the other outcomes (overall survival, TRM, and chronic GVHD). In agreement with our results, a high cell surface expression of one mismatched allele in recipient has previously been associated with increased risks of aGVHD[9,10,21,22] and sometimes with a decreased incidence of relapse.[10] Furthermore, our results are consistent with previous findings that the level of expression of the patient's mismatched HLA-DPB1 allele correlates with outcome more so than the expression level of the donor's mismatched allele and follows a biological GvH recognition.[22] Differences between fully matched transplants, permissive and nonpermissive TCE mismatches were previously described, with some heterogeneity across studies, for survival and TRM,[4-8] aGVHD,[5,7,8,22,23] and relapse.[8,23,24] Some studies also did not observe any significant differences for TCE3 or TCE4 for these outcomes.[20,25] Our results are thus consistent regarding aGVHD and relapse; we do not retrieve a signal for the other clinical end points.

For the most part, our results on relapse accompanied the ones observed for aGVHD, although with minor differences regarding groups that were associated with a significant *P* value. We thus focus on aGVHD to discuss in more detail the specific contributions of each biological model. The concordance between expression and TCEs (ie, at classifying high-risk vs low-risk mismatches) in our cohort was 68%, a percentage similar to those already reported.[12,22] We detected more PIRCHE I and II with TCE nonpermissive GvH mismatches than with permissive and nonpermissive HvG mismatches, similar to a previous report by Thus et al.[15] In addition, the number of PIRCHE I and II was the largest in recipients with a highly expressed mismatched allele, especially when the donor-mismatched allele had a low expression (supplemental Figure 1). It is thus possible that PIRCHE acts as a partial surrogate for TCEs and expression regarding clinical outcomes. Our results suggest that the presence of at least one PIRCHE II was sufficient to affect significantly the risks of GVHD and that this was not driven by the number of potential binders (Tables 2 and 3); this theory remains to be formally investigated. The study of Thus et al[15]

reported a significant correlation between the presence of PIRCHE I and PIRCHE II and the increased incidence of aGVHD. However, this was observed in a much smaller group of patients (n = 88). Of interest, we observed a trend toward increased incidence of aGVHD with the presence of PIRCHE I early posttransplant, but the differences were not significant globally and at later stages (results not shown). The importance of PIRCHE I therefore warrants confirmation by independent studies and needs to be contrasted with the influence of PIRCHE II.

Recent comparative analyses using expression and TCEs,[9,12,21,22] on the one hand, or TCEs and PIRCHE,[15] on the other hand, have proposed that their combination provides more information than either model alone and helps to better stratify the risks. This is because each model emphasizes only part of the complex mechanisms of T-cell alloreactivity against incompatible HLA-DPB1 molecules. Indeed, the biological models are analyzing distinct components of the alloreactive response but with some overlap (Figure 1). For instance, recent evidence suggests that TCE permissive mismatches present less divergent immunopeptidomes than nonpermissive mismatches, with a role for HLA-DM mediating peptide editing.[26] Components of both direct and indirect allorecognition are thus probably involved in the TCE model, similarly to the expression model, whereas PIRCHE is strictly restricted to the indirect pathway of recognition. Moreover, the different models do not cover the same breadth of information; for example, expression and PIRCHE focus specifically on incompatibilities in the GvH direction, and expression does not account for more than one HLA-DPB1 mismatch. Actually, their relative contributions were not straightforward to interpret when combined two-by-two in multivariable regressions. Each model seemed to play a significant role but more on an individual basis rather than by acting in concert. Previous studies reported similar complex relationships,[21,22] whereas others have suggested that the biological models could be prioritized according to their performance for different clinical outcomes (eg, expression with aGVHD or TCE4 with survival, respectively).[9] A recent analysis observed increased risks for GVHD/relapse-free survival, nonrelapse mortality, and aGVHD and reduced risks for relapse in grafts with two DPB1 mismatches combined to TCE3 nonpermissiveness in GvH direction.[23] An additive effect of expression combined with a TCE3 nonpermissive allotype was associated with aGVHD and relapse. In contrast to our data, the risks were assessed by combining both 10/10 and 9/10 grafts. A synthetic look at aGVHD in Table 3 found that expression adds new information beyond the information provided by PIRCHE II or by the TCE status, whereas the reverse situation is not true (ie, the TCE status or presence of PIRCHE II does not add new information once the level of expression is determined).

Also, once the TCE status is defined, the presence of PIRCHE II adds new information, whereas TCEs provide additional information only in the presence of PIRCHE II. This led us to propose a tentative algorithm for selecting unrelated donors with lowest aGVHD risks in Switzerland, as presented in Figure 2. Although the decision tree is mainly devised for patients with a low risk of relapse, adapting the selection to a donor carrying either two DPB1 mismatches, a mismatch against a highly expressed allele in the recipient, a nonpermissive TCE mismatch, or having a high PIRCHE II potential could also be beneficial in situations in which the risk of relapse/progression in the patient is preponderant over the risk of acute GVHD and should be minimized. For instance, permissive mismatches have been proposed to be associated with a limited alloreactivity sufficient to elicit GVL, thus maintaining treatment efficacy, without the deleterious effects of clinically uncontrollable GVHD.[17] However, we do not see any effect on overall survival of the different models, but reduction of the risks of aGVHD would be associated with less immunosuppression, better immune reconstitution, lower risk of concomitant infections, and other complications.[27-30] Thus, a flexible and individualized approach should always be considered along these general guidelines. Feasibility of TCE permissive matching for selecting prospectively among unrelated donors who were equally matched has previously been shown,[31] and several donor algorithms (eg, HapLogic,[32] OptiMatch [https://search.wmda.info/login]) already include information about TCE permissiveness in their match grade. Considerations about incorporating the other models in these algorithms should arise given our results and other recent studies addressing this issue.[9,22]

Our data also sustain a role of HLA-DRB3/4 matching for predicting the risks of aGVHD as an independent factor and as a covariable in the univariate and multivariable analyses, respectively.

Our study has some limitations despite the large size of the cohort. For instance, a few groups were relatively small for the analyses that combined the biological models two-by-two. We performed power and sample size calculations (not shown), and we used 2 different groupings for the TCE model with distinct sample sizes, which make us confident that our results are robust. For the same reason, we could not perform analyses combining the 3 models together. Also, grade 3 to 4 aGVHD is of main clinical importance due to its potential for severe sequelae, including death. However, this concerns rather rare events and because of this, we could not retrieve significant signals in our data, although the results resembled the significant observations made on grade 2 to 4 aGVHD. In addition, the tool for PIRCHE II is based on peptides derived from the mismatched DPB1 alleles presented on DRB1, DRB3, DRB4, and DRB5 and not on DQB1 or DPB1 because DRA is monomorphic and the polymorphic DQA1 and DPA1 genes were not typed to allow the software to make binding predictions for DQ and DP heterodimers. A transplant center effect was detected with several clinical outcomes. Although the centers share common practices (eg, donor selection process, HLA compatibility), clinical protocols can differ among them. The multivariable analyses were adjusted for this effect.

Although several studies have already compared the TCEs and the expression model[9,12,21-23] or PIRCHE and TCEs,[15] the current study, to the best of our knowledge, is the first to include the 3 models. In summary, HLA-DPB1 matching for both alleles (ie,12/12) remains the best option to prevent aGVHD. In the context of one HLA-DPB1 allele match, the donor with the lowest aGVHD risk would be mismatched to an allele with low cell surface expression in the recipient, followed by a permissive TCE3 or TCE4 mismatch and/or by the absence of PIRCHE II potential against the recipient. In the context of 2 DPB1 allele mismatches, the donor with the lowest aGVHD risk would have a permissive TCE3 or TCE4 mismatch and/or no PIRCHE II potential against the recipient. Because the 3 models are significant and exhibit complex relationships, this should be confirmed by independent studies, at least for the PIRCHE model, which has not been extensively tested in large cohorts. Donor selection includes immunogenetic and other factors, and the more we know, the more we will be able to personalize the choice for the best outcome. It would also be interesting to see if those

findings will still be observed with the growing use of cyclophosphamide posttransplantation as GVHD prophylaxis. Finally, T cells from a donor with a DPB1 mismatch could be an interesting tool in the future for cellular immunotherapy after HSCT.[17,33] The initial choice of such a donor could be part of a global strategy, including prevention of complications and potential for post-HSCT therapy to intervene in case of relapse.

Acknowledgments

The authors are grateful to the technicians of the National Reference Laboratory for Histocompatibility (LNRH) for their most efficient support for HLA typing.

This study was supported by the Swiss National Science Foundation (grant #310030_173237/1), IRGHET (International Research Group on Unrelated Hematopoietic Stem Cell Transplantation), and the Dr. Henri Dubois-Ferrière Dinu Lippatti foundation.

Authorship

Contribution: J.V., S.F.-L., and S.B. designed the study; J.R.P., H.B., and S.B. performed statistical analysis; and S.B. and J.V. drafted the manuscript; and all authors assembled the data, critically reviewed and edited the manuscript, and approved the final version.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: S.B., 0000-0001-6675-5287; H.B., 0000-0002-6554-483X; S.F.-L., 0000-0002-2183-3349; J.H., 0000-0003-4304-1956; Y.C., 0000-0001-9341-8104; T.G., 0000-0002-3261-1186; J.R.P., 0000-0001-7092-3351; J.V., 0000-0003-2667-7506.

Correspondence: Jean Villard, Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Geneva University Hospitals, Gabrielle-Perret-Gentil 4, 1211 Geneva 4, Switzerland; e-mail: jean.villard@hcuge.ch.

References

1. Fernández-Viña MA, Klein JP, Haagenson M, et al. Multiple mismatches at the low expression HLA loci DP, DQ, and DRB3/4/5 associate with adverse outcomes in hematopoietic stem cell transplantation. *Blood.* 2013;121(22):4603-4610.
2. Petersdorf EW, Gooley T, Malkki M, et al. The biological significance of HLA-DP gene variation in haematopoietic cell transplantation. *Br J Haematol.* 2001;112(4):988-994.
3. Shaw BE, Gooley TA, Malkki M, et al. The importance of HLA-DPB1 in unrelated donor hematopoietic cell transplantation. *Blood.* 2007;110(13):4560-4566.
4. Ludajic K, Balavarca Y, Bickeböller H, et al. Impact of HLA-DPB1 allelic and single amino acid mismatches on HSCT. *Br J Haematol.* 2008;142(3):436-443.
5. Zino E, Frumento G, Marktel S, et al. A T-cell epitope encoded by a subset of HLA-DPB1 alleles determines nonpermissive mismatches for hematologic stem cell transplantation. *Blood.* 2004;103(4):1417-1424.
6. Crocchiolo R, Zino E, Vago L, et al. Italian Bone Marrow Donor Registry. Nonpermissive HLA-DPB1 disparity is a significant independent risk factor for mortality after unrelated hematopoietic stem cell transplantation. *Blood.* 2009;114(7):1437-1444.
7. Fleischhauer K, Shaw BE, Gooley T, et al. International Histocompatibility Working Group in Hematopoietic Cell Transplantation. Effect of T-cell-epitope matching at HLA-DPB1 in recipients of unrelated-donor haemopoietic-cell transplantation: a retrospective study. *Lancet Oncol.* 2012;13(4):366-374.
8. Pidala J, Lee SJ, Ahn KW, et al. Nonpermissive HLA-DPB1 mismatch increases mortality after myeloablative unrelated allogeneic hematopoietic cell transplantation. *Blood.* 2014;124(16):2596-2606.
9. Lorentino F, Sacchi N, Oldani E, et al. Comparative evaluation of biological human leukocyte antigen DPB1 mismatch models for survival and graft-*versus*-host disease prediction after unrelated donor hematopoietic cell transplantation. *Haematologica.* 2020;105(4):e186-e189.
10. Petersdorf EW, Malkki M, O'Huigin C, et al. High HLA-DP expression and graft-versus-host disease. *N Engl J Med.* 2015;373(7):599-609.
11. Fleischhauer K. Immunogenetics of HLA-DP–a new view of permissible mismatches. *N Engl J Med.* 2015;373(7):669-672.
12. Meurer T, Arrieta-Bolaños E, Metzing M, et al. Dissecting genetic control of HLA-DPB1 expression and its relation to structural mismatch models in hematopoietic stem cell transplantation. *Front Immunol.* 2018;9:2236.
13. Geneugelijk K, Thus KA, Spierings E. Predicting alloreactivity in transplantation. *J Immunol Res.* 2014;2014:159479.
14. Geneugelijk K, Thus KA, van Deutekom HWM, et al. Exploratory study of predicted indirectly recognizable HLA epitopes in mismatched hematopoietic cell transplantations. *Front Immunol.* 2019;10:880.
15. Thus KA, Ruizendaal MT, de Hoop TA, et al. Refinement of the definition of permissible HLA-DPB1 mismatches with predicted indirectly recognizable HLA-DPB1 epitopes. *Biol Blood Marrow Transplant.* 2014;20(11):1705-1710.
16. Thus KA, Te Boome L, Kuball J, Spierings E. Indirectly recognized HLA-C mismatches and their potential role in transplant outcome. *Front Immunol.* 2014;5:210.
17. Fleischhauer K, Shaw BE. HLA-DP in unrelated hematopoietic cell transplantation revisited: challenges and opportunities. *Blood.* 2017;130(9):1089-1096.
18. Crivello P, Zito L, Sizzano F, et al. The impact of amino acid variability on alloreactivity defines a functional distance predictive of permissive HLA-DPB1 mismatches in hematopoietic stem cell transplantation. *Biol Blood Marrow Transplant.* 2015;21(2):233-241.

19. Schöne B, Bergmann S, Lang K, et al. Predicting an HLA-DPB1 expression marker based on standard DPB1 genotyping: linkage analysis of over 32 000 samples [published correction appears in *Hum Immunol*. 2020;81(10-11):661]. *Hum Immunol.* 2018;79(1):20-27.

20. Bettens F, Passweg J, Schanz U, et al. Impact of HLA-DPB1 haplotypes on outcome of 10/10 matched unrelated hematopoietic stem cell donor transplants depends on MHC-linked microsatellite polymorphisms. *Biol Blood Marrow Transplant.* 2012;18(4):608-616.

21. Morishima S, Shiina T, Suzuki S, et al. Japan Marrow Donor Program. Evolutionary basis of HLA-DPB1 alleles affects acute GVHD in unrelated donor stem cell transplantation. *Blood.* 2018;131(7):808-817.

22. Petersdorf EW, Bengtsson M, De Santis D, et al. International Histocompatibility Working Group in Hematopoietic-Cell Transplantation. Role of HLA-DP expression in graft-versus-host disease after unrelated donor transplantation. *J Clin Oncol.* 2020;38(24):2712-2718.

23. Mytilineos D, Tsamadou C, Neuchel C, et al. The human leukocyte antigen-DPB1 degree of compatibility is determined by its expression level and mismatch permissiveness: a German multicenter analysis. *Front Immunol.* 2021;11:614976.

24. Gagne K, Loiseau P, Dubois V, et al. Is there any impact of HLA-DPB1 disparity in 10/10 HLA-matched unrelated hematopoietic SCT? Results of a French multicentric retrospective study. *Bone Marrow Transplant.* 2015;50(2):232-236.

25. Loiseau P, Busson M, Balere ML, et al. HLA association with hematopoietic stem cell transplantation outcome: the number of mismatches at HLA-A, -B, -C, -DRB1, or -DQB1 is strongly associated with overall survival. *Biol Blood Marrow Transplant.* 2007;13(8):965-974.

26. Meurer T, Crivello P, Metzing MF, et al. Permissive HLA-DPB1 mismatches in HCT depend on immunopeptidome divergence and editing by HLA-DM. *Blood.* 2021;137(7):923-928.

27. Chaudhry MS, Velardi E, Malard F, van den Brink MR. Immune reconstitution after allogeneic hematopoietic stem cell transplantation: time to T up the thymus. *J Immunol.* 2017;198(1):40-46.

28. Hakki M, Riddell SR, Storek J, et al. Immune reconstitution to cytomegalovirus after allogeneic hematopoietic stem cell transplantation: impact of host factors, drug therapy, and subclinical reactivation. *Blood.* 2003;102(8):3060-3067.

29. Miller HK, Braun TM, Stillwell T, et al. Infectious risk after allogeneic hematopoietic cell transplantation complicated by acute graft-versus-host disease. *Biol Blood Marrow Transplant.* 2017;23(3):522-528.

30. Phan TL, Pritchett JC, Leifer C, et al. HHV-6B infection, T-cell reconstitution, and graft-vs-host disease after hematopoietic stem cell transplantation. *Bone Marrow Transplant.* 2018;53(12):1508-1517.

31. Tram K, Stritesky G, Wadsworth K, Ng J, Anasetti C, Dehn J. Identification of DPB1 permissive unrelated donors is highly likely. *Biol Blood Marrow Transplant.* 2017;23(1):81-86.

32. Dehn J, Setterholm M, Buck K, et al. HapLogic: a predictive human leukocyte antigen-matching algorithm to enhance rapid identification of the optimal unrelated hematopoietic stem cell sources for transplantation. *Biol Blood Marrow Transplant.* 2016;22(11):2038-2046.

33. Chabannon C, Kuball J, Bondanza A, et al. Hematopoietic stem cell transplantation in its 60s: a platform for cellular therapies. *Sci Transl Med.* 2018;10(436):eaap9630.

frontiers in Immunology

ORIGINAL RESEARCH
published: 20 January 2021
doi: 10.3389/fimmu.2020.588741

CD8+ T-Cell Repertoire in Human Leukocyte Antigen Class I-Mismatched Alloreactive Immune Response

Florence Bettens, Zuleika Calderin Sollet, Stéphane Buhler and Jean Villard *

Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Geneva University Hospitals, Geneva, Switzerland

In transplantation, direct allorecognition is a complex interplay between T-cell receptors (TCR) and HLA molecules and their bound peptides expressed on antigen-presenting cells. In analogy to HLA mismatched hematopoietic stem cell transplantation (HSCT), the TCR CDR3β repertoires of alloreactive cytotoxic $CD8^+$ responder T cells, defined by the cell surface expression of CD137 and triggered *in vitro* by HLA mismatched stimulating cells, were analyzed in different HLA class I mismatched combinations. The same HLA mismatched stimulatory cells induced very different repertoires in distinct but HLA identical responders. Likewise, stimulator cells derived from HLA identical donors activated $CD8^+$ cells expressing very different repertoires in the same mismatched responder. To mimic *in vivo* inflammation, expression of HLA class I antigens was upregulated *in vitro* on stimulating cells by the inflammatory cytokines TNFα and IFNβ. The repertoires differed whether the same responder cells were stimulated with cells treated or not with both cytokines. In conclusion, the selection and expansion of alloreactive cytotoxic T-cell clonotypes expressing a very diverse repertoire is observed repeatedly despite controlling for HLA disparities and is significantly influenced by the inflammatory status. This makes prediction of alloreactive T-cell repertoires a major challenge in HLA mismatched HSCT.

Keywords: T-cell alloreactivity, human leukocyte antigen (HLA), T-cell repertoire, T-cell receptor, hematopoietic stem cell transplantation (HSCT)

OPEN ACCESS

Edited by:
Luca Vago,
San Raffaele Scientific Institute (IRCCS), Italy

Reviewed by:
Esteban Arrieta-Bolaños,
Essen University Hospital, Germany
Leo Luznik,
Johns Hopkins University, United States

***Correspondence:**
Jean Villard
jean.villard@hcuge.ch

Specialty section:
This article was submitted to Alloimmunity and Transplantation, a section of the journal Frontiers in Immunology

Received: *29 July 2020*
Accepted: *26 November 2020*
Published: *20 January 2021*

Citation:
Bettens F, Calderin Sollet Z, Buhler S and Villard J (2021) CD8+ T-Cell Repertoire in Human Leukocyte Antigen Class I-Mismatched Alloreactive Immune Response. Front. Immunol. 11:588741. doi: 10.3389/fimmu.2020.588741

INTRODUCTION

Solid-organ and hematopoietic stem cell transplantations (HSCT) are characterized by an immune response mediated by direct, indirect and semi-direct T-cell allorecognition (1–5). In the context of HSCT, HLA compatibility between the donor and recipient is critical to prevent severe complications such as graft versus host disease (GVHD). The current standard of HLA compatibility includes the loci HLA-A, B, C, DRB1, and DQB1, with HLA-DPB1 matching considered additionally according to the number of potential compatible donors. Although HLA compatibility (i.e., the so-called 10/10 matching) is the best option, a mismatched situation can also

be considered, usually involving one mismatch at a single HLA class I or class II locus (i.e., 9/10) (6–12).

HLA class I molecules have extremely high allelic polymorphism (ebi.ac.uk/ipd/imgt/hla/stats.html) (13), potentially influencing direct alloreactivity in HLA mismatched situations. Indeed, we and others (14–16) have previously shown that HLA alloantigens can induce variable strengths of alloreactive T-cell response in cellular *in vitro* assays and *in vivo*, emphasizing the potential role of the TCR repertoire of the alloreactive responder T-cell population.

The T-cell repertoire is initially shaped in the thymus by positive and negative selection of maturing T cells on self-peptide-HLA complexes and then modulated overtime at the periphery by the cumulative history of foreign antigenic exposures (2–5). Cross-reactivity and flexibility of the T-cell receptor (TCR) allow each TCR to potentially interact with many HLA-peptide complexes (5, 17). While the CDR1 and CDR2 loops of the TCR interact primarily with cognate HLA molecules, the most variable region of the TCR, encoded by the third complementary region (CDR3) of α and β chains, is specifically involved in antigenic peptide recognition. Its nucleotide sequence, generated by somatic rearrangements of V(D)J gene segments and the random insertion/deletion of nucleotides, allows to characterize unique T-cell clonotypes. Powerful high-throughput T-cell receptor sequencing technology has been proposed as an approach to study T-cell response at the clonal level (18).

In HLA mismatched HSCT, TCRs of the donor's T cells can cross-react with non-self-HLA-peptide complexes expressed on the recipient's cells and thereby elicit a direct alloreactive immune response, which can induce a strong clinical complication called GVHD (19).

In the context of semi or fully HLA mismatched situations, quantitative analyses of *in vitro* induced alloimmune responses have revealed that up to one-tenth of circulating $CD4^+$ and $CD8^+$ T-cell clones are potentially alloreactive, accounting thus for the large diversity of the alloresponse (20, 21). In a single HLA-DPB1 mismatched situation, Arrieta-Bolanos et al. (22) demonstrated that alloreactive $CD4^+$ T-cell repertoires had virtually no overlapping TCR rearrangements in three different HLA-DPB1*04:02 individuals when stimulated by two different HeLa cells (i.e., expressing either HLA-DPB1*02:01 or DPB1*09:01). The clonal diversity was independent of the level of alloreactivity and was not based on HLA-DPB1 alloantigen structure and dissimilarity between responder and stimulator cells.

In this study, we have investigated the specificity of the alloreactive cytotoxic $CD8^+$ T-cell repertoire by using as an *in vitro* model a one-way mixed lymphocyte reaction (MLR) assay. We have performed the analyses on specific HLA class I incompatibilities according to two scenarios. In the first one, T cells derived from a given anonymous blood donor (responder cells) were stimulated with cells from distinct blood donors (stimulator cells) mismatched for the same HLA. In the second one, responder T cells of different HLA-matched anonymous blood donors were stimulated with HLA-mismatched cells from the same given blood donor. This experimental approach investigates the specific cytotoxic $CD8^+$ T-cell response in a more physiological environment involving other leukocytes like helper T cells and monocytes, representing an approximation of events occurring during the *in vivo* direct alloreactive immune response. In addition, to mimic the effect of inflammation induced by a clinical event such as infection or conditioning regimen, HLA molecules expressed by stimulator cells were upregulated by transiently incubating cells with the inflammatory cytokines, namely tumor necrosis factor alpha (TNFα) and interferon beta (IFNβ) (23).

MATERIALS AND METHODS

Cells

Peripheral blood mononuclear cells (PBMCs) were purified using standard Ficoll procedure from blood collected from anonymous donors who have been HLA genotyped at loci A, B, C, DRB1, DRB3/4/5, DQB1 and DPB1 at high resolution by the Swiss National Reference Laboratory for Histocompatibility (LNRH), while searching potential unrelated HSC donors. Cells were cryopreserved in RPMI 1640 medium (Gibco, Life Technologies, Oslo Norway) supplemented with 10 mM L-glutamine, 100 units/ml, penicillin/streptomycin (Gibco), 10% heat-inactivated human AB serum (own preparation) and 10% DMSO (Merk, Darmstadt, Germany). HLA typing was performed by reverse PCR-sequence-specific oligonucleotide microbeads arrays and high throughput sequencing (One Lambda, Canoga Park, USA). Unstimulated total $CD8^+$ T cells (average purity of 95.8%) were isolated from PBMCs by negative selection using a CD8 cell magnetic microbeads isolation kit (No. 130-096-495) (Miltenyi Biotec, Bergisch Gladbach, Germany).

Mixed Lymphocyte Reactions

One way MLRs were performed as previously described (14, 24). Briefly, responder PBMC cells ($2x10^6$) were stimulated at a ratio of 1:1 with 30Gy irradiated stimulator PBMC cells in RPMI 1640 medium (Gibco) supplemented with 10 mM L-glutamine, 100 units/ml penicillin/streptomycin (Gibco) and 10% human AB serum (own preparation). Twenty units per milliliter rIL-2 (Peprotech, London, UK) were added at days 3, 7, and 11. After 13 days of culture, responding T cells were restimulated overnight at a ratio of 1:1 with irradiated PKH-2 (Sigma-Aldrich, Buchs, Switzerland)-labeled PHA blasts obtained by activation of non-irradiated stimulatory PBMCs with one µg/ml PHA (Gibco). As a control, part of the cells was also restimulated with autologous PHA blasts. The percentage of CD137-positive PKH-2 negative CD8-positive CD56-negative viable T cells was quantified by flow cytometry. The level of alloreactivity was measured as % $CD137^+CD8^+$ cells. It corresponds to the delta between the % $CD137^+CD8^+$ cells measured at day 14, after restimulation on day 13 with allogeneic cells, minus % $CD137^+CD8^+$ measured at day 14, after restimulation with autologous cells (14, 24). To upregulate the HLA expression of stimulator cells, stimulator cells were incubated in culture

medium overnight with or without 50 ng/ml TNFα and 100 ng/ml IFNβ (PrepoTech, London, UK) prior irradiation and mixing with the responder cells.

Immunofluorescence

To label activated cytotoxic CD8 cells, APC-labeled anti-human CD8a, (clone HT8a) PerCP/Cy5.5-labeled anti-human CD56 (clone HCD56) (BioLegend, Fell, Germany) and FITC-labeled anti-human CD137 (clone 4B4-1) (Milteny Biotec) antibodies, as well as APC- and FITC-labeled murine IgG1 isotype controls (clone MOPC) (BD Bioscience, Switzerland) were used.

HLA class I surface expression was determined on $CD3^+$ T cells using the monoclonal antibodies APC-labeled anti-human CD3 and FITC-labeled anti-HLA-ABC (Miltenyi Biotec) and their corresponding isotype control. HLA- C surface expression was determined on $CD3^+$ T cells using the monoclonal antibodies APC-labeled anti-human CD3 and anti-HLA-C (clone DT9) (Milliport, Darmstadt Germany) and FITC-labeled anti-mouse IgG2b and their corresponding isotype controls (Lucernachem, Luzern, Switzerland). Data acquisition was performed on gated mononuclear cells, using the ACCURI-C6 cytometer (BD) and the CFLOWPLUS analysis software (BD Bioscience, Allschwil, Switzerland).

Cell Sorting

Activated $CD8^+$ $CD137^+$ T cells were sorted after staining with anti-CD137-FITC and anti-CD8-PEVio770 (clone BW135/80) (Miltenyi Biotec), as CD137-positive CD8-positive PKH negative cells on a BioRad s3 cell sorter (BioRad, Hercules, USA). The gating strategy is presented in **Supplementary Figure 1**. A mean of 6937 ± 5757 CD137+CD8 cells were isolated, depending on the strength of the alloresponse (average purity was 96.1 ± 1.5%).

DNA Extraction of Sorted Cell

Genomic DNA was extracted using the Genomic DNA extraction kit NucleoSpin (Machery-Nagel, Düren, Germany)

T-Cell Receptor Immunosequencing

High throughput sequencing of the TCR CDR3β region was carried out at survey resolution on the Illumina HiSeq system (Illumina, San Diego, USA) following a multiplex PCR (ImmunoSEQ© assay, Adaptive Biotechnologies, Seattle, USA). We used 400 ng of DNA from unstimulated isolated CD8+ cells and the total amount of extracted DNA from sorted alloreactive $CD137+CD8^+$ cells (see **Supplementary Table 1**). Sequencing results were sent to Adaptive Biotechnologies for analysis and datasets were downloaded from the Adaptive Biotechnologies platform for further investigations. Sample overview indicating number of productive templates, rearrangements, maximal productive frequencies and clonalities is provided in **Supplementary Table 1**. T-cell repertoire diversity was estimated by Shannon clonality, defining maximum diversity (i.e., polyclonal samples) at 0 and minimum diversity (i.e., monoclonal samples) at one. Analyses were performed using the online ImmunoSEQ Analyzer 3.0 software provided by Adaptive Biotechnologies. TCR overlap analyses were based on the amino acid sequences of the CDR3 region. Repertoire overlap between two samples (S1 and S2) was calculated with the following formulas: % TCR overlap is the number of shared clonotypes between S1 and S2 divided by the total number of clonotypes in S1 and S2. This formula is similar to the Jaccard index. Respective % overlap was calculated as the number of shared clonotypes between S1 and S2 divided by the number of clonotypes in S1 or S2, respectively. In addition, the Morisita's index (25) was estimated with ImmunoSEQ Analyzer 3.0 software. This index measures the overlap based on the statistical dispersion of clonotypes in the samples and is expected to vary between 0 (no similarity) and 1 (complete similarity). To compare rearrangements (i.e., unique CDR3β amino acid sequences) with significantly increased or decreased frequencies between two samples or experimental conditions, binomial differential abundance analysis was performed with the ImmunoSEQ Analyzer 3.0 software and as specified in (26). Respective cumulative frequencies of shared clonotypes in different experiment conditions were calculated based on abundance scatter files. Scatterplots, barplots, and boxplots were generated using R version 3.5.1.

Statistics

Paired *t*-tests were performed with GraphPad prism software version 8.01 (GraphPad, San Diego, CA, USA) A threshold of 5% was considered for statistical significance. Clonotypes with p-values lower than 0.01 were identified as being differentially abundant between two samples or experimental conditions according to the differential abundance tool.

RESULTS

To investigate alloreactive T-cell repertoires, peripheral blood mononuclear cells (PBMC) of HLA genotyped anonymous blood donors were cultured *in vitro* in a classical one-way MLR assay. After specific restimulation, $CD8^+$ responder T cells expressing the activation antigen CD137 were isolated by flow cytometry and their repertoire determined by high throughput sequencing of the TCR CDR3β region.

Clonotype Frequency Distribution of Alloreactive CD137+CD8+ T Cells

Alloreactive repertoires were determined in nine different MLR cultures and compared between responder cells, which were either activated with allogeneic stimulator cells harboring one or two HLA class I mismatches or with fully HLA mismatched stimulator cells (**Figures 1** and **2** and **Supplementary Tables 1** and **2**). In **Figures 1A, B**, representative of two prototypical MLR cultures, we observe the presence of a majority of low frequency (<0.1%) clonotypes in the unstimulated $CD8^+$ population of the responder isolated before the MLR cultures. By contrast, after allogeneic stimulation, most clonotypes were retrieved at much higher frequencies (up to 74.6%) in the alloreactive $CD137^+CD8^+$ responder population. Furthermore, some clonotypes with specific TCR rearrangements, although not

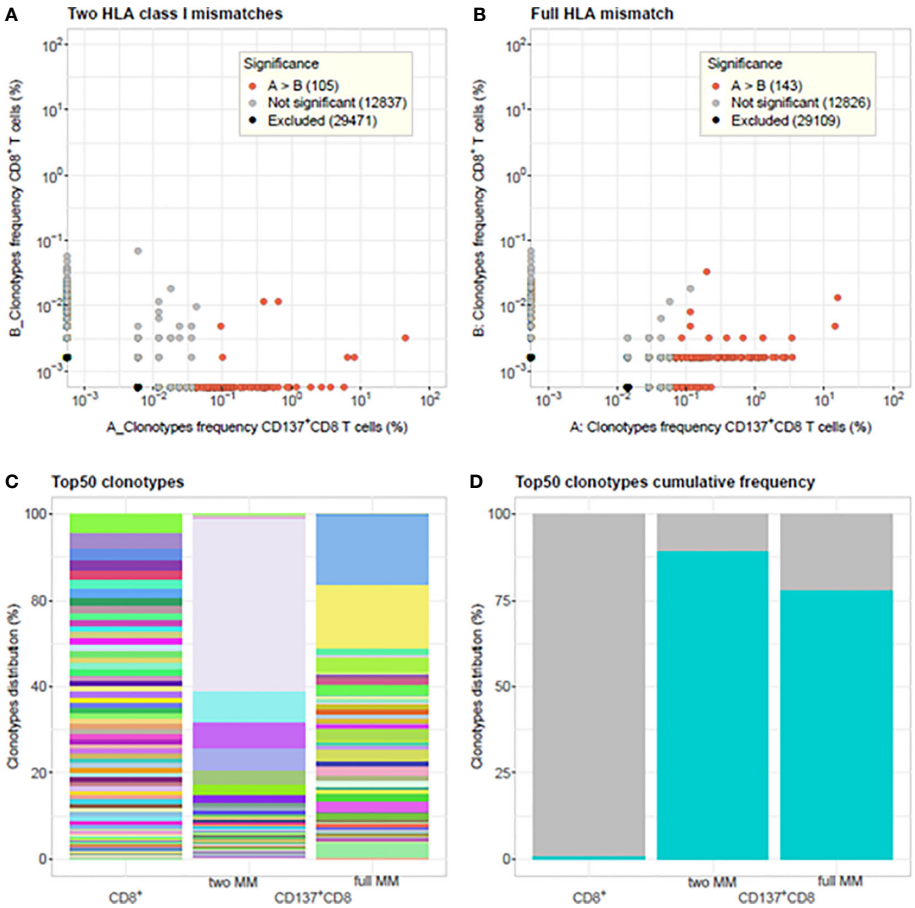

FIGURE 1 | Comparison of T-cell clonotype frequency distributions between unstimulated $CD8^+$ T cells and activated $CD137^+CD8^+$ T cells. Clonotype frequency scatterplots of $CD8^+$ T cells isolated from unstimulated PBMC of the responder blood donor before culture and $CD137^+CD8^+$ T cells isolated at day 14 from MLR cultures between responder and stimulator cells harboring: **(A)** two HLA class I mismatches A*11:01/24:02 and C*01:01/04:01(i.e., 8/10 HLA matched, 2 MM) and **(B)** full HLA mismatch (i.e., 0/10 HLA matched, full MM). In **(A)** 4.68% and in **(B)** 31.97% of the $CD137^+CD8^+$ clonotypes are shared with the unstimulated $CD8^+$ T-cells clonotypes. The cumulative frequencies of the shared $CD137^+CD8^+$ clonotypes observed in the unstimulated population are in **(A)** 0.61 and in **(B)** 0.88. The differential abundance tool of ImmunoSEQ Analyzer 3.0 was used to analyze clonotype frequencies (26). Red dots represent clonotypes that are observed with a statistically significant greater frequency in sample A compared to sample **(B)** Grey dots represent clonotypes that are not found to be differentially abundant. Black dots represent clonotypes that are excluded from the analyses. **(C)** Barplots showing the % clonal frequency distribution of the 50 most frequent clonotypes in unstimulated $CD8^+$ T cells versus $CD137^+CD8^+$ activated cells of MLR with two HLA class I mismatches and full HLA mismatch. Same colors represent same clonotypes. **(D)** Cumulative frequency of the top 50 rearrangements (cyan-blue bars) from unstimulated $CD8^+$ T cells (1.2%), $CD137^+CD8^+$ activated cells in 2 HLA class I MM MLR (89.5%) and full MM MLR (78.4%), respectively. TCR clonalities are 0.006, 0.55, and 0.31 for the unstimulated $CD8^+$ T cells, the $CD137^+CD8^+$ T cells in the 2 HLA class I MM, and full MM MLR, respectively. Morisita indices among the three samples are <0.0015.

detected in the unstimulated $CD8^+$ population, were isolated after stimulation representing the expansion of very low-frequency (< 0.001%) clonotypes. Conversely, some other clonotypes were observed only in the unstimulated population but not after stimulation. Overall, repertoire overlap between unstimulated and activated cells was very low, with Morisita indices below 0.0015. Compared to unstimulated cells, activated $CD137^+CD8^+$ T-cell repertoires showed a substantial increase in clonality: from 0.006 to 0.55 (**Figure 1A**) and from 0.006 to 0.31 (**Figure 1B**), respectively. The clonal distribution of unstimulated cells was even with clonotype frequencies varying between 0.001% and 0.1%, while a few dominant clonotypes were observed at much higher frequencies in activated cells (**Figure 1C**). The cumulative frequencies of the clonotypes retrieved after allogeneic stimulation and observed in the unstimulated population are of 61% and 88% in the two MLR, respectively. This is characteristic of the strong alloreactive expansion to high frequencies of a few clonotypes among a broad repertoire of low frequency clonotypes in the unstimulated T-cell population. Accordingly, while the 50 most frequent clonotypes in unstimulated CD8 cells represent 1.2% of all the clonotypes, they represent more than 78% of the clonotypes retrieved from activated $CD137^+CD8^+$ cells (**Figure 1D**). The proportion of specifically activated cells was not significantly ($p>0.05$) different

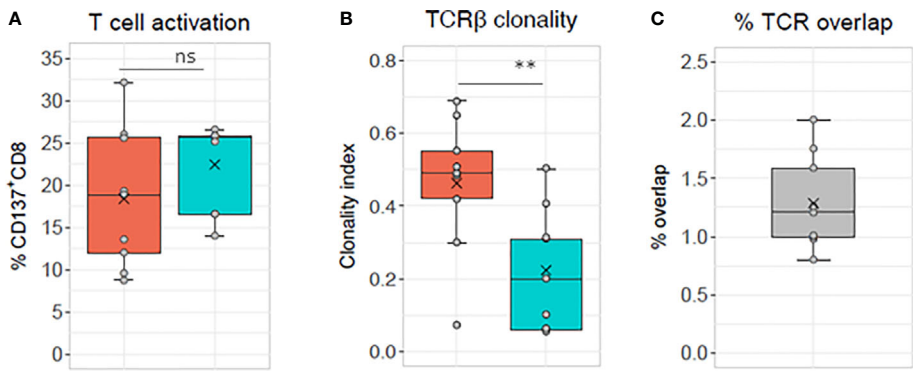

FIGURE 2 | TCR analyses of sorted CD137$^+$CD8$^+$ T cells. The same responder cells are stimulated with allogeneic cells mismatched for either one or two HLA alleles (orange box, i.e., 8–9/10 HLA matched, 1–2 MM) or for all HLA alleles (blue box, i.e., 0/10 HLA matched, full MM) in nine different MLR cultures (see **Supplementary Table 2** for HLA genotyping of the cells used for the experiments and **Supplementary Table 1** for TCR sequencing sample overview). **(A)** Percentage of CD137$^+$CD8$^+$ cells measured after specific minus autologous restimulation. Mean ± SD: 1–2 MM 18.39 ± 8.1, full MM 22.43 ± 5.1, ns, not significant paired *t* test p>0.05. **(B)** TCR clonality mean ± SD: 1–2 MM 0.46 ± 0.18, full MM 0.22 ± 0.16, **significant paired *t* test p = 0.01 and **(C)** percentage of CD137$^+$CD8$^+$ T cell clonotypes shared between the two culture conditions (mean, 1.28 ± 0.4).

whether the responder cells were stimulated with cells harboring one or two (mean, 18.39 ± 8.1) HLA mismatched alleles or cells fully (mean, 22.43 ± 5.1) HLA mismatched (**Figure 2A**). By contrast, the clonality was significantly lower in cells activated with fully mismatched stimulator cells (mean, 0.22 ± 0.16) than stimulator cells with one or two HLA class I mismatches (mean, 0.46 ± 0.18, p = 0.01, **Figure 2B**). Note: the CDR3β amino acid identity, referred to in this study as "% TCR overlap" (also named Jaccard index), represents the percentage of clonotypes shared among the total number of clonotypes observed in two repertoires under comparison. **Figure 2C** shows that the % TCR overlap between the activated CD137$^+$CD8$^+$ T cells in one to two HLA class I mismatched or fully mismatched MLR was low in all MLR pairs (mean, 1.28 ± 0.4%).

We investigated in a few MLR cultures whether CMV-specific T cell clonotypes could be enriched after the *in vitro* stimulation. However, based on a set of 18,855 potentially public CMV-specific T-cell clonotypes gathered from two resources (27, 28), we did not observed over representation of such clonotypes (data not shown).

Reproducibility of Clonal Stimulation

To evaluate whether the low TCR overlap observed in different MLR conditions is genuinely a consistent result, we monitored our experiments' reproducibility. Interestingly, MLR repeats showed heterogeneous CDR3β clonal distributions, although similar ranges of T-cell activation and clonality were retrieved (**Supplementary Table 1A** and **3** and **Supplementary Figure 2**). Nevertheless, up to 10% of shared CD137+CD8+ T-cell clonotypes were detected in triplicate MLRs. In pairwise comparisons, TCR overlap as high as 23.6% was observed, also reflected by Morisita's index up to 0.45 and respective cumulative frequencies of shared clonotypes up to 83% and above 50% in all but one replicate. Shared and non-shared clonotypes among the replicates were present across all ranges of frequencies.

T-Cell Receptor Repertoires of Alloreactive T-cell Clonotypes in Mixed Lymphocyte Reaction Cultures

To further investigate the allogeneic repertoire specificity, different combinations of HLA mismatched responder/stimulator cells were tested in distinct MLR cultures. The TCR repertoires of purified activated CD137$^+$CD8$^+$ cells were sequenced and analyzed (**Figure 3** and **Table 1**).

First, responder T cells isolated from two distinct HLA identical blood donors (i.e., HLA-A, B, C, DRB1, DRB3/4/5, and DQB1 matched at high resolution) were stimulated in parallel cultures by cells isolated from a third HLA mismatched donor (see the illustrative chart in **Figure 3A** and the top half of **Table 1**). Second, cells isolated from one donor were stimulated in parallel cultures with HLA mismatched stimulator cells isolated from two HLA identical donors (see the illustrative chart in **Figure 3B** and bottom part of **Table 1**).

As shown in **Table 1**, the clonality and percentage of activated CD137$^+$CD8$^+$ T cells were constrained within close ranges of values for each pair of MLRs. Interestingly, the percentage of activated clonotypes sharing a CDR3β amino acid sequence (i.e., % TCR overlap) was low when comparing pairs of MLR culture performed at the same time in parallel, ranging between 0 and 1.9%. In addition, the percentage of activated clonotypes shared in one MLR culture compared to the other culture ranged between 0 and 5.8% (i.e., respective % TCR overlap). Along this line, very low cumulative frequencies of shared clonotypes were measured between pairs of MLR. Accordingly, the Morisita's indices were very low and varied from 10^{-6} to 0.044. The clonal distribution of the 50 most frequent clonotypes revealed dominant clonotypes in each MLR condition (**Figure 3**, right panel), but these clonotypes did not share the same CDR3β amino acid sequence. Of note, the CD8$^+$ T-cell repertoires of the two HLA matched responders' pairs (responders 1 and 2 or responders 3 and 4, respectively) were

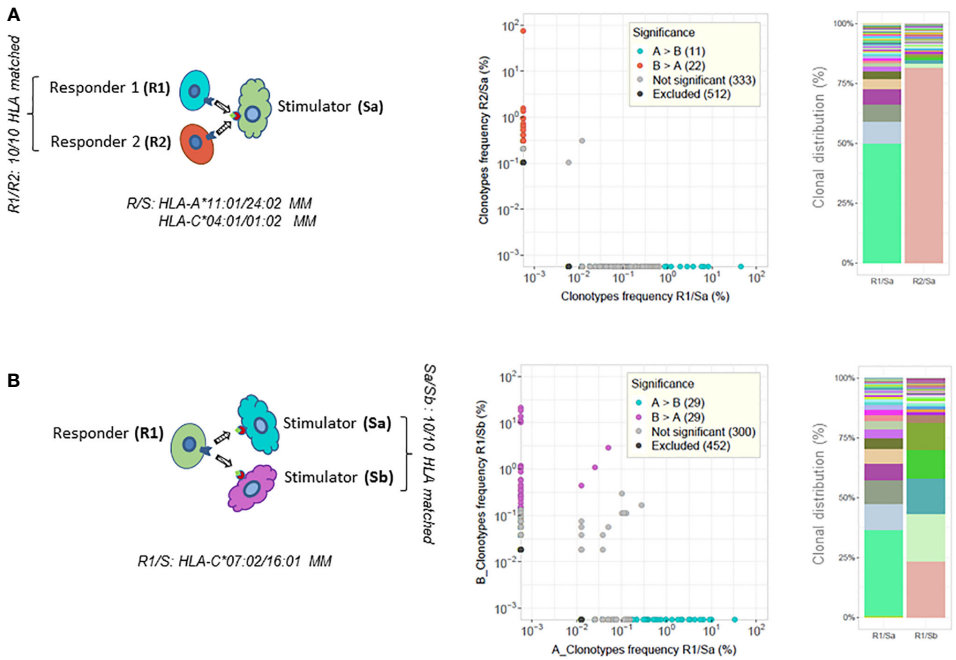

FIGURE 3 | TCR analysis of paired MLRs. $CD137^{+}CD8^{+}$ T-cell clonotype frequency scatterplot comparisons between paired MLRs and clonal distribution of the top 50 most frequent clonotypes retrieved from each culture (illustrative examples of results taken from **Table 1**). **(A)** Two MLR cultures of different HLA 10/10 matched responder cells (R1 or R2) stimulated by the same mismatched stimulator cells (Sa). R1/R2 cells are HLA-A*11:01 versus 24:02 and HLA-C*04:01 versus 01:02 mismatched with Sa cells. **(B)** Two MLR cultures of the same responder cells (R1) stimulated with cells from two distinct stimulators (Sa and Sb). Sa and Sb are 10/10 HLA matched. R1 cells are HLA-C*07:02 versus 16:01 mismatched with Sa/Sb cells. The top 50 clonotypes represent 89.5% in R1/Sa and 82.3% in R2/Sa (top experiments shown in panel **A**), and 91.9% in R1/Sa and 90.5% in R1/Sb (bottom experiments shown in panel **B**) of the respective repertoires. Colored dots in scatterplots represent clonotypes with frequencies differing significantly between paired MLRs. Cyan-blue dots represent clonotypes observed with a statistically significant greater frequency in sample A compared to sample B, red and purple dots clonotypes have a statistically significant greater frequency in sample B compared to sample A, while gray dots represent clonotypes not differing significantly in frequency. Black dots clonotypes are excluded from the analyses. Frequency analysis was performed with the differential abundance tool in ImmunoSEQAnalyzer 3.0. Barplots represent the clonal distribution of the 50 most frequent clonotypes.

very distinct sharing only 0.1% to 1.1% of TCR rearrangements (results not shown).

Effect of Human Leukocyte Antigen Upregulation on the T-Cell Receptor Repertoire of the Alloreactive T-cell Response

Under *in vivo* conditions, HLA expression, which has previously been shown to influence the allogeneic immune response (14–16), might be affected by inflammation driven by infection or GVHD. The HLA expression of stimulator cells was upregulated *in vitro* by overnight incubation with TNFα alpha and IFNβ to mimic inflammation. RNA sequencing (results not shown) and cytofluorometric analysis (**Supplementary Figure 3**) revealed that, although expressed at variable levels, all HLA class I alleles, including HLA-C alleles, were upregulated after TNFα/IFNβ induction to similar extent. This was confirmed by correlation coefficients of r = 0.84 and r = 0.92 between unstimulated and upregulated HLA-ABC and HLA-C expression, respectively. The mean fold cell surface expression upregulation measured on PBMCs, isolated from 56 different blood donors, was 1.62 for total HLA class antigens and 1.9 for HLA-C antigens on PBMCs from 32 of these 56 blood donors. Compared to untreated stimulator cells, cytokine TNFα/IFNβ treated stimulator cells, induced similar percentages of activated $CD137^{+}CD8^{+}$ T cells in 20 parallel MLR cultures (**Supplementary Figure 4**). The mean percentage of $CD137^{+}CD8^{+}$ T cells were 18.5 ± 9.8 and 14.6 ± 11.2 ($p = 0.051$) in MLR cultures of untreated versus TNFα/IFNβ treated stimulator cells, although the cell surface expression was significantly increased in treated stimulator cells ($p<0.0001$).

The clonal analysis of alloreactive $CD137^{+}CD8^{+}$ T cells induced by the same mismatched HLA stimulator cells with basal or elevated HLA antigen expression is shown for one representative experiment in the upper panels of **Figure 4**. The percentage of shared clonotypes was only 1.5% (Morisita index of 0.011). Cumulative frequencies of clonotypes shared was of 33% in the MLR culture toward cells with upregulated HLA expression compared with clonotypes of the paired MLR toward cells of basal HLA class I expression. Of note, four shared clonotypes were present at frequencies above 0.01 in one or the other cultures. The clonal distribution of the 50 most frequent clonotypes is represented in **Figure 4** (upper right

TABLE 1 | Paired alloreactive T-cell response.

MLR[a]		HLA MM[b]	% $CD137^+CD8$	TCR clonality	%TCR overlap[c]	Morisita index	Respective% TCR overlap[d]	Cumulative freq. of shared clonotypes
Responder 1	Stimulator a	HLA-A*11:01 vs 24:02 HLA-C*04:01 vs 01:02	12.0	0.55	0.2	10^{-6}	0.3	0.0002
Responder 2		HLA-A*11:01 vs 24:02 HLA-C*04:01 vs 01:02	8.8	0.64			1.6	0.004
Responder 3	Stimulator b	HLA-C*07:01 vs 12:03	12.9	0.31	0	–	–	–
Responder 4		HLA-C*07:01 vs 12:03	13.1	0.30			–	–
Responder 5[e]	Stimulator c	HLA-C*07:02 vs 16:01	25.6	0.51	1.9	0.0002	3.1	0.0099
	Stimulator d	HLA-C*07:02 vs 16:01	18.8	0.49			4.6	0.054
Responder 5[e]	Stimulator c	HLA-C*07:02 vs 16:01	22.2	0.38	1.9	0.044	4.6	0.10
	Stimulator d	HLA-C*07:02 vs 16:01	11.3	0.41			3.1	0.08
Responder 3	Stimulator b	HLA-C*07:01 vs 12:03	12.9	0.31	1.7	0.022	5.8	0.10
	Stimulator e	HLA-C*07:01 vs 12:03	13.5	0.31			2.3	0.05

[a]*Responder 1 cells are 10/10 HLA matched with responder 2 cells, dito responder 3 with responder 4, stimulator c with stimulator d and stimulator b with stimulator e cells.*
[b]*mismatched HLA class I alleles between responder and stimulator cells.*
[c]*total % shared T cell clonotypes.*
[d]*% shared T cell clonotypes in one MLR compared to the other MLR.*
[e]*MLRs performed in duplicate.*
Responder 1and responder 2 TCR clonalities of unstimulated CD8+ T cells are 0.006 and 0.023 respectively with 0.1% overlapping T-cell clonotypes and responder 3 and responder 4 TCR clonalities are 0.04 and 0.06, respectively, with 1,1% overlapping T-cell clonotypes. TCR clonality of responder 5 cells is 0.026.

panel) and shows a few unshared dominant clonotypes in both culture conditions. The results of five different MLR cultures with basal or elevated HLA antigen expression are shown in **Figure 4** (lower panels). The mean percentage and clonality of activated $CD137^+CD8^+$ T cells did not significantly differ when induced by stimulating cells treated or not with the cytokines. The % TCR overlap of activated $CD37^+CD8^+$ cells in untreated and treated conditions was low and ranged between 1.4% and 2.9% in three out of the five cultures. However, a slightly higher proportion (3.3% to 6.3%) of the TCR rearrangements was observed in the $CD137^+CD8^+$ T cells induced by the TNFα/IFNβ treated stimulator compared to the cells detected in the parallel MLR induced by untreated stimulator cells (**Figure 4**, lower, right panel). The cumulative frequencies of the shared clonotypes between cultures stimulated with TNFα/IFNβ treated stimulator versus untreated stimulator cells was 0.18 ± 0.15. It is important to note that in two out of the five cultures (**Figure 4**, dots circled in red), the repertoires were not analyzed based on activated $CD137^+CD8^+$ cells but on the total of $CD8^+$ sorted responder T cells, as the number of $CD137^+CD8^+$ cells was to low to be sorted. This explains the greater overlap ranging between 10.4% and 22.9% in these two experiments.

DISCUSSION

The data of this study suggest that in donor/recipient HLA mismatched situations (like it might occur in HSCT), the same mismatched stimulatory cells (in the case of HSCT: recipient antigen-presenting cells) induce the proliferation of different responder T-cell clonotypes (donor T-cell clonotypes) with a little percentage of shared TCRs in distinct HLA identical (i.e., 10/10 matched) responder cells. Similarly, distinct HLA identical (i.e., 10/10 matched) stimulators cells (recipient antigen-presenting cells) induce the proliferation of different responders TCR clonotypes (donor T-cell clonotypes) and repertoire in the same HLA mismatched responder/stimulator configuration. Random selection and expansion of alloreactive T-cell clonotypes were also observed in MLR repeats and triplicates, although at a lesser extent Moreover, the level of mismatched HLA cytokine-modulated expression on the cell surface does not influence the strength of the T-cell response, but it affects the repertoire of the alloreactive $CD137^+CD8^+$ T cells retrieved after culture. Our results are in line with previous reports (3–5, 17) indicating the very high flexibility of alloreactive TCRs in the alloimmune response, as well as the

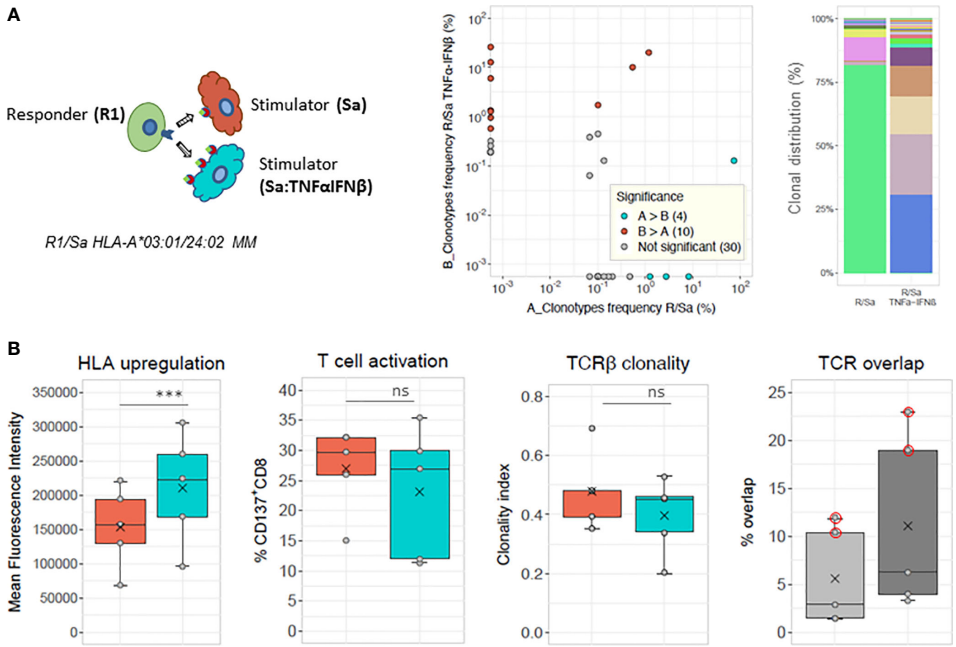

FIGURE 4 | Modulation of allogeneic T cell response by HLA class I upregulation. **(A)** Chart of MLR stimulations: responder cells (R) derived from the same blood donor are stimulated in parallel cultures with allogeneic stimulator cells, with basal (Sa) or upregulated HLA surface expression after overnight TNFα and IFNβ cytokine treatment (Sa: TNFαINFβ). Scatterplot and barplot are a representative example of one of the five parallel MLRs shown in **(B)**. Frequency scatterplot showing 1.5% (8/531) of shared clonotypes between the two conditions. Colored dots in scatterplots represent clonotypes with frequencies differing significantly between paired MLRs (i.e., cyan-blue dots represent clonotypes that have a statistically significant greater frequency in sample A compared to sample B, red dots clonotypes have a statistically significant greater frequency in sample B compared to sample A), while gray dots represent clonotypes not differing significantly in frequency. The frequency analysis was performed with the differential abundance tool in ImmunoSEQAnalyzer 3.0). The cumulative frequency of the shared clonotypes in MLR stimulation with Sa : TNFαINFβ is 0.33 and with untreated Sa is 0.75, respectively. (Right panel) Clonal distribution of the top 50 most frequent clonotypes in each culture (i.e., with basal (R/Sa) or upregulated (R/SaTNFαINFβ) HLA cell surface expression of stimulator cells). Responder and stimulator cells are mismatched for HLA-A*03:01 versus 24:02. The top 50 clonotypes represent: 91.3% in R/Sa and 87.4% in R/Sa+TNFαINFβ of the respective repertoires. **(B)** HLA cell surface upregulation, T-cell activation, TCRβ clonality and % TCR overlap determined in five pairs of MLRs performed with stimulator cells treated with TNFα and IFNβ (cyan-blue boxes) or not treated before stimulation (orange boxes). The light and dark gray boxes in the last panel represent total and respective TCR overlap (in %), respectively. The mean value of Morisita's indices for TCR overlap is 0.024. Of note, among the five pairs of MLR, two pairs were performed with fully HLA mismatched stimulator cells, while the other stimulator cells were mismatched for one HLA-C (07:02 versus 16:01) or one HLA-A (03:01 versus 24:02) allele. TCR overlap was determined based on $CD137^+CD8^+$ T cells except in two out of the five MLR cultures where total $CD8^+$ T cells were isolated. This is shown by the red circled dots on the lower right panel. No statistical differences (ns) in T-cell activation or TCR clonality is observed between MLRs with basal or upregulated HLA expression (paired *t*-tests). ***) HLA induced upregulation was statistically significant (paired *t* test, p = 0.0053).

random selection and expansion of alloreactive cytotoxic T-cell clones, which is influenced by the complex interplay between the TCRs and mismatched-HLA-peptide complexes, alongside additional factors such as the fitness of clones, cytokines and inflammatory events.

To monitor the cytotoxic T-cell response in transplantation, T cells were isolated from defined HLA genotyped healthy donors and stimulated with specific HLA-mismatched allogeneic cells *in vitro*. We are aware that this study's *in vitro* model might represent only a snapshot of the possible *in vivo* alloimmune response. Indeed, the observed expansion of dominant clonotypes can be affected by the clones' microenvironment and fitness. In this study, the starting PBMC number was 2 million and might not always include the total number of cells with alloreactive potential. This is especially true when considering that the clonotypes observed with increased frequencies in the responding cell population were not necessarily detected in the starting population because of their supposed too low frequencies. Only an increased number of cultures performed in parallel would probably clarify this matter and minimize technical limitations (29–31). However, working with a higher number of cells from two individuals with a single HLA mismatch would be cumbersome and require higher blood withdrawals, which is not ethically acceptable. Emerson et al. (32) distinguished low (i.e., not seen before stimulation and representing 84% of the reacting clonotypes) and high abundance clonotypes (i.e., seen before and after stimulation) in the alloresponse of the total T-cell population. They reported repetitive detection of the high abundance clonotypes only.

Similarly to others (33), we observed shared and non-shared clonotypes at various frequencies in activated $CD137^+CD8^+$ cells of MLR triplicates (**Supplementary Figure 2**). Conscious of the technical restrictions, we expected that the very abundant alloreactive clones would be repeatedly detected, and we have reduced the possible randomization of the results by testing

multiple alloresponses toward different single HLA antigens. With these precautions in mind, we are confident that the results obtained in this study stand to estimate events occurring during the alloimmune response. To avoid inter-individual variability and to better define the diversity of the TCR repertoire of alloreactive cells, distinct stimulations of the same cells were examined. Furthermore, alloreactive activated $CD137^+CD8^+$ T cells were sorted after being specifically restimulated to minimize bystander stimulation of non-specific clones. We can however not exclude that, while sorting activated cells, we also sorted some auto-specific cells. Note that when measuring T cell activation, we did stimulate the cultures with autologous cells to monitor specificity. Accordingly, the referred activation represents % $CD137^+CD8^+$ in specifically restimulated cultures minus % $CD137^+CD8^+$ T cells in the same culture after autologous restimulation. The % after autologous stimulation never exceeded 10% of $CD137^+CD8^+$ (data not shown).

Using such assays and similar to others (21, 22, 33), we identified highly frequent clonotypes among alloreactive $CD8^+$ T cells previously not detected in unstimulated cells. This confirms the highly efficient expansion of reactive T-cell clones in the allogeneic immune response. The increased clonality observed in **Figures 1** and **2** for T cells stimulated with a single or a few alloantigens compared to T cells stimulated with a large number of alloantigens is concordant with the results of de Wolf et al. (21). These authors reported a reduced diversity and thus increased clonality of the alloresponsive $CD4^+$ and $CD8^+$ T cells stimulated by haploidentical cells (i.e., half-matched for HLA) as compared to cells stimulated with fully HLA mismatched cells. This supports the belief that a higher number of HLA mismatches correspond to a higher number of potential alloantigens, thus inducing a more diverse alloreactive repertoire at fault for a stronger immune response in transplantation. Nevertheless, it has to be kept in mind that a single mismatched HLA molecule might bind a large variety of peptides and may represent many different alloantigens, possibly interacting with different TCRs. On the other side of this multifaceted ligand/receptor cognate interaction, one given TCR was reported to react with more than 100 different peptides (17, 34). Other authors reported up to 1 million peptides being potentially recognized by each TCR. This TCR's particularity probably also influences the specificity of the alloresponse. In an HLA class II mismatched situation, Arrieta-Bolanos et al. (22) reported that although higher levels of alloresponse were detected in the same HLA-DPB1*04:02 individual after stimulation with Hela cells expressing HLA-DPB1*09:01 rather than HLA-DPB1*02:01, very different clonotypes of similar TCR clonality were measured among the responding CD4+ T cells. Their results suggest that the number of amino acid mismatches between the HLA of the responder and stimulator cell does not influence the clonality of the responding T cells, leading to the same conclusion as we do regarding the pleiotropy of the TCR repertoire when stimulated under similar controlled conditions.

Furthermore, in this study, we have described that the very same HLA mismatched stimulatory cells (thus expressing the same HLA-peptide alloantigens) induce the proliferation of clonotypes with a very low percentage of shared TCR in distinct but HLA 10/10 matched responder cells. This was also the case when the alloresponse was restricted to a single allelic mismatch. These results suggest an individual selection of these clonotypes, which could be influenced by environmental factors and/or immunological memory of the responding cells' past immune responses, similar to the phenomenon described in twins (35). The preferential thymic selection of different repertoires as well as the distinctive memory versus naïve $CD8^+$ T-cell repertoires could also influence the results (36, 37). Likewise to us, Arrieta-Bolanos et al. (22) reported that the same HLA-DP mismatched expressing cells induce $CD4^+$ T cells with different TCR rearrangements in different individuals.

We have also shown that HLA identical (10/10 matched) stimulatory cells isolated from distinct healthy donors induced the proliferation of different T-cell clonotypes in the same HLA mismatched responder cells. Such diversity stands for the "peptide-centric model" described by Cole and al (38). as it probably reflects the T-cell response toward heterogeneous peptidomes bound by the same HLA alleles of different individuals. A refined peptidome analysis would be required to investigate this hypothesis further. Along this line, Koyama et al. (39) reported that different T-cell clonotypes were detected in patients undergoing GVHD in the skin, colon, or blood of the very same patient, suggesting that HLA molecules present distinct peptides in various organs. In contrast, Michalek et al. (40) reported the expansion of a unique specific $CD4^+$ T-cell clone in the blood of an HSCT patient undergoing GVHD.

Since additional factors such as concomitant infection or any ongoing inflammatory processes could play a role in the *in vivo* allogeneic response, we have monitored inflammatory cytokines' effect on the alloresponse. To do so, the levels of HLA expression of stimulatory cells were boosted *in vitro* with a combination of TNFα and IFNβ. We could observe that all HLA class I alleles' transcription (results not shown) and translation were increased with the same order of magnitude. Whether increased HLA expression alters the peptidome presented, has not yet been investigated. Nevertheless, comparing the alloresponses of the same $CD8^+$ T cells induced by stimulating cells with or without upregulated HLA surface expression revealed no significant change in the response's strength and clonality. Some TCR rearrangements were shared, revealing mean cumulative frequencies of 0.38 ± 0.32 and 0.18 ± 0.15 for shared clonotypes of the same CD8+ T cells induced by stimulating cells without or with upregulated HLA surface expression. The overlap of activated $CD8^+$ cells in these two conditions was however weak, with a maximum of 6.3% of shared clonotypes. Although the low overlap might be influenced by the randomness observed in culture repeats, it supports a random selection and expansion of responding T cells with high TCR flexibility to bind foreign peptide-HLA complexes in alloimmune processes. Additionally, the results suggest that inflammation, occurring for example, after a bystander viral infection in transplantation, might also alter the allogeneic immune response as it might induce the activation of new responding alloreactive T cells. Potential effect of allele-specific expression

which is a topic of debate (41–43) has not been investigated in this study.

Our study focused on specific HLA class I incompatibilities, mainly at the HLA-C locus. This could be considered a limitation of the study, however as HLA mismatched transplantations at the HLA-C locus have been privileged for many years in our institution, while HLA-B mismatches avoided, we did not have cells to perform HLA-B mismatch experiments and only a few experiments could involve HLA-A incompatibilities. In addition, it was not possible to assess every HLA-C mismatch combination due to the extensive HLA polymorphism. In HSCT, it is well known that particular mismatch combinations are so-called "permissive" and are leading to different transplantation outcomes (10, 14, 44). In our previous study (14), permissive HLA mismatched combinations did not induce $CD137^{+}CD8^{+}$ positive T cells suggesting any TCR repertoire bias before or after stimulation.

In conclusion, our data demonstrated the vast diversity of the alloimmune response regarding the expansion of T-cell clones. However, we are aware of the limitation of *in vitro* studies not allowing to fully extrapolate to *in vivo* situations. Factors that influence this expansion such as infection, immunosuppression and GVHD are under investigation (45–47). Interestingly, we recently presumed about such factors when we reported that the expansion of TCR clones in reconstructing the repertoire after HSCT was correlated to CMV reactivation in patients one-year post-HSCT, without being CMV specific (45). Another recent publication also demonstrates the difficulties to clearly associate the TCR repertoire and clinical events (46). Thus, the prediction of alloreactive T-cell response based on the TCR repertoire before and after transplantation remains a major challenge in HLA mismatched situations. Clinical protocols or pharmacological agents targeting specifically alloreactive T cells to control clinical complications such as GVHD or the graft-versus-leukemia effect (GVL) could be very challenging to establish.

DATA AVAILABILITY STATEMENT

The data sets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**. DOI: 10.21417/FSB2020FI. The URL is: clients.adaptivebiotech.com/pub/ studer-bettens-2020-fi.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical committee of University Hospitals of Geneva (CER 06–208 and 08–208R). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

FB designed research studies, conducted experiments, and acquired and analyzed the data. ZS made the graphs. FB, SB, and JV wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Swiss National Science Foundation (grant #310030_173237/1), the Fondation privée of HUG, IRGHET (International Research Group on unrelated Hematopoietic stem cell Transplantation), and the Dr. Henri Dubois-Ferrière Dinu Lippatti foundation.

ACKNOWLEDGMENTS

The authors thank the technicians of the LNRH for their efficient support for HLA typing and the freezing of the cells. They also thank the cytofluorometry platform for their help in sorting cells.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2020.588741/full#supplementary-material

REFERENCES

1. Boardman DA, Jacob J, Smyth LA, Lombardi G, Lechler RI. What Is Direct Allorecognition? *Curr Transplant Rep* (2016) 3(4):275–83. doi: 10.1007/s40472-016-0115-8
2. DeWolf S, Sykes M. Alloimmune T cells in transplantation. *J Clin Invest* (2017) 127(7):2473–81. doi: 10.1172/JCI90595
3. Mason D. A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunol Today* (1998) 19(9):395–404. doi: 10.1016/S0167-5699(98)01299-7
4. Colf LA, Bankovich AJ, Hanick NA, Bowerman NA, Jones LL, Kranz DM, et al. How a single T cell receptor recognizes both self and foreign MHC. *Cell* (2007) 129(1):135–46. doi: 10.1016/j.cell.2007.01.048
5. Degauque N, Brouard S, Soulillou JP. Cross-Reactivity of TCR Repertoire: Current Concepts, Challenges, and Implication for Allotransplantation. *Front Immunol* (2016) 7:89. doi: 10.3389/fimmu.2016.00089
6. Loiseau P, Busson M, Balere ML, Dormoy A, Bignon JD, Gagne K, et al. HLA Association with hematopoietic stem cell transplantation outcome: the number of mismatches at HLA-A, -B, -C, -DRB1, or -DQB1 is strongly associated with overall survival. *Biol Blood Marrow Transplant* (2007) 13(8):965–74. doi: 10.1016/j.bbmt.2007.04.010
7. Lee SJ, Klein J, Haagenson M, Baxter-Lowe LA, Confer DL, Eapen M, et al. High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood* (2007) 110(13):4576–83. doi: 10.1182/blood-2007-06-097386
8. Petersdorf EW. Optimal HLA matching in hematopoietic cell transplantation. *Curr Opin Immunol* (2008) 20(5):588–93. doi: 10.1016/j.coi.2008.06.014

9. Shaw BE, Mayor NP, Russell NH, Apperley JF, Clark RE, Cornish J, et al. Diverging effects of HLA-DPB1 matching status on outcome following unrelated donor transplantation depending on disease stage and the degree of matching for other HLA alleles. *Leukemia* (2010) 24(1):58–65. doi: 10.1038/leu.2009.239
10. Fleischhauer K, Shaw BE, Gooley T, Malkki M, Bardy P, Bignon JD, et al. Effect of T-cell-epitope matching at HLA-DPB1 in recipients of unrelated-donor haemopoietic-cell transplantation: a retrospective study. *Lancet Oncol* (2012) 13(4):366–74. doi: 10.1016/S1470-2045(12)70004-9
11. Tiercy JM. How to select the best available related or unrelated donor of hematopoietic stem cells? *Haematologica* (2016) 101(6):680–7. doi: 10.3324/haematol.2015.141119
12. Petersdorf EW, Carrington M, O'HUigin C, Bengtsson M, De Santis D, Dubois V, et al. Role of HLA-B exon 1 in graft-versus-host disease after unrelated haemopoietic cell transplantation: a retrospective cohort study. *Lancet Haematol* (2020) 7(1):e50–60. doi: 10.1016/S2352-3026(19)30208
13. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA Database. *Nucleic Acids Res* (2020) 48(D1):D948–d55. doi: 10.1093/nar/gkz950
14. Bettens F, Buhler S, Tiercy JM. Allorecognition of HLA-C Mismatches by CD8 + T Cells in Hematopoietic Stem Cell Transplantation Is a Complex Interplay between Mismatched Peptide-Binding Region Residues, HLA-C Expression, and HLA-DPB1 Disparities. *Front Immunol* (2016) 7:584. doi: 10.3389/fimmu.2016.00584
15. Petersdorf EW, Gooley TA, Malkki M, Bacigalupo AP, Cesbron A, Du Toit E, et al. HLA-C expression levels define permissible mismatches in hematopoietic cell transplantation. *Blood* (2014) 124(26):3996–4003. doi: 10.1182/blood-2014-09-599969
16. Morishima S, Kashiwase K, Matsuo K, Azuma F, Yabe T, Sato-Otsubo A, et al. High-risk HLA alleles for severe acute graft-versus-host disease and mortality in unrelated donor bone marrow transplantation. *Haematologica* (2016) 101(4):491–8. doi: 10.3324/haematol.2015.136903
17. Birnbaum ME, Mendoza JL, Sethi DK, Dong S, Glanville J, Dobbins J, et al. Deconstructing the peptide-MHC specificity of T cell recognition. *Cell* (2014) 157(5):1073–87. doi: 10.1016/j.cell.2014.03.047
18. Robins HS, Campregher PV, Srivastava SK, Wacher A, Turtle CJ, Kahsai O, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* (2009) 114(19):4099–107. doi: 10.1182/blood-2009-04-217604
19. Gras S, Kjer-Nielsen L, Chen Z, Rossjohn J, McCluskey J. The structural bases of direct T-cell allorecognition: implications for T-cell-mediated transplant rejection. *Immunol Cell Biol* (2011) 89(3):388–95. doi: 10.1038/icb.2010.150
20. Suchin EJ, Langmuir PB, Palmer E, Sayegh MH, Wells AD, Turka LA. Quantifying the frequency of alloreactive T cells in vivo: new answers to an old question. *J Immunol* (2001) 166(2):973–81. doi: 10.4049/jimmunol.166.2.973
21. DeWolf S, Grinshpun B, Savage T, Lau SP, Obradovic A, Shonts B, et al. Quantifying size and diversity of the human T cell alloresponse. *JCI Insight* (2018) 3(15). doi: 10.1172/jci.insight.121256
22. Arrieta-Bolaños E, Crivello P, Metzing M, Meurer T, Ahci M, Rytlewski J, et al. Alloreactive T Cell Receptor Diversity against Structurally Similar or Dissimilar HLA-DP Antigens Assessed by Deep Sequencing. *Front Immunol* (2018) 9(280). doi: 10.3389/fimmu.2018.00280
23. Johnson DR. Locus-specific constitutive and cytokine-induced HLA class I gene expression. *J Immunol* (2003) 170(4):1894–902. doi: 10.4049/jimmunol.170.4.1894
24. Bettens F, Schanz U, Tiercy JM. Lack of recognition of HLA class I mismatches outside alpha1/alpha2 domains by CD8+ alloreactive T lymphocytes: the HLA-B44 paradigm. *Tissue Antigens* (2013) 81(6):414–8. doi: 10.1111/tan.12102
25. Morisita M. Measuring of the dispersion of individuals and analysis of the distributional patterns. *Mem Fac Sci Kyushu Univ Ser E (Biol)* (1959) 2:215–35.
26. DeWitt WS, Emerson RO, Lindau P, Vignali M, Snyder TM, Desmarais C, et al. Dynamics of the cytotoxic T cell response to a model of acute viral infection. *J Virol* (2015) 89(8):4517–26. doi: 10.1128/JVI.03474-14
27. Emerson RO, DeWitt WS, Vignali M, Gravley J, Hu JK, Osborne EJ, et al. Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire. *Nat Genet* (2017) 49(5):659–65. doi: 10.1038/ng.3822
28. Bagaev DV, Vroomans RMA, Samir J, Stervbo U, Rius C, Dolton G, et al. VDJdb in 2019: database extension, new analysis infrastructure and a T-cell receptor motif compendium. *Nucleic Acids Res* (2020) 48(D1):D1057–d62. doi: 10.1093/nar/gkz874
29. Warren RL, Freeman JD, Zeng T, Choe G, Munro S, Moore R, et al. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res* (2011) 21(5):790–7. doi: 10.1101/gr.115428.110
30. Greiff V, Miho E, Menzel U, Reddy ST. Bioinformatic and Statistical Analysis of Adaptive Immune Repertoires. *Trends Immunol* (2015) 36(11):738–49. doi: 10.1016/j.it.2015.09.006
31. Rytlewski J, Deng S, Xie T, Davis C, Robins H, Yusko E, et al. Model to improve specificity for identification of clinically-relevant expanded T cells in peripheral blood. *PloS One* (2019) 14(3):e0213684. doi: 10.1371/journal.pone.0213684
32. Emerson RO, Mathew JM, Konieczna IM, Robins HS, Leventhal JR. Defining the alloreactive T cell repertoire using high-throughput sequencing of mixed lymphocyte reaction culture. *PloS One* (2014) 9(11):e111943. doi: 10.1371/journal.pone.0111943
33. Morris H, DeWolf S, Robins H, Sprangers B, LoCascio SA, Shonts BA, et al. Tracking donor-reactive T cells: Evidence for clonal deletion in tolerant kidney transplant patients. *Sci Trans Med* (2015) 7(272):272ra10. doi: 10.1126/scitranslmed.3010760
34. Veerapathran A, Pidala J, Beato F, Yu XZ, Anasetti C. Ex vivo expansion of human Tregs specific for alloantigens presented directly or indirectly. *Blood* (2011) 118(20):5671–80. doi: 10.1182/blood-2011-02-337097
35. Brodin P, Jojic V, Gao T, Bhattacharya S, Angel CJ, Furman D, et al. Variation in the human immune system is largely driven by non-heritable influences. *Cell* (2015) 160(1-2):37–47. doi: 10.1016/j.cell.2014.12.020
36. Robins HS, Srivastava SK, Campregher PV, Turtle CJ, Andriesen J, Riddell SR, et al. Overlap and effective size of the human CD8+ T cell receptor repertoire. *Sci Trans Med* (2010) 2(47):47ra64. doi: 10.1126/scitranslmed.3001442
37. Zvyagin IV, Pogorelyy MV, Ivanova ME, Komech EA, Shugay M, Bolotin DA, et al. Distinctive properties of identical twins' TCR repertoires revealed by high-throughput sequencing. *Proc Natl Acad Sci USA* (2014) 111(16):5980–5. doi: 10.1073/pnas.1319389111
38. Cole DK, Miles KM, Madura F, Holland CJ, Schauenburg AJ, Godkin AJ, et al. T-cell receptor (TCR)-peptide specificity overrides affinity-enhancing TCR-major histocompatibility complex interactions. *J Biol Chem* (2014) 289(2):628–38. doi: 10.1074/jbc.M113.522110
39. Koyama D, Murata M, Hanajiri R, Akashi T, Okuno S, Kamoshita S, et al. Quantitative Assessment of T Cell Clonotypes in Human Acute Graft-versus-Host Disease Tissues. *Biol Blood Marrow Transplant* (2019) 25(3):417–23. doi: 10.1016/j.bbmt.2018.10.012
40. Michalek J, Collins RH, Hill BJ, Brenchley JM, Douek DC. Identification and monitoring of graft-versus-host specific T-cell clone in stem cell transplantation. *Lancet* (2003) 361(9364):1183–5. doi: 10.1016/S0140-6736(03)12917-0
41. Aguiar VRC, Cesar J, Delaneau O, Dermitzakis ET, Meyer D. Expression estimation and eQTL mapping for HLA genes with a personalized pipeline. *PloS Genet* (2019) 15(4):e1008091. doi: 10.1371/journal.pgen.1008091
42. Bettens F, Brunet L, Tiercy JM. High-allelic variability in HLA-C mRNA expression: association with HLA-extended haplotypes. *Genes Immun* (2014) 15(3):176–81. doi: 10.1038/gene.2014.1
43. Greene JM, Wiseman RW, Lank SM, Bimber BN, Karl JA, Burwitz BJ, et al. Differential MHC class I expression in distinct leukocyte subsets. *BMC Immunol* (2011) 12:39. doi: 10.1186/1471-2172-12-39
44. Joris MM, Lankester AC, von dem Borne PA, Kuball J, Bierings M, Cornelissen JJ, et al. Translating in vitro prediction of cytotoxic T cell alloreactivity to hematopoietic stem cell transplantation outcome. *Transpl Immunol* (2014) 30(2–3):59–64. doi: 10.1016/j.trim.2013.08.006
45. Buhler S, Bettens F, Dantin C, Ferrari-Lacraz S, Ansari M, Mamez AC, et al. Genetic T-cell receptor diversity at 1 year following allogeneic hematopoietic

stem cell transplantation. *Leukemia* (2019) 34(5):1422–32. doi: 10.1038/s41375-019-0654-y

46. Link-Rachner CS, Eugster A, Rücker-Braun E, Heidenreich F, Oelschlägel U, Dahl A, et al. T-cell receptor-α repertoire of CD8+ T cells following allogeneic stem cell transplantation using next-generation sequencing. *Haematologica* (2019) 104(3):622–31. doi: 10.3324/haematol.2018.199802
47. Kanakry CG, Ganguly S, Zahurak M, Bolaños-Meade J, Thoburn C, Perkins B, et al. Aldehyde dehydrogenase expression drives human regulatory T cell resistance to posttransplantation cyclophosphamide. *Sci Trans Med* (2013) 5 (211):211ra157. doi: 10.1126/scitranslmed.3006960

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PLOS GENETICS

RESEARCH ARTICLE

Regulation of HLA class I expression by non-coding gene variations

Florence Bettens[1], Halit Ongen[2], Guillaume Rey[2], Stéphane Buhler[1], Zuleika Calderin Sollet[1], Emmanouil Dermitzakis[2], Jean Villard[1]*

1 Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Geneva University Hospitals, Geneva, Switzerland, **2** Department of Genetic Medicine and Development, University of Geneva, Geneva, Switzerland

* jean.villard@hcuge.ch

Citation: Bettens F, Ongen H, Rey G, Buhler S, Calderin Sollet Z, Dermitzakis E, et al. (2022) Regulation of HLA class I expression by non-coding gene variations. PLoS Genet 18(6): e1010212. https://doi.org/10.1371/journal.pgen.1010212

Editor: Jukka Partanen, Finnish Red Cross Blood Service, FINLAND

Received: September 27, 2021

Accepted: April 20, 2022

Published: June 6, 2022

Data Availability Statement: All data files are available from the Yareta database (https://doi.org/10.26037/yareta:ry65havvezhmnlwa2hzyvxv2ga).

Funding: This research received funding from The Swiss National Science Foundation grant #310030_173237/1 to FB and JV. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Abstract

The Human Leukocyte Antigen (HLA) is a critical genetic system for different outcomes after solid organ and hematopoietic cell transplantation. Its polymorphism is usually determined by molecular technologies at the DNA level. A potential role of HLA allelic expression remains under investigation in the context of the allogenic immune response between donors and recipients. In this study, we quantified the allelic expression of all three HLA class I loci (HLA-A, B and C) by RNA sequencing and conducted an analysis of expression quantitative traits loci (eQTL) to investigate whether HLA expression regulation could be associated with non-coding gene variations. HLA-B alleles exhibited the highest expression levels followed by HLA-C and HLA-A alleles. The max fold expression variation was observed for HLA-C alleles. The expression of HLA class I loci of distinct individuals demonstrated a coordinated and paired expression of both alleles of the same locus. Expression of conserved HLA-A~B~C haplotypes differed in distinct PBMC's suggesting an individual regulated expression of both HLA class I alleles and haplotypes. Cytokines TNFα /IFNβ, which induced a very similar upregulation of HLA class I RNA and cell surface expression across alleles did not modify the individually coordinated expression at the three HLA class I loci. By identifying *cis* eQTLs for the HLA class I genes, we show that the non-coding eQTLs explain 29%, 13%, and 31% of the respective HLA-A, B, C expression variance in unstimulated cells, and 9%, 23%, and 50% of the variance in cytokine-stimulated cells. The eQTLs have significantly higher effect sizes in stimulated cells compared to unstimulated cells for HLA-B and HLA-C genes expression. Our data also suggest that the identified eQTLs are independent from the coding variation which defines HLA alleles and thus may be influential on intra-allele expression variability although they might not represent the causal eQTLs.

Author summary

In transplantation, the allogenic immune response is linked to the HLA compatibility between donor and recipient, HLA genes being highly polymorphic. The impact of this allelic polymorphism on MHC cell surface expression and the potential role of expression

in the T cell activation and on clinical outcomes remains poorly investigated. In this study, we documented the individual variability of allelic HLA class I expression in PBMCs by RNA sequencing. To mimic the inflammatory clinical situation after transplantation, we performed a similar analysis following cytokine (TNFα/IFNβ) stimulation of PBMCs. The results demonstrated a coordinated and paired expression of both alleles of the same locus in all individuals. The levels of expression of matched HLA-A~B~C haplotypes differed in distinct PBMCs suggesting that expression of both HLA class I alleles and haplotypes is regulated individually. To identify *cis* regulatory elements of expression we performed an eQTL analysis on unstimulated and stimulated PBMCs. These eQTLs accounted for up to 9%, 23% and 50% of the respective HLA-A, B and C expression variance in stimulated PBMCs. Of interest, we could show that they are independent from the SNPs encoding allelic variation, meaning that they could explain a different portion of the variance in HLA expression.

Introduction

HLA class I molecules are expressed constitutively on nucleated cells and function as antigen presenting molecules to cytotoxic T cells in immune responses to pathogens, cancer cells and autoantigens [1]. They also regulate the innate immune response by influencing NK cell activation [2]. HLA disparities between donors and patients induce allogenic immune responses, leading to rejection or graft versus host disease (GVHD) in different transplantation settings [3]. HLA class I molecules are characterized by an extremely high polymorphism and variable levels of expression [4–7], which potentially influence their function(s). For instance, tumor cells downregulate HLA expression to escape immune-surveillance by T cells [8,9]. *In vitro* allogenic immune responses were previously shown by us and others [10–12] to be partially dependent on HLA-C expression. Intra- and inter-individual variation of HLA cell surface expression on T and B cells was also described to impact antibody dependent cytotoxic immune response [13]. More recently, HLA-expression was also shown to impact crossmatch outcomes in transplantation diagnostic [14]. Two retrospective clinical studies tested the impact of HLA expression on clinical outcome in the setting of HLA-C mismatched unrelated hematopoietic stem cell transplantation (HSCT). Both studies used as proxy of HLA-C expression the mean cell surface expression reported by Apps et al. [4]. While Petersdorf et al [15] found an association between highly expressed HLA-C*03 or C*14 allotypes and increased mortality, Morishima et al [16] did not. In other clinical settings, a high HLA-C expression was correlated to more efficient recognition of HIV by cytotoxic cells and lower viremia in patients [4,17], while lower levels of HLA-C expression had a protective effect from Crohn's disease [7]. Similarly, higher HLA-B*27 expression levels were reported in patients with ankylosing spondylitis compared to healthy donors [18].

HLA expression regulation underlies complex mechanisms, involving genetic polymorphisms as well as transcriptional and translational factors [19–23]. To date, no consensus explanation exists for the variable levels of HLA expression that are observed [24].

Initial reports on HLA class I expression were based on cell surface and gene expression and relied on allotypes, without the possibility to discriminate between the two alleles carried by heterozygous donors at a given locus. [4,7,17,25,26]. The advent of more recent technologies using RNA sequencing has allowed qualitative, quantitative and equivalent inter-allelic expression analyses [14,27–30].

An extended analysis of single nucleotide polymorphisms (SNPs) in the non-coding regions of HLA genes, which tag transcript abundance (expression quantitative traits loci, eQTLs) was performed by Aguiar et al. in lymphoblastic cells of the GEUVADIS consortium [28]. They identified HLA-C associated eQTL's in strong linkage disequilibrium with the previously described genetic variants rs9264942 and rs2395471 and deletion (263del/ins) [17,31,32]. Vandiedonck et al. 2011 [33] analyzed the specific transcriptional variation of haplotypes associated with autoimmune diseases in three known homozygous cell lines PGF, COX and QBL.

In the present study, HLA class I allelic RNA expression, analyzed in peripheral blood mononuclear cells (PBMCs) of healthy blood donors, was quantified using RNA sequencing. To mimic clinical situations with inflammation driven by events such as alloreactivity, autoimmunity or infection, the analyses were performed on PBMCs before and after stimulation with the pro-inflammatory cytokines Tumor Necrosis Factor alpha (TNFα) and Interferon beta (IFNβ). Furthermore, we investigated whether variations in the non-coding genome (eQTLs) affect HLA class I expression. The eQTLs were first analyzed on T cells (169 samples) from the Blueprint consortium and then compared to the ones retrieved in PBMCs of the current study (54 samples), before and after stimulation.

Results

HLA class I expression at the allelic and individual levels

To evaluate and compare the expression of HLA class I genes, RNA from 63 healthy blood donors was isolated from PBMCs and sequenced. Expression levels for the alleles detected in our cohort are provided in Fig 1A. The highest mean expression was measured for HLA-B alleles (1870±541 transcripts per million, tpm) followed by HLA-C (1238±532 tpm) and HLA-A (866±224 tpm) alleles. Variation in expression was also the highest among HLA-B alleles with a maximum fold variation ratio of 7.86 between HLA-B*18:01 and B*56:01 which comprise the highest and lowest tpm values, respectively. The expression of HLA-A alleles was more even with a ratio of 3.2 between the highest (A*01:01) and lowest (A*31:01) tpm values. HLA-C was intermediate with a ratio of 5.2 between the highest (C*04:01) and lowest (C*01:02) tpm values. A more conservative comparison of expression between alleles using median values showed the highest fold ratio for HLA-C (2.7) between C*14:02 and C*03:03, followed by HLA-B (2.0) between B*13:02 and B*56:01 and HLA-A (1.8) between A* 33:03 and A*31:01. Expression according to HLA antigens is shown in S1 Fig.

HLA expression per locus and per donor is represented in Fig 1B. A coordinated expression of both alleles is clearly visible in each heterozygous donor for the three HLA class I genes. The intra-individual variation of expression was the lowest at HLA-A and the highest at HLA-C. The intra-individual variation at HLA-B was also low except in two donors. Along this line, a high correlation between pairs of alleles in heterozygotes at HLA-A, and B and a slightly lower correlations at HLA-C was measured (HLA-A: spearman $\rho = 0.66$ $p = 3.6x10^{-7}$, HLA-B: $\rho = 0.67$ $p = 1.0x10^{-9}$ and HLA-C: $\rho = 0.51$ $p = 1.7x10^{-5}$) (S2 Fig). Calculation of allele specific expression (ASE) as the ratio of the lowest expressed allele towards the total expression of both alleles at a given locus in each individual revealed a very balanced pattern of expression with median values of 0.47, 0.47 and 0.44 for HLA-A, B and C, respectively (Fig 2A). For one individual, the ASE was exceptionally low (i.e., 0.1) due to the low expression of one HLA-B*56:01 allele. Furthermore, in order to test whether the balanced pattern of expression of HLA class I alleles was different from what could be expected by chance, we applied a resampling procedure. The empirical distributions of the simulated ASE obtained after 1000 permutation replicates are shown in Fig 2B (see the figure legend for more details). The observed ASE at

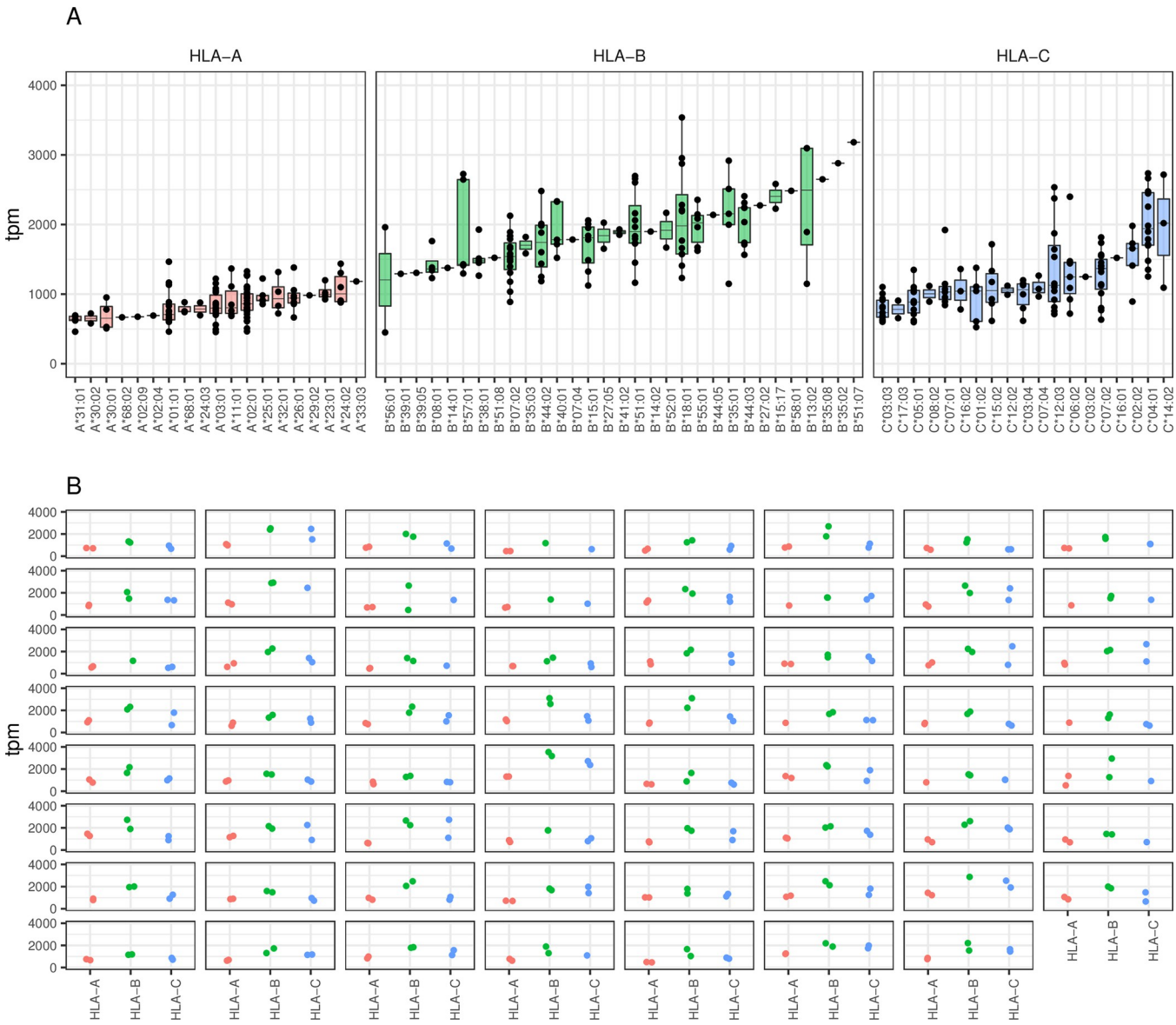

Fig 1. Gradient HLA class I expression at the allelic level. Panel (A) HLA class I RNA expression measured in samples of PBMCs obtained from 63 individuals is plotted as tpm (transcript per million) per allele (indicated on the horizontal axis) for HLA-A, B and C. Nine-teen HLA-A, 31 HLA-B and 19 HLA-C unique alleles are detected. Mean expressions are 866±224, 1870±541and 1238±532 tpm for HLA-A, B and C, respectively. Each dot represents the expression of an allele in one individual. Note that the HLA typing inferred from RNA sequencing corresponded to the available high-resolution typing performed on DNA. Panel (B) Allelic expression is plotted per locus (horizontal axis) and per individual. Each facet represents one individual and each dot one allele.

https://doi.org/10.1371/journal.pgen.1010212.g001

HLA-A, B and C are significantly closer to the maximum possible balance of expression (i.e., 0.5) than any of the replicates (p-values < 0.001).

Upregulated HLA expression

To further investigate the influence of pro-inflammatory cytokines on HLA expression, PBMCs of 56 blood donors were stimulated overnight without or with the cytokines TNFα and IFNβ. Their respective expression was determined and is shown in Fig 3. Mean expression of HLA-A, B and C alleles was upregulated to a similar extent (Fig 3A). Mean fold upregulation

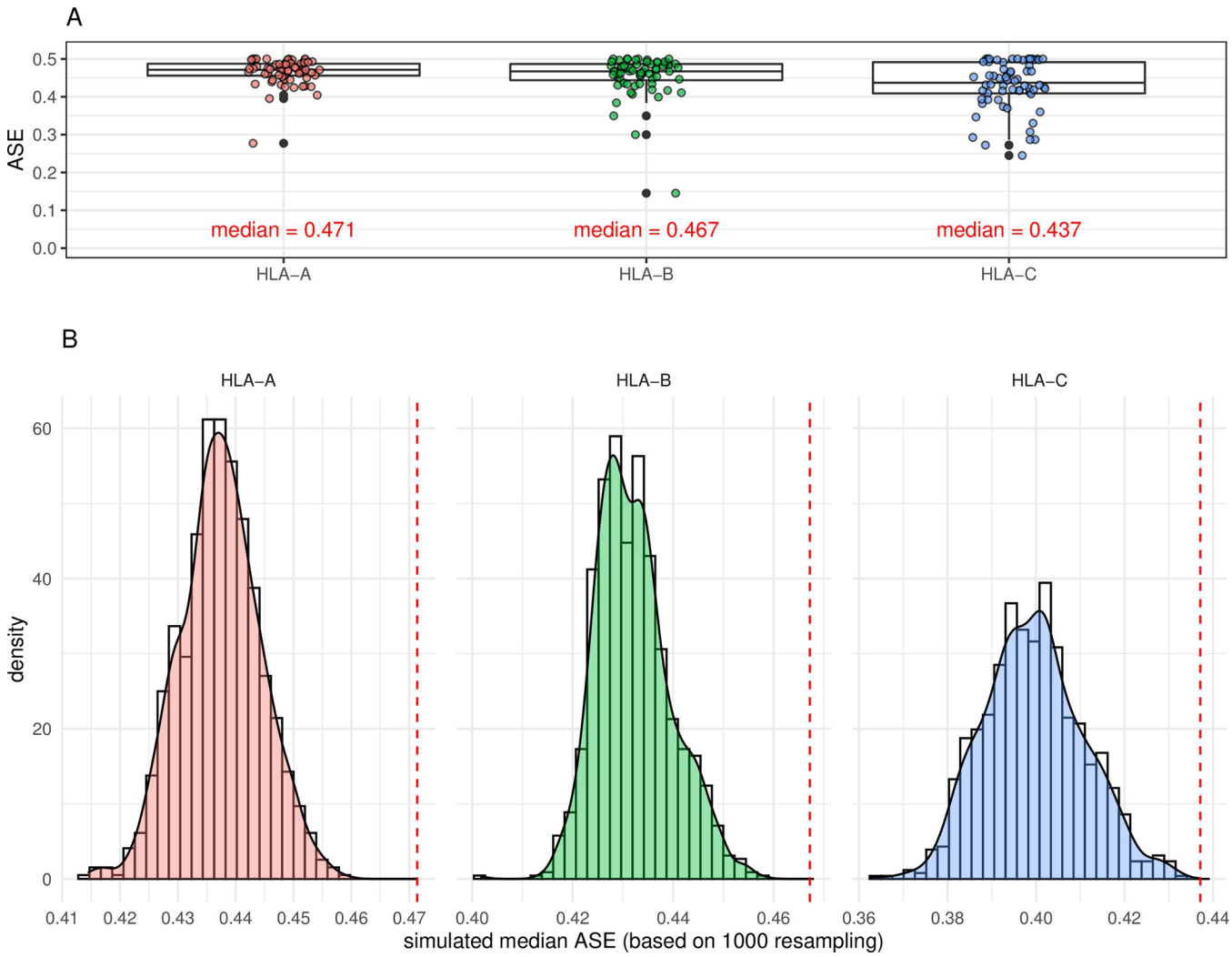

Fig 2. ASE and permutations. Panel (A) Allele specific expression (ASE) of the lowest expressed allele in relation to the expression of both alleles for the three HLA class I genes is plotted. Each dot represents one individual. The observed median is also shown for each locus. Panel (B) Empirical distributions of simulated ASE following a permutation procedure. In this approach, tpm values were shifted randomly among individuals if belonging to the same allele (i.e., permuting the observed tpm values among all carriers of a given allele and doing so for the whole cohort simultaneously). The procedure was replicated 1000 times and for each replicate ASE was computed in every individual for the three loci. The simulated median ASE was computed at each locus and plotted along a density curve and according to a discrete histogram distribution. The dotted red line represents the observed median ASE at a given locus (see panel (A) in comparison to the empirical distribution obtained through the resampling process.

https://doi.org/10.1371/journal.pgen.1010212.g002

was 2.5, 2.04 and 2.11 for the tested HLA-A, B and C alleles, respectively, with spearman ϱ coefficient of 0.4, 0.31 and 0.58 and p-values of $2.9x10^{-4}$, $1.0x10^{-3}$ and $2.2x10^{-16}$. The detailed analysis of the upregulation per allele is shown in S3 Fig and revealed only a few possible outliers. Namely, HLA-A*29:02, 30:01, 32:01 and HLA-C*16:01 had higher fold expression of 4.1, 3.4, 3.5 and 2.7, respectively, while HLA-B* 15:01 had a lower fold expression of 1.3 following stimulation. Otherwise, although the results were gathered from 56 different PBMC cultures performed at different times, a very homogenous and conserved upregulation of expression was observed among alleles. In accordance with RNA expression, cell surface expression of the corresponding HLA-class I molecules was upregulated to a similar extent with a ratio of mean fold upregulation of 1.6, as determined by flow cytometry (Figs 3B and S4). The Spearman correlation for HLA cell surface expression between cells stimulated or non-stimulated with

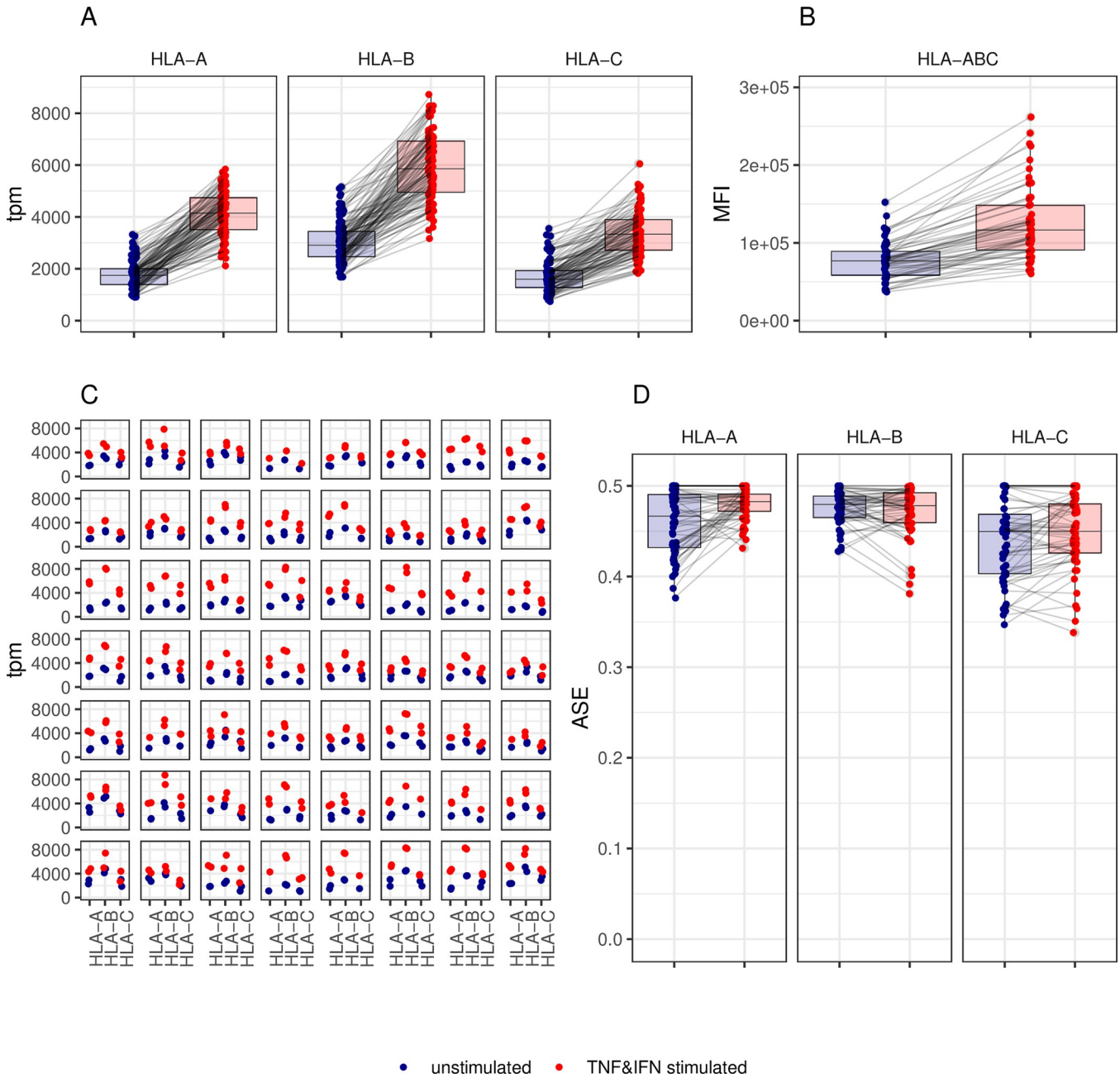

Fig 3. Upregulation of expression induced by TNFα and IFNβ. Panel (A) RNA expression of HLA class I alleles (given as tpm) is plotted in 56 PBMC samples stimulated overnight with (red dots) or without (blue dots) pro-inflammatory cytokines TNFα and IFNβ Each dot represents one allele of one PBMC sample. Panel (B) Cell surface expression of HLA-ABC as determined by flow cytometry is plotted and expressed as MFI for the same PBMCs tested just prior to RNA extraction. Panel (C) Allelic expression is plotted per locus (horizontal axis) and per individual. Each facet represents one individual, each dot one allele. The dots are colored according to the same code as in Panel (A). Panel (D) Allele specific expression (ASE) at the three HLA class I genes is plotted according to both stimulation conditions.

https://doi.org/10.1371/journal.pgen.1010212.g003

cytokines was 0.84 with a p-value of $1.0x10^{-16.}$ The coordinated expression of alleles in heterozygous donors persisted after induced upregulation (Figs 3C and S5). Moreover, the balance of expression among alleles within individuals, as measured by ASE, was similar to the data previously shown in Fig 2 for the uncultured PBMCs, with median ASE comprised between 0.45

and 0.48 in unstimulated and stimulated PBMCs (Fig 3D). Upregulation through overnight cytokines stimulation did not change the much conserved balance of expression for HLA-B and C or even slightly reinforced it as observed for HLA-A (paired Wilcoxon rank sum test, p-value < 0.001, median ASE for unstimulated PBMCs = 0.467, median ASE for stimulated PBMCs = 0.483).

Allelic expression in HLA-A~B~C~DRB1 haplotypes

The RNA expression measured at the allelic level in PBMCs was grouped according to the expected segregation of alleles on three common HLA haplotypes (i.e. inferred at least five times in our cohort and estimated with a frequency above 1% in a large Swiss cohort [34]. The three common haplotypes considered were HLA-A*01:01~B*08:01~C*07:01~DRB1*03:01, HLA-A*03:01~B*07:02~C*07:02~DRB1*15:01 and HLA-A*02:01~B*07:02~C*07:02~DRB1*15:01. The RNA expression of alleles within the common haplotypes in freshly isolated PBMCs or in cells kept in culture without stimulation overnight are shown in the upper panels of Fig 4. A stacked expression was seen across loci for alleles belonging to the same haplotype. In contrast,

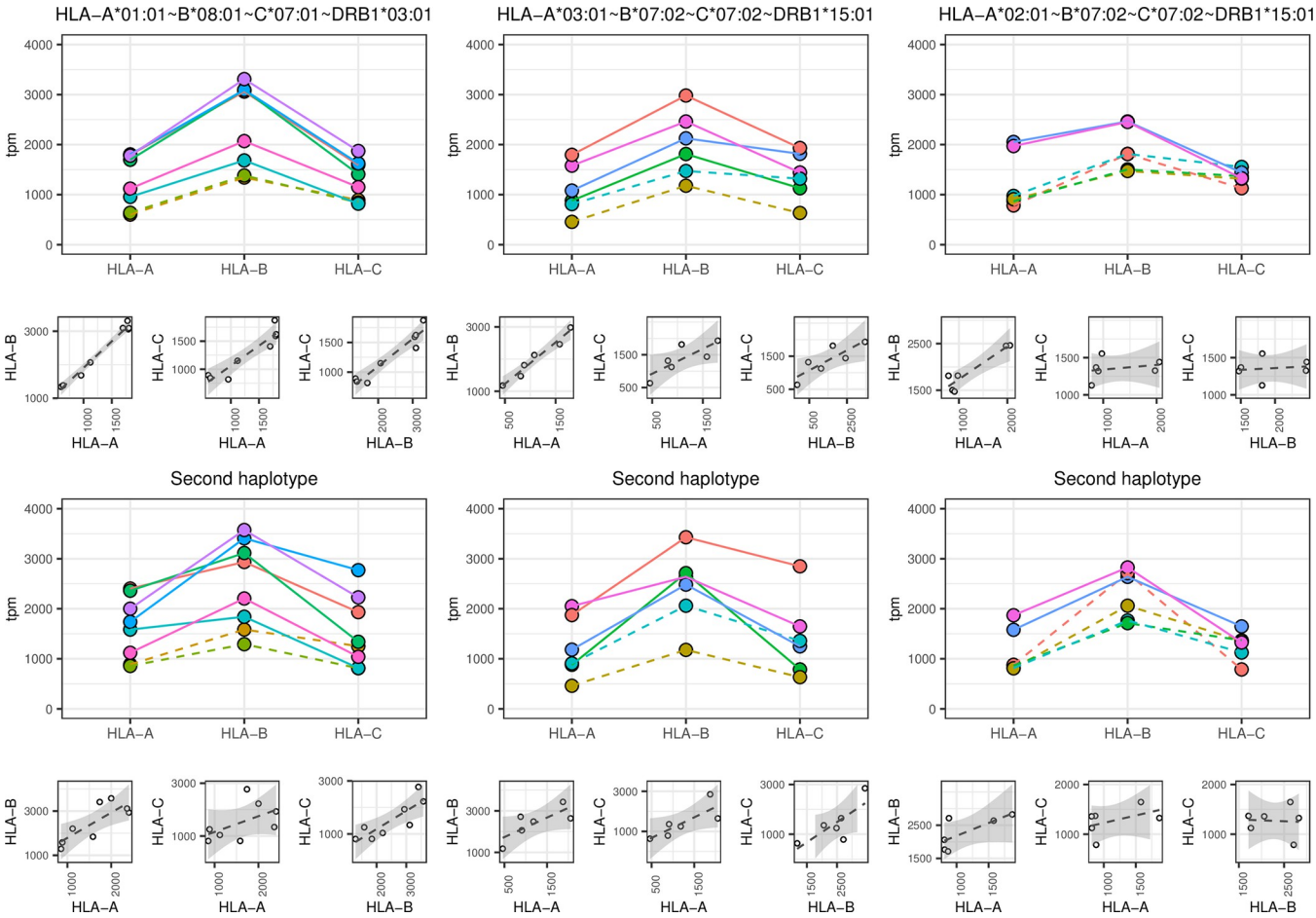

Fig 4. Allelic expression in PBMC carrying common HLA haplotypes. The upper panels represent the RNA expression (in tpm) of alleles segregating on three common HLA haplotypes as indicated. In the lower panels, the RNA expression (in tpm) of the alleles belonging to the second haplotype is shown. Each individual carrying a given haplotype is shown with a different color code. Straight and dotted lines correspond to RNAs of cells kept in culture without stimulation overnight or of freshly isolated PBMCs, respectively. The HLA-typing of the second haplotype differ between samples sharing a common haplotype. Small facets represent the pairwise linear regression of expression between loci. The standard error of the regression is indicated as grey shedding.

https://doi.org/10.1371/journal.pgen.1010212.g004

the corresponding expression of alleles on the second carried haplotype was less concordant as these haplotypes differed among individuals sharing the same common haplotype (Fig 4 lower panels). This is also suggested by the broader standard errors of regression between pairs of loci for the second carried haplotype when compared to the common one (Fig 4, small facets).

Expression quantitative loci analysis of the HLA class I region

Since the PBMC dataset generated in this study has a small sample size, eQTL analyses were carried out also on an external, larger, dataset comprising T cells from the Blueprint consortium (www.blueprint-epigenome.eu).

Genome-wide significant eQTLs (FDR = 5%, Table 1) were identified in 294, 428, and 7502 genes in unstimulated PBMCs, stimulated PBMCs, and Blueprint T cells, respectively.

There are no genome-wide significant eQTLs for any of the HLA class I genes in unstimulated PBMCs. Stimulated PBMCs have a genome-wide significant eQTL for HLA-B (6:31370329:C:A, p = 2.6e-7) and HLA-C (6:31243785:G:T, p = 4.1e-9) genes. In Blueprint T cells, 2, 3, and 2 independent eQTLs for HLA-A, B, and C, were identified respectively (Table 2 and Fig 5). The nominal p-values of the T-cell eQTLs in PBMCs for the best eQTL variant per independent signal are shown in Fig 6A. The p-values of the best eQTL variant in PBMCs that correspond to one of the significant variants in T cell independent signals are shown in Fig 6B. Two SNPs seen as significant in T cells were also seen in unstimulated PBMCs as significant (Fig 6A) and the best SNP's seen in unstimulated and stimulated PBMCs are also seen in T cells (Fig 6B). We tested how well all T-cell eQTLs are replicated in the PBMCs using the π_1 estimates ([35]) and estimated that of the T-cell eQTLs, 34% are active in unstimulated PBMCs, and 38% are active in stimulated PBMCs. To perform the aforementioned π_1 statistic the best eQTL found in stimulated PBMC which corresponds to one of the most significant eQTLs of the Blueprint T-cell were chosen.

The eQTLs identified account for 29%, 13%, and 31% of the variance in HLA-A, B, C expression in unstimulated cells, respectively, and 9%, 23%, and 50% of the variance in stimulated cells. Note that as described in Table 2, there are no significant eQTLs for the unstimulated cells and HLA-A is not significant for the stimulated cells. In order to test whether this variance is mostly independent of the variance of the coding variants that determine an individual's HLA allele, the pairwise correlation was calculated between the eQTL variant and the coding variants for the HLA alleles found in 1000 European genomes of this study. The correlation between the eQTLs and the coding variants that dictate HLA-alleles are low (S6 Fig) indicating that the intra-HLA allele expression variance may be explained by the non-coding variants involved in the regulation of HLA class I gene expression.

Lastly, significant difference in the effect sizes of HLA class I eQTLs between unstimulated and stimulated PBMCs was assessed. We used a linear mixed model to check the interaction between genotype and stimulation status of the samples. We observed that one, two and two eQTLs for HLA-A, B and C, respectively, show significantly higher effect sizes in stimulated cells vs. unstimulated cells (Fig 7). For one signal in HLA-B and all signals in HLA-C significant interaction p-values were observed, indicating that there is a significant difference between the effect sizes of these variant in unstimulated vs. stimulated cells.

Table 1. *Cis* eQTL.

Cell Type	No of Samples	No of Genes	eQTLs FDR 5%
PBMC unstimulated	55	19117	294
PBMC stimulated	55	18737	428
Blueprint T Cells	169	15488	7202

https://doi.org/10.1371/journal.pgen.1010212.t001

Table 2. Best eQTL SNPs shared between T cells and PBMCs.

Gene	HLA-A		HLA-B			HLA-C	
Independent Signal	HLA-A-1	HLA-A- 2	HLA-B-1	HLA-B-2	HLA-B-3	HLA-C-1	HLA-C-2
T cell SNP	6: 29917673:G:A	6: 29923522:C:T	6:31327723:G:A	6:31331829:C:T	6:31378510:G:A	6:31236051:G:A	6:31241002:A:G
T cell p-value	9.42E-14	4.31E-07	3.46E-09	8.81E-07	1.26E-05	5.53E-19	1.87E-08
Unstimulated PBMC p-value	0.1828	0.0316	0.8610	0.8904	0.0239	0.8416	NA
Stimulated PBMC p-value	0.4846	0.6724	0.6432	0.9589	0.6924	0.4203	NA
Independent Signal	HLA-A-0	HLA-A-1	HLA-B-0	HLA-B-1	HLA-B-2	HLA-C-0	HLA-C-1
Best SNP in unstimulated PBMCs [1)]	6:29920713:T:C	6:29918841:G:A	6:31328795:A:T	6:31347798:G:A	6:31381533:C:T	6:31240712:A:C	6:31243785:G:T
Best SNP unstimulated PBMCs p-value	0.0001	0.0097	0.0197	0.1194	0.0239	0.0002	1.90E-05
Best SNP in stimulated PBMCs [1)]	6:29910801:C:A	6:29892854:T:G	6:31334945:T:C	6:31347798:G:A	6:31378510:G:A	6:31274027:C:T	6:31243785:G:T
Best SNP stimulated PBMCs p-value	0.0243	0.1540	0.0034	0.0046	0.6924	2.93E-06	4.11E-09

[1)] the best eQTL in unstimulated or stimulated PBMCs that correspond to T cell eQTL's

https://doi.org/10.1371/journal.pgen.1010212.t002

Discussion

HLA expression at the cell surface might influence the alloreactivity induced in transplantation when T cells recognize non-self HLA. Indeed, Petersdorf et al [15] found an association between highly expressed HLA-C allotypes and increased mortality, while Morishima [16] et al did not. Likewise, high expression of HLA class II (HLA-DPB1) antigens is reported to be risk factor for developing acute graft versus host disease in HSCT [36–38].

In this study, we quantified allelic expression of the three HLA class I genes on an individual basis by RNA sequencing and analyzed whether non-coding genome variations like *cis* eQTL's could explain the observed levels of expression in PBMCs. HLA expression in T cells of the Blueprint Consortium was additionally investigated to increase sample size and the upregulated expression in cytokine-stimulated PBMCs was measured to allow a better balance of the results towards expression. We focused on HLA class I alleles only, as they are constitutively expressed in all PBMCs in contrast to HLA class II alleles, which are only constitutively expressed in specific cell types such as B cells and professional antigen presenting cells. HLA class II allele expression might thus be dependent on subpopulation distribution and cell activation status in PBMCs. Note that Yamamoto et al [29] as well as Johansson et al [30] did not consider this caveat when they analyzed HLA class II expression by RNA-Seq capture methodology in PBMC samples. *Cis* regulating elements associated to HLA-DQB1 allele-specific expression variability after autoimmune T-cell activation were reported by Gutierrrez et al [39]. In contrast to their publication, we were not able to define dynamic ASE (dynASE), which can be associated with HLA expression. We speculate that the power to detect these dynASE sites could be due, on one side, to the difference between HLA class I and HLA class II expression specificity and, on the other side, to the time course experiment they performed with 8 time points, which we do not have in our cohort.

Our results confirm the variability of expression among HLA class I alleles as previously reported [5,14,28,30,40–42]. Nevertheless, considering the mean expression of alleles at the locus level, our data are not in agreement with other publications classifying HLA-C mRNA expression as the lowest [30,43]. Furthermore, the levels of expression that we observe across distinct alleles differ from the values reported in some publications [5,30,40,42], while they concord, with the results published by Garcia-Romano et al [44] and Yamamoto et al [29],

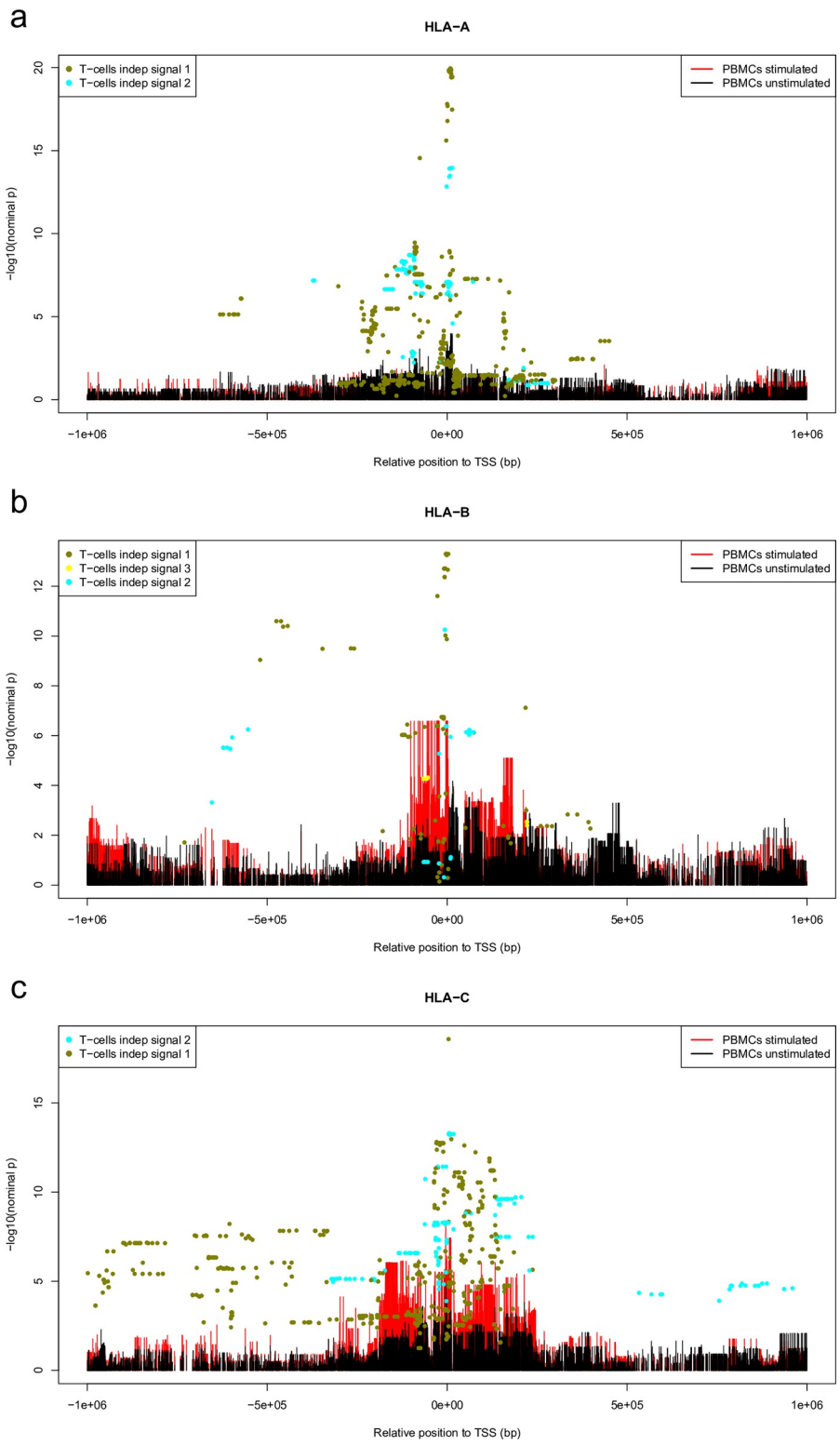

Fig 5. The genomic locations of the eQTLs. The genomic locations of the eQTL associations in the *cis* regions (± 1 Mb from the transcription start site) of HLA-A (**a**), HLA-B (**b**), and HLA-C (**c**). The x-axes are the relative position of the variants to the TSS, and the y-axes represent the significance of the eQTL association (-log10 nominal p-value). The colored bars show all the variants in the region and their eQTL p-values in stimulated and unstimulated PBMC cells. The colored points represent all the variants for each of the genome-wide significant independent eQTL signals in Blueprint T cells.

https://doi.org/10.1371/journal.pgen.1010212.g005

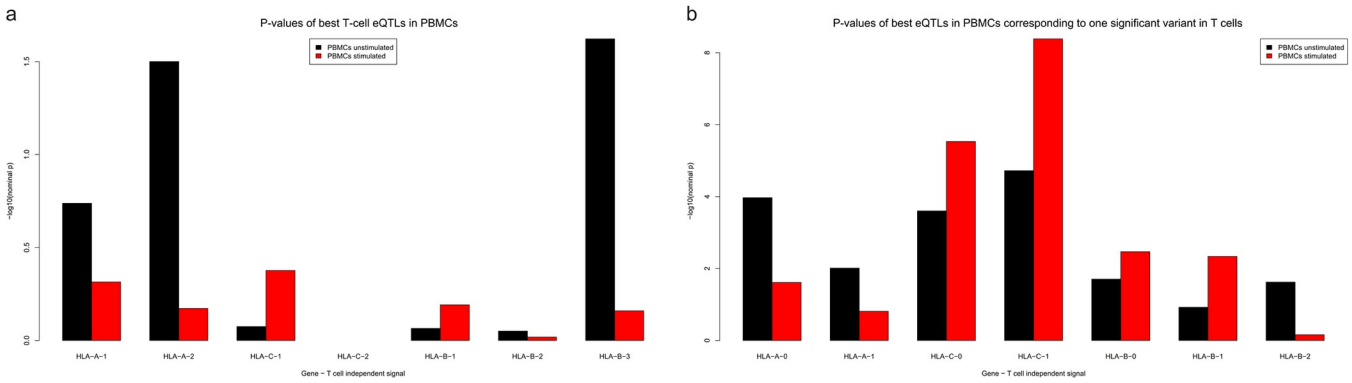

Fig 6. Comparison of best SNPs between T cells, unstimulated PBMCs and stimulated PBMCs. (**a**) The -log10 nominal p-values of the best variant for a T-cell independent signal in PBMCs. (**b**) The -log10 nominal p-values of the best PBMC variant amongst all the significant variants for each of T-cell independent signals.

https://doi.org/10.1371/journal.pgen.1010212.g006

who suggest like us that HLA-B alleles possesses the highest and HLA-A alleles the lowest mean expression, respectively. Due to the limited numbers of samples analyzed, the study was not designed to evaluate the specific expression of alleles at each locus. Lee et al [42] reported expression of allele with more than 10 individuals per allele in the GEUVADIS cohort. Nevertheless, if we classify the alleles as high or low expressers, we obtain similar subgroups to the ones reported by others [4,29,42]. For example, the highest/lowest mean expression among HLA-C alleles was seen for HLA-C*04 and 14 and C*03 (Fig 1), respectively. While the lowest HLA-B expression was seen for the HLA-B*56 allele. This steadiness of expression at the allelic level suggests that allele-specific regulation might exist alongside the individual-dependent regulation of expression discussed in the next paragraph. Indeed, concerning HLA-C alleles, Aguiar et al [28] showed that the rs 41561715-T SNP was associated to the HLA-C*04 lineage in lymphoblastoid cells. In our study, this SNP has a p-value of 3.1175e-06 and 0.0382403 for HLA-C in stimulated and unstimulated PBMCs, respectively. Unfortunately, we cannot comment further on lineage specific effects, as we do not have great enough sample size.

In all three cell cohorts used in this study (i.e., freshly isolated PBMCs, PBMCs kept overnight in culture medium with or without TNFα and IFNβ), we observed a coordinated HLA class I expression at the individual level rather than an allele-specific regulation. Indeed, calculation of allele-specific expression (ASE), measured as the proportion of the lowest expressed allele per locus per individual, revealed a very balanced and conserved expression of both alleles in heterozygotes with ratio close to 0.5 in uncultured cells as well as in cells cultured overnight with or without pro-inflammatory cytokines. The suggested individual regulated expression is sustained by high correlation and a very low variance of expression between paired alleles of the same PBMC sample. A random assignment of allelic pairs by permutations (even when constraining the permutations to keep the same HLA typing within each individual) showed that the observed ASE is significantly closer to the maximum of 0.5 (i.e., representing a fully balanced expression of both alleles) than any of the permutation replicates. The coordinated expression of alleles from distinct HLA class I loci in a given individual was best seen in Fig 3 when alleles belonging to the same haplotype were compared. This confirms our previous observation suggesting an association between HLA-C expression and extended HLA haplotypes [6]. The low number of samples that share the same haplotype(s) did not allow analyzing whether haplotype-specific regulating non-coding elements (eQTL's) exists. Moreover, independently from haplotype segregation, eQTL analysis did not reveal variants affecting the bulk expression of the three HLA class I genes together. Each locus seemed thus to have his

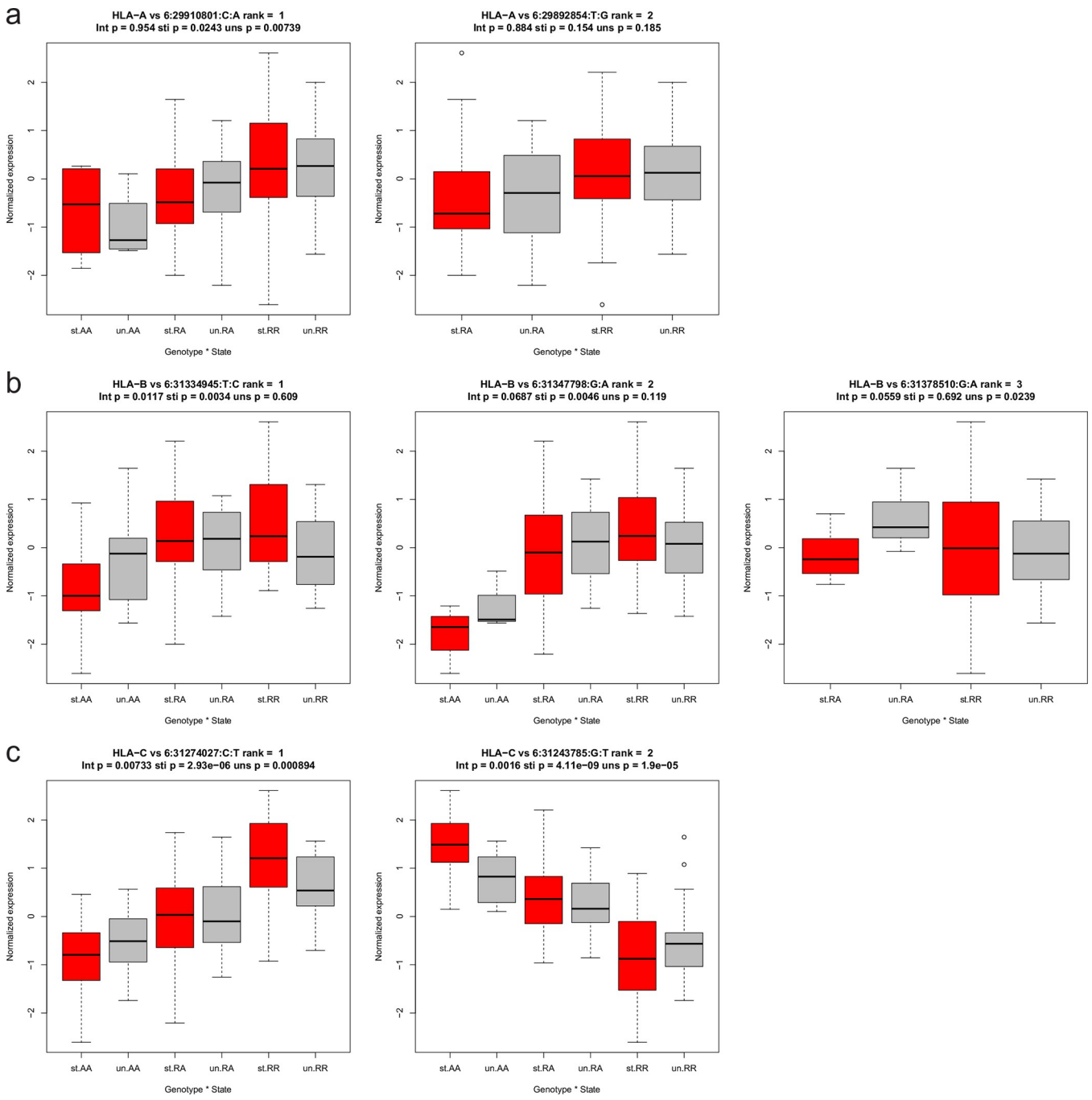

Fig 7. Association between eQTL genotypes and HLA expression in unstimulated versus stimulated PBMCs. Interaction analysis boxplots showing the association between the independent eQTL genotypes (AA = homozygous alternative, RA = heterozygous, RR = homozygous reference) and normalized expression for HLA-A (**a**), HLA-B (**b**), and HLA-C (**c**). The stimulated cells are shown in red, and the unstimulated cells are in grey. The interaction p-value and the p-values in stimulated and unstimulated cells are provided in the titles of the boxplots.

https://doi.org/10.1371/journal.pgen.1010212.g007

own regulatory transcription elements. The coordinated expression at the individual level observed in this study is also in agreement with previous studies of Vandidonck et al [33] and Lam et al. [45], who reported HLA haplotype-specific regulation of gene expression in distinct blastoid cell lines. Gene clustering and *cis* regulatory domains were proposed to explain the

allelic co-expression. [45,46]. However, no higher co-expression of alleles within the same haplotypes compared to alleles on different haplotypes were reported by Aguiar et al [28]. Nevertheless, when they considered allele-specific expression (ASE), as we do in Figs 2 and 3, they observed a rather balanced expression profile for HLA-A, B and C alleles. This sustains our results suggesting the paired expression of HLA class I alleles in most individuals.

We are aware that analysis on different cohorts, namely T cells from the Blueprint consortium and PBMC's represent different cell type composition and might not be directly comparable. Indeed, eQTLs were shown to be possibly related to differences in cell type composition across individuals [47]. Nevertheless, the eQTL association patterns we observed in T cells are similar to the eQTL association patterns in stimulated PMBCs for HLA-B and HLA-C genes. In addition, as seen in Fig 7, two SNPs (6:31274027 and 6: 31243785) were best predictor of HLA C expression. Overall, stimulated cells show stronger effect sizes when compared to unstimulated cells, suggesting that the effect of eQTLs on expression would be magnified in an active immune system.

In conclusion, our data showed that non-coding variations regulating HLA class I genes could be independent of the coding variations that define alleles. As the eQTLs identified in PBMCs mostly have low effect sizes, they may only imperfectly capture the true signals. Thus, confirmatory data on a larger cohort are warranted. More importantly, however, our results demonstrated a coordinated and paired expression of both alleles of the same locus in each individual, which is maintained under conditions of inflammation. Although our study was not designed to answer to this question, the poor correlation between eQTL and distinct HLA allelic expression could suggest that allelic regulation is mediated by sequences, which have not been preferentially selected during evolution. Only a larger sample size may allow the assignment of specific non-coding variants, which would be directly responsible for the intra-allelic variation in expression.

Methodology

Ethic statement

The studies involving human participants were reviewed and approved by Ethical committee of University Hospitals of Geneva (CER 06–208 and 08–208R). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Cells

Peripheral blood mononuclear cells (PBMCs) were purified using standard Ficoll procedure from blood collected from healthy donors who were genotyped at loci HLA-A, B, C, DRB1, DRB3/4/5, DQB1 and DPB1 at high resolution by the Swiss National Reference Laboratory for Histocompatibility (LNRH), while searching for potential unrelated HSC donors.

To upregulate HLA expression, PBMCs were incubated in RPMI 1640 culture medium (Gibco) supplemented with 10mM L-glutamine 100 units/ml penicillin /streptomycin (Gibco) and 10% human AB serum (own preparation) overnight with 50 ng/ml TNFα and 100 ng/ml IFNβ (PrepoTech, London, UK) prior to RNA extraction.

Immunofluorescence

HLA cell surface expression was determined on $CD3^+$ T cells using the monoclonal antibodies APC-labeled anti-human CD3 (clone BW264/56) and FITC-labeled anti-HLA-ABC (clone REA230) (Milteny Biotec) and their corresponding isotype controls. Data acquisition was

performed on gated mononuclear cells, using the ACCURI-C6 cytometer (BD) and the CFLOWPLUS analysis software (BD Bioscience, Allschwil, Switzerland).

DNA extraction and high-resolution HLA typing

DNA was extracted on an automatic system (QIAGEN GmbH, Hilden, Germany) from Ficoll purified peripheral blood mononuclear cells (PBMCs). HLA typing was performed by reverse PCR-sequence-specific oligonucleotide microbeads arrays and high throughput sequencing (One Lambda, Canoga Park, USA).

RNA extraction

Total RNA was extracted using the RNeasy Micro kit (Qiagen, Valencia, CA, USA) according to manufacturer's instructions.

Genotyping and imputation

The genotyping of the samples was conducted with the Illumina Infinium Global Screening Array v2.0. The genotype calls were done using Illumina GenomeStudio 2.0. We filtered variants with minor allele frequency (MAF) less than 1%, genotype missingness greater than 2.5%, and Hardy-Weinberg equilibrium (HWE) p-value less than 1e-5. This resulted in 484886 variants which were imputed using the Michigan Imputation Server with the HRC (Version r1.1 2016) reference panel. Post-imputation filters, r^2 less than 0.3, MAF less than 1%, and HWE p-value less than 1e-6 were applied, resulting in 7924221 variants. Principle component analysis of the samples, together with 1000 genomes samples [48] revealed that 2 samples have non-European ancestry, thus were excluded from further analyses. The human reference genome build used was GRCh37.

RNA-sequencing and quality control

RNA-sequencing was conducted on Illumina HiSeq 4000, according to manufacturer's instructions. The RNA-seq QC was conducted according to Lappalainen T [49]. We also assessed the matching between the genotypes and the RNA-seq experiments using QTL tools mbv [50]. We observed low concordance at heterozygous sites ($< 80\%$) for 6 samples, which were excluded from downstream analyses.

HLA- allele-specific mRNA expression

The RNA sequencing method has previously extensively been discussed by Aguiar et al [28]. Briefly we used HLApers to create a personalized genome based on HLA alleles, mapped the RNA-seq reads against these genomes using STAR [51], and then quantified gene expression using Salmon [52] including allele specific HLA transcripts. The human reference genome build used was GRCh37, and gene annotation used was GENCODE v33 lifted over to b37GENCODE reference annotation for the human and mouse genomes [53]. We excluded genes that were not quantified in more than half of the samples.

eQTL analyses

All analyses were conducted with QTLtools [46]. We calculated population principal components (PCs) from genotypes and technical RNA-seq covariate PCs, using QTLtools pca, and all PCs were centered and scaled. The number of RNA-seq was chosen such that they maximized the number of cis eQTLs genome-wide. We used 5 RNA-seq PCs and 3 population PCs, as

technical covariates in PBMCs, and 30 RNA-seq PCs and 3 population PCs in T cells. Cis eQTL analysis was conducted using QTLtools cis with 1000 permutations.

Graphs

Graphs were generated using R version 4.0.2.

Statistics

Statistical paired *t* tests were performed with R version 4.0.2 and GraphPad prism version 8.0.

Supporting information

S1 Fig. Gradient expression of HLA antigens. HLA class I RNA expression measured in samples of PBMCs obtained from 63 individuals is plotted as tpm (transcript per million). Alleles are grouped according to their serological specificity (i.e., HLA-A, B and C antigens, as indicated on the horizontal axis). Each dot represents the expression of a given allele/antigen in one individual.
(TIF)

S2 Fig. Correlation of expression between pairs of HLA alleles. The RNA expression of pairs of HLA -A, B and C alleles measured in 63 different PBMC samples are plotted as transcript per million (tpm) against each other. The Spearman coefficient ϱ and associated p-value are indicated.
(TIF)

S3 Fig. TNFα/IFNβ induced HLA class I upregulation per allele. HLA class I RNA expression measured in 56 different PBMC samples stimulated with or without TNFα+IFNβ overnight is plotted for each HLA-A, B and C allele taken individually. The numbers in the plots represent median fold upregulation ratios. Alleles are given on the top of each plot.
(TIF)

S4 Fig. Cytofluorometric gating strategy of TNFα/IFNβ induced HLA class I upregulation. HLA cell surface expression on gated $CD3^+$ lymphocytes (upper panels) from PBMC stimulated over night without (grey histograms) or with TNFα/IFNβ (light blue histograms). HLA class I typing of the corresponding PBMC's are: HLA-A*03:01,32:01 HLA*B 08:01,44:03 HLA*C 04:01,07:01 (left lower panel), HLA-A*03:01,24:02 HLA*B 07:02,38:01 HLA*C 07:02,12:03 (middle lower panel), HLA-A*02:01,24:02 HLA*B 39:06,44:02 HLA*C 05:01,07:02 (right lower panel). Corresponding mean fluorescence intensities (MFI) are indicated.
(TIF)

S5 Fig. Correlation of expression between pairs of HLA alleles. The RNA expression of pairs of HLA -A, B and C alleles measured in 56 different PBMC samples stimulated with or without TNFα+IFNβ overnight are plotted against each other. The Spearman coefficient ϱ is indicated.
(TIF)

S6 Fig. The correlation between the non-coding eQTL variants and the coding variants defining alleles. The correlation (r^2, calculated from 1000 genomes European samples) between the best PBMC variants (indicated on the top of the panels) corresponding to a T-cell independent eQTL signal and all the coding variants responsible for an individual's allele type of HLA-A, HLA-B, and HLA-C.
(TIF)

Author Contributions

Conceptualization: Florence Bettens, Emmanouil Dermitzakis, Jean Villard.

Formal analysis: Florence Bettens, Halit Ongen, Guillaume Rey, Stéphane Buhler.

Funding acquisition: Jean Villard.

Investigation: Florence Bettens, Halit Ongen, Zuleika Calderin Sollet.

Methodology: Florence Bettens, Halit Ongen, Guillaume Rey, Stéphane Buhler.

Supervision: Florence Bettens.

Writing – original draft: Florence Bettens, Halit Ongen.

Writing – review & editing: Florence Bettens, Stéphane Buhler, Jean Villard.

References

1. Klein J, Sato A. The HLA system. First of two parts. N Engl J Med. 2000; 343(10):702–9. Epub 2000/09/07. https://doi.org/10.1056/NEJM200009073431006 PMID: 10974135.
2. Parham P. MHC class I molecules and KIRs in human history, health and survival. NatRevImmunol. 2005; 5(3):201–14. https://doi.org/10.1038/nri1570 PMID: 15719024
3. Petersdorf EW. Genetics of graft-versus-host disease: the major histocompatibility complex. Blood Rev. 2013; 27(1):1–12. Epub 2012/11/28. https://doi.org/10.1016/j.blre.2012.10.001 PMID: 23182478; PubMed Central PMCID: PMC3626414.
4. Apps R, Qi Y, Carlson JM, Chen H, Gao X, Thomas R, et al. Influence of HLA-C expression level on HIV control. Science. 2013; 340(6128):87–91. Epub 2013/04/06. https://doi.org/10.1126/science.1232685 PMID: 23559252; PubMed Central PMCID: PMC3784322.
5. Ramsuran V, Hernandez-Sanchez PG, O'HUigin C, Sharma G, Spence N, Augusto DG, et al. Sequence and Phylogenetic Analysis of the Untranslated Promoter Regions for HLA Class I Genes. J Immunol. 2017; 198(6):2320–9. https://doi.org/10.4049/jimmunol.1601679 PMID: 28148735; PubMed Central PMCID: PMC5340644.
6. Bettens F, Brunet L, Tiercy JM. High-allelic variability in HLA-C mRNA expression: association with HLA-extended haplotypes. Genes Immun. 2014; 15(3):176–81. https://doi.org/10.1038/gene.2014.1 PMID: 24500399.
7. Kulkarni S, Qi Y, O'HUigin C, Pereyra F, Ramsuran V, McLaren P, et al. Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. Proc Natl Acad Sci U S A. 2013; 110(51):20705–10. https://doi.org/10.1073/pnas.1312237110 PMID: 24248364; PubMed Central PMCID: PMC3870724.
8. Bubeník J. Tumour MHC class I downregulation and immunotherapy (Review). Oncol Rep. 2003; 10(6):2005–8. Epub 2003/10/10. PMID: 14534734.
9. Esteban F, Concha A, Delgado M, Pérez-Ayala M, Ruiz-Cabello F, Garrido F. Lack of MHC class I antigens and tumour aggressiveness of the squamous cell carcinoma of the larynx. Br J Cancer. 1990; 62(6):1047–51. Epub 1990/12/01. https://doi.org/10.1038/bjc.1990.437 PubMed Central PMCID: PMC1971557. PMID: 2257212
10. Bettens F, Buhler S, Tiercy JM. Allorecognition of HLA-C Mismatches by CD8+ T Cells in Hematopoietic Stem Cell Transplantation Is a Complex Interplay between Mismatched Peptide-Binding Region Residues, HLA-C Expression, and HLA-DPB1 Disparities. Front Immunol. 2016; 7:584. https://doi.org/10.3389/fimmu.2016.00584 PMID: 28018351; PubMed Central PMCID: PMC5151176.
11. Bettens F, Calderin Sollet Z, Buhler S, Villard J. CD8+ T-Cell Repertoire in Human Leukocyte Antigen Class I-Mismatched Alloreactive Immune Response. Front Immunol. 2020; 11:588741. https://doi.org/10.3389/fimmu.2020.588741 PMID: 33552048; PubMed Central PMCID: PMC7856301.
12. Israeli M, Roelen DL, Carrington M, Petersdorf EW, Claas FH, Haasnoot GW, et al. Association between CTL Precursor Frequency to HLA-C Mismatches and HLA-C Antigen Cell Surface Expression. Front Immunol. 2014; 5:547. Epub 2014/11/12. https://doi.org/10.3389/fimmu.2014.00547 PMID: 25386183; PubMed Central PMCID: PMC4209872.
13. Honger G, Krahenbuhl N, Dimeloe S, Stern M, Schaub S, Hess C. Inter-individual differences in HLA expression can impact the CDC crossmatch. Tissue antigens. 2015; 85(4):260–6. https://doi.org/10.1111/tan.12537 PMID: 25786570.

14. Montgomery MC, Liu C, Petraroia R, Weimer ET. Using Nanopore Whole-Transcriptome Sequencing for Human Leukocyte Antigen Genotyping and Correlating Donor Human Leukocyte Antigen Expression with Flow Cytometric Crossmatch Results. The Journal of molecular diagnostics: JMD. 2020; 22 (1):101–10. Epub 2019/11/02. https://doi.org/10.1016/j.jmoldx.2019.09.005 PMID: 31669229.

15. Petersdorf EW, Gooley TA, Malkki M, Bacigalupo AP, Cesbron A, Du Toit E, et al. HLA-C expression levels define permissible mismatches in hematopoietic cell transplantation. Blood. 2014; 124(26):3996–4003. https://doi.org/10.1182/blood-2014-09-599969 PMID: 25323824; PubMed Central PMCID: PMC4271183.

16. Morishima S, Kashiwase K, Matsuo K, Azuma F, Yabe T, Sato-Otsubo A, et al. High-risk HLA alleles for severe acute graft-versus-host disease and mortality in unrelated donor bone marrow transplantation. Haematologica. 2016; 101(4):491–8. https://doi.org/10.3324/haematol.2015.136903 PMID: 26768690.

17. Thomas R, Apps R, Qi Y, Gao X, Male V, O'HUigin C, et al. HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. Nat Genet. 2009; 41(12):1290–4. Epub 2009/11/26. https://doi.org/10.1038/ng.486 PMID: 19935663; PubMed Central PMCID: PMC2887091.

18. Cauli A, Dessole G, Fiorillo MT, Vacca A, Mameli A, Bitti P, et al. Increased level of HLA-B27 expression in ankylosing spondylitis patients compared with healthy HLA-B27-positive subjects: a possible further susceptibility factor for the development of disease. Rheumatology (Oxford). 2002; 41(12):1375–9. Epub 2002/12/07. https://doi.org/10.1093/rheumatology/41.12.1375 PMID: 12468816.

19. van den Elsen PJ. Expression regulation of major histocompatibility complex class I and class II encoding genes. Front Immunol. 2011; 2:48. Epub 2011/01/01. https://doi.org/10.3389/fimmu.2011.00048 PMID: 22566838; PubMed Central PMCID: PMC3342053.

20. Kobayashi KS, van den Elsen PJ. NLRC5: a key regulator of MHC class I-dependent immune responses. Nat Rev Immunol. 2012; 12(12):813–20. Epub 2012/11/24. https://doi.org/10.1038/nri3339 PMID: 23175229.

21. Jongsma MLM, Guarda G, Spaapen RM. The regulatory network behind MHC class I expression. Mol Immunol. 2019; 113:16–21. Epub 2017/12/12. https://doi.org/10.1016/j.molimm.2017.12.005 PMID: 29224918.

22. Carey BS, Poulton KV, Poles A. Factors affecting HLA expression: A review. Int J Immunogenet. 2019. Epub 2019/06/12. https://doi.org/10.1111/iji.12443 PMID: 31183978

23. Dersh D, Phelan JD, Gumina ME, Wang B, Arbuckle JH, Holly J, et al. Genome-wide Screens Identify Lineage- and Tumor-Specific Genes Modulating MHC-I- and MHC-II-Restricted Immunosurveillance of Human Lymphomas. Immunity. 2021; 54(1):116–31.e10. Epub 2020/12/04. https://doi.org/10.1016/j.immuni.2020.11.002 PMID: 33271120; PubMed Central PMCID: PMC7874576.

24. Petersdorf EW, O'HUigin C. The MHC in the era of next-generation sequencing: Implications for bridging structure with function. Hum Immunol. 2019; 80(1):67–78. Epub 2018/10/16. https://doi.org/10.1016/j.humimm.2018.10.002 PMID: 30321633; PubMed Central PMCID: PMC6542361.

25. Ramsuran V, Kulkarni S, O'Huigin C, Yuki Y, Augusto DG, Gao X, et al. Epigenetic regulation of differential HLA-A allelic expression levels. Human molecular genetics. 2015; 24(15):4268–75. https://doi.org/10.1093/hmg/ddv158 PMID: 25935001; PubMed Central PMCID: PMC4492392.

26. Gentle NL, Paximadis M, Puren A, Tiemessen CT. Genetic variability in markers of HLA-C expression in two diverse South African populations. PLoS One. 2013; 8(7):e67780. Epub 2013/07/19. https://doi.org/10.1371/journal.pone.0067780 PMID: 23861805; PubMed Central PMCID: PMC3702582.

27. Boegel S, Lower M, Bukur T, Sorn P, Castle JC, Sahin U. HLA and proteasome expression body map. BMC Med Genomics. 2018; 11(1):36. https://doi.org/10.1186/s12920-018-0354-x PMID: 29587858; PubMed Central PMCID: PMC5872580.

28. Aguiar VRC, César J, Delaneau O, Dermitzakis ET, Meyer D. Expression estimation and eQTL mapping for HLA genes with a personalized pipeline. PLOS Genetics. 2019; 15(4):e1008091. https://doi.org/10.1371/journal.pgen.1008091 PMID: 31009447

29. Yamamoto F, Suzuki S, Mizutani A, Shigenari A, Ito S, Kametani Y, et al. Capturing Differential Allele-Level Expression and Genotypes of All Classical HLA Loci and Haplotypes by a New Capture RNA-Seq Method. Frontiers in Immunology. 2020; 11(941). https://doi.org/10.3389/fimmu.2020.00941 PMID: 32547543

30. Johansson T, Yohannes DA, Koskela S, Partanen J, Saavalainen P. HLA RNA Sequencing With Unique Molecular Identifiers Reveals High Allele-Specific Variability in mRNA Expression. Front Immunol. 2021; 12:629059. Epub 2021/03/16. https://doi.org/10.3389/fimmu.2021.629059 PMID: 33717155; PubMed Central PMCID: PMC7949471.

31. Kulkarni S, Savan R, Qi Y, Gao X, Yuki Y, Bass SE, et al. Differential microRNA regulation of HLA-C expression and its association with HIV control. Nature. 2011; 472(7344):495–8. Epub 2011/04/19. https://doi.org/10.1038/nature09914 PMID: 21499264; PubMed Central PMCID: PMC3084326.

32. Vince N, Li H, Ramsuran V, Naranbhai V, Duh FM, Fairfax BP, et al. HLA-C Level Is Regulated by a Polymorphic Oct1 Binding Site in the HLA-C Promoter Region. Am J Hum Genet. 2016; 99(6):1353–8. Epub 2016/11/08. https://doi.org/10.1016/j.ajhg.2016.09.023 PMID: 27817866; PubMed Central PMCID: PMC5142108.

33. Vandiedonck C, Taylor MS, Lockstone HE, Plant K, Taylor JM, Durrant C, et al. Pervasive haplotypic variation in the spliceo-transcriptome of the human major histocompatibility complex. Genome Res. 2011; 21(7):1042–54. Epub 2011/06/02. https://doi.org/10.1101/gr.116681.110 PMID: 21628452; PubMed Central PMCID: PMC3129247.

34. Buhler S, Baldomero H, Ferrari-Lacraz S, Nunes JM, Sanchez-Mazas A, Massouridi-Levrat S, et al. High-resolution HLA phased haplotype frequencies to predict the success of unrelated donor searches and clinical outcome following hematopoietic stem cell transplantation. Bone Marrow Transplant. 2019; 54(10):1701–9. Epub 2019/04/07. https://doi.org/10.1038/s41409-019-0520-6 PMID: 30953025; PubMed Central PMCID: PMC7198472.

35. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A. 2003; 100(16):9440–5. Epub 2003/07/29. https://doi.org/10.1073/pnas.1530509100 PMID: 12883005; PubMed Central PMCID: PMC170937.

36. Petersdorf EW, Malkki M, O'HUigin C, Carrington M, Gooley T, Haagenson MD, et al. High HLA-DP Expression and Graft-versus-Host Disease. The New England journal of medicine. 2015; 373(7):599–609. https://doi.org/10.1056/NEJMoa1500140 PMID: 26267621; PubMed Central PMCID: PMC4560117.

37. Petersdorf EW, Bengtsson M, De Santis D, Dubois V, Fleischhauer K, Gooley T, et al. Role of HLA-DP Expression in Graft-Versus-Host Disease After Unrelated Donor Transplantation. J Clin Oncol. 2020; 38(24):2712–8. https://doi.org/10.1200/JCO.20.00265 PMID: 32479188; PubMed Central PMCID: PMC7430213.

38. Stéphane Buhler HB, Sylvie Ferrari-Lacraz, Anne-Claire Mamez, Stavroula Masouridi-Levrat, Dominik Heim, Jörg Halter, Gayathri Nair, Yves Chalando, Urs Schanz, Tayfun Güngör, Grazia Nicoloso, Passweg Jakob R, Jean Villard* Analysis of biological models to predict clinical outcomes based on HLA-DPB1 disparities in unrelated transplantation. Blood Advances. 2021. https://doi.org/10.1182/bloodadvances.2020003998 PMID: 34448833

39. Gutierrez-Arcelus M, Baglaenko Y, Arora J, Hannes S, Luo Y, Amariuta T, et al. Allele-specific expression changes dynamically during T cell activation in HLA and other autoimmune loci. Nat Genet. 2020; 52(3):247–53. Epub 2020/02/19. https://doi.org/10.1038/s41588-020-0579-4 PMID: 32066938; PubMed Central PMCID: PMC7135372.

40. Pan N, Lu S, Wang W, Miao F, Sun H, Wu S, et al. Quantification of classical HLA class I mRNA by allele-specific, real-time polymerase chain reaction for most Han individuals. Hla. 2018; 91(2):112–23. Epub 2017/11/28. https://doi.org/10.1111/tan.13186 PMID: 29178661.

41. Rene C, Lozano C, Villalba M, Eliaou JF. 5' and 3' untranslated regions contribute to the differential expression of specific HLA-A alleles. Eur J Immunol. 2015; 45(12):3454–63. https://doi.org/10.1002/eji.201545927 PMID: 26399450.

42. Lee W, Plant K, Humburg P, Knight JC. AltHapAlignR: improved accuracy of RNA-seq analyses through the use of alternative haplotypes. Bioinformatics (Oxford, England). 2018; 34(14):2401–8. Epub 2018/03/08. https://doi.org/10.1093/bioinformatics/bty125 PMID: 29514179; PubMed Central PMCID: PMC6041798.

43. Johnson DR. Differential expression of human major histocompatibility class I loci: HLA-A, -B, and -C. Hum Immunol. 2000; 61(4):389–96. Epub 2000/03/15. https://doi.org/10.1016/s0198-8859(99)00186-x PMID: 10715516.

44. Garcia-Ruano AB, Mendez R, Romero JM, Cabrera T, Ruiz-Cabello F, Garrido F. Analysis of HLA-ABC locus-specific transcription in normal tissues. Immunogenetics. 2010; 62(11–12):711–9. Epub 2010/09/16. https://doi.org/10.1007/s00251-010-0470-z PMID: 20842357.

45. Lam TH, Shen M, Tay MZ, Ren EC. Unique Allelic eQTL Clusters in Human MHC Haplotypes. G3 (Bethesda). 2017; 7(8):2595–604. https://doi.org/10.1534/g3.117.043828 PMID: 28600441; PubMed Central PMCID: PMC5555465.

46. Delaneau O, Ongen H, Brown AA, Fort A, Panousis NI, Dermitzakis ET. A complete tool set for molecular QTL discovery and analysis. Nat Commun. 2017; 8:15452. Epub 2017/05/19. https://doi.org/10.1038/ncomms15452 PMID: 28516912; PubMed Central PMCID: PMC5454369.

47. Võsa U, Esko T, Kasela S, Annilo T. Altered Gene Expression Associated with microRNA Binding Site Polymorphisms. PloS one. 2015; 10(10):e0141351. Epub 2015/10/27. https://doi.org/10.1371/journal.pone.0141351 PMID: 26496489; PubMed Central PMCID: PMC4619707.

48. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015; 526(7571):68–74. Epub 2015/10/04. https://doi.org/10.1038/nature15393 PMID: 26432245; PubMed Central PMCID: PMC4750478.

49. Lappalainen T, Sammeth M, Friedländer MR, 't Hoen PAC, Monlong J, Rivas MA, et al. Transcriptome and genome sequencing uncovers functional variation in humans. Nature. 2013; 501:506. https://doi.org/10.1038/nature12531 PMID: 24037378

50. Fort A, Panousis NI, Garieri M, Antonarakis SE, Lappalainen T, Dermitzakis ET, et al. MBV: a method to solve sample mislabeling and detect technical bias in large combined genotype and sequencing assay datasets. Bioinformatics. 2017; 33(12):1895–7. Epub 2017/02/12. https://doi.org/10.1093/bioinformatics/btx074 PMID: 28186259; PubMed Central PMCID: PMC6044394.

51. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29(1):15–21. Epub 2012/10/30. https://doi.org/10.1093/bioinformatics/bts635 PMID: 23104886; PubMed Central PMCID: PMC3530905.

52. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nature methods. 2017; 14(4):417–9. Epub 2017/03/07. https://doi.org/10.1038/nmeth.4197 PMID: 28263959; PubMed Central PMCID: PMC5600148.

53. Uszczynska-Ratajczak B, Lagarde J, Frankish A, Guigó R, Johnson R. Towards a complete map of the human long non-coding RNA transcriptome. Nature reviews Genetics. 2018; 19(9):535–48. Epub 2018/05/26. https://doi.org/10.1038/s41576-018-0017-y PMID: 29795125; PubMed Central PMCID: PMC6451964.