



Thèse

2024

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

KIR repertoire diversity and its implication in clinical course after allogeneic hematopoietic stem cell transplantation

Schaefer, Antonia Marie

How to cite

SCHAEFER, Antonia Marie. KIR repertoire diversity and its implication in clinical course after allogeneic hematopoietic stem cell transplantation. Doctoral Thesis, 2024. doi: 10.13097/archive-ouverte/unige:180365

This publication URL: <https://archive-ouverte.unige.ch/unige:180365>

Publication DOI: [10.13097/archive-ouverte/unige:180365](https://doi.org/10.13097/archive-ouverte/unige:180365)

Section de médecine clinique

Département de médecine

Service de néphrologie

Thèse préparée sous la direction du Professeur Jean Villard

**KIR repertoire diversity and its implication in clinical course after allogeneic
hematopoietic stem cell transplantation**

Thèse

Présentée à la Faculté de Médecine

de l'Université de Genève

pour obtenir le grade de Docteur en Science Médicales (MD-PhD)

par

Antonia Marie SCHÄFER

de

Lausanne (VD), Suisse

Thèse n°57

Genève

2024

DOCTORAT EN SCIENCES MEDICALES

« MD-PhD »

Thèse de :

Antonia Marie SCHÄFER

Originaire de Lausanne, Suisse

Intitulée :

« KIR repertoire diversity and its implication in clinical course after allogeneic hematopoietic stem cell transplantation »

La Faculté de médecine, sur les préavis du Comité directeur du MD-PhD, autorise l'impression de la présente thèse, sans prétendre par là émettre d'opinion sur les propositions qui y sont énoncées.

Genève, le 11 septembre 2024

Thèse n° **57**



Pr Antoine Geissbuhler
Doyen

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives à la présentation des thèses de doctorat à l'Université de Genève".

**KIR repertoire diversity and its implication in clinical course after allogeneic
hematopoietic stem cell transplantation**

by

Antonia Marie Schäfer

Transplant immunology unit

Faculty of medicine, Department of medicine

University of Geneva

Geneva, Switzerland

Principal supervisor:

Prof. Jean Villard, MD, PhD

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

Thesis co-committee:

Prof. Sophie De Seigneux, MD, PhD

Service and Laboratory of Nephrology, Department of Internal Medicine Specialties and of Physiology and Metabolism, Geneva University Hospitals

Prof. Yves Chalandon, MD

Service of Haematology, Department of Oncology, Geneva University Hospitals

External expert:

Prof. Kay Poulton, PhD FRCPATH

Department of Histocompatibility and Immunogenetics, Transplantation Laboratory, Manchester Royal Infirmary, Manchester University NHS Foundation Trust

RÉSUMÉ

La greffe de cellules souches hématopoïétiques vise à restaurer un nouveau système hématopoïétique et à exercer un effet bénéfique du greffon contre la tumeur. Le succès de cette thérapie cellulaire est cependant encore considérablement entravé par l'apparition de complications immunologiques, infectieuses et surtout de la rechute de la maladie initiale.

Les cellules NK (Natural Killer) sont des lymphocytes du système inné ayant une fonction antitumorale et antivirale. Les capacités d'immunosurveillance et de cytotoxicité de ces cellules sont déclenchées par l'absence du HLA de classe I du soi « missing-self » médiée par leurs récepteurs KIR (Killer-cell immunoglobulin-like receptors). Dans le cadre d'une transplantation, des configurations du donneur et du receveur avec un mismatch KIR et HLA possèdent le potentiel d'induire une alloréactivité des cellules NK contre des cibles tumorales et virales. De nombreuses études ont tenté de prédire cette alloréactivité sur la base des facteurs génétiques KIR et HLA, cependant avec de nombreux résultats contradictoires. La complexité génétique du système KIR, qui englobe des variations dans le type de gène, le nombre de copies et un polymorphisme allélique étendu, ainsi que le manque de résolution et de rendement du séquençage, ont toutefois entravé les recherches dans cette direction. La première partie de cette thèse vise à investiguer le pouvoir prédictif du génotypage KIR à haute résolution pour informer le potentiel d'alloréactivité des cellules NK après la transplantation et analyser son impact sur les complications post-transplantation. Nous avons pu démontrer que les interactions fonctionnelles KIR2DS4 ont un effet néfaste sur plusieurs complications post-transplantation. Un impact différentiel sur le risque de rechute dépendait de la force d'interaction entre le KIR3DL1 et HLA-Bw4 ainsi que de la présence d'interactions KIR2DS1.

Outre le contexte génétique, des facteurs non génétiques se sont révélés être des inducteurs tout aussi importants des variations du système immunitaire. Cependant, la susceptibilité différentielle et le moment de la modulation immunitaire par ces facteurs dans la reconstitution du système immunitaire n'ont pas encore été entièrement élucidés. La deuxième partie de cette thèse vise à disséquer la reconstitution du répertoire des cellules NK et du TCR (T cell receptor) après la transplantation et analyser l'impact d'événements cliniques sur la trajectoire de reconstitution. En utilisant de l'immunophénotypage et du séquençage à haute résolution, nous avons démontré que ces répertoires s'établissent précocement après la transplantation et qu'ils subissent des changements inflationnistes et coordonnés lors d'une réactivation du cytomégalovirus.

En résumé, cette thèse est une tentative de contribuer au besoin urgent d'améliorer les taux de réussite des transplantations de cellules souches en explorant le potentiel du système immunogénétique KIR pour affiner l'algorithme de sélection des donneurs et en améliorant les connaissances biologiques sur le rétablissement complexe des compartiments immunitaires post-transplantation.

ABSTRACT

An allogeneic hematopoietic stem cell transplantation (HSCT) aims to restore a new hematopoietic system and exert a beneficial graft-versus-tumor (GVT) effect. The success of this cellular therapy is, however, still significantly hampered by the occurrence of immunological, infectious complications and especially disease relapse.

NK cell immunosurveillance and killing capabilities are triggered by an HLA class I missing-self situation mediated by their germline-encoded killer cell immunoglobulin-like receptors (KIR), which can be harnessed in a transplant setting. Donor and recipient KIR and HLA configurations mimicking a state of missing ligand are thought to unleash NK cell alloreactivity against tumor and viral target cells post-transplant. The genetic complexity of the KIR system encompassing gene, copy number and allelic variation and the lack of sequencing resolution depth and scalability have however challenged research investigations in this direction, with many conflicting results reported so far. To tackle this issue, the first part of this thesis interrogates the predictive power of high-resolution KIR and HLA genotyping to inform post-transplant NK cell alloreactivity and analyze its impact on transplant outcomes. We could demonstrate that functional KIR2DS4 interactions have a detrimental effect on various transplant outcomes. A differential impact on the relapse propensity was driven by the strength of KIR3DL1 – HLA-Bw4 as well as the presence of KIR2DS1 interactions.

Besides the genetic background, non-genetic factors have been shown as equally important inducers of immune system variations. However, the differential susceptibility and timing to immune modulation by non-genetic factors in reconstituting immune systems have not been fully understood yet. The second part of this thesis aims to dissect the temporal restoration of the NK cell and TCR repertoire post-transplant and trace relations between immune state shifts and clinical events in the unprecedented case study of HSCT. Leveraging high-dimensional immunophenotyping and next-generation sequencing, we showed that these repertoires establish at an early timepoint post-transplant and undergo coordinated inflationary changes upon CMV reactivation.

Altogether, this thesis is an attempt to contribute to the pressing need of improving HSCT success rates by exploring the potential of the KIR immunogenetic system to refine the donor selection algorithm and by increasing learnings on the complex re-establishment of immune compartments and its interplay between immune regulators.

CONTENTS

ABBREVIATIONS.....	5
ACKNOWLEDGEMENTS.....	8
CHAPTER I General scientific background.....	10
1. Basic NK cell biology	10
2. Basic T cell biology	22
3. HSCT as a therapeutic option for immune and hematological related disorders.....	29
Thesis aims.....	33
CHAPTER II High-resolution KIR genotyping in hematopoietic stem cell transplantation.....	35
1. Current state of research.....	36
2. Manuscript I	41
CHAPTER III Comparative assessment of mass cytometry and spectral flow cytometry	69
1. Current state of research.....	81
2. Manuscript II	84
CHAPTER IV NK and T cell repertoire after hematopoietic stem cell transplantation	109
1. Current state of research.....	110
2. Manuscript III	114
PERSPECTIVES	171
REFERENCES	175

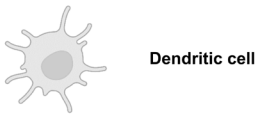
ABBREVIATIONS

ADCC	antibody-dependant cellular cytotoxicity
aGvHD	acute graft-versus-host disease
ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
APC	antigen-presenting cell
ATG	anti-thymocyte globulin
BM	bone marrow
BMF	bone marrow failure
CB	cord blood
CD	cluster of differentiation
CD159	C-type lectin superfamily
CDR3	complementarity determining region
CDR3aa	CDR3 amino acid
cGVHD	chronic graft-versus-host disease
CLL	chronic lymphoblastic leukemia
CML	chronic myeloid leukemia
CMV	cytomegalovirus
CNV	copy number variation
CSB	cell staining buffer
CTL	cytotoxic T lymphocytes
CyTOF	cytometry by time - of - flight
D	donor
DC	dendritic cells
ddNTP	dideoxynucleotides triphosphates
DMSO	dimethylsulfoxide
dNTP	deoxynucleotides triphosphates
EBV	epstein-barr virus
FasL	fas ligand
FCS	fetal calf serum
FMO	fluorescence minus one
GrzmB	granzyme B

GvHD	graft-versus-host disease
GvL	graft-versus-leukemia
GvT	graft-versus-tumor
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HLA	human leucocyte antigen
HSCT	hematopoietic stem cell transplantation
HSV-2	herpes simplex virus
ICP	inductively coupled plasma
IFN γ	interferon γ
Ig	immunoglobulin
IL	interleukin
IMGT	international immunogenetics information system
IPD-KIR	immuno polymorphism database - killer-cell immunoglobulin-like receptor
IQR	interquartile range
ITAM	tyrosine-based activation motif
ITIM	tyrosine-based inhibitory motif
KIR	killer cell immunoglobulin-like receptor
KLRG1	killer cell lectin-like receptor G1
LAG-3	lymphocyte activation gene 3
LIR	leukocyte immunoglobulin-like receptor
Ln	lanthanide
MAC	myeloablative conditioning
McPAS	manually curated catalogue of pathology-associated TCR sequences
MDS	myelodysplastic syndrome
MHC	major histocompatibility complex
MHC-I	major histocompatibility complex class I
MICA	major histocompatibility complex class I chain-related protein A
MICB	major histocompatibility complex class I chain-related protein B
MPN	myeloproliferative neoplasms
NCR	natural cytotoxicity receptors
NGS	next-generation sequencing
NHL	non-hodgkin lymphoma
NK	natural killer

NKG2	natural killer group 2
NRM	non-relapse mortality
OS	overall survival
PBMC	peripheral blood mononuclear cells
PBS	phosphate-buffered saline
PBSC	Peripheral blood stem cell
PCD	plasma cell disorder
Pd	palladium
PD-1	programmed cell death protein 1
PFS	progression-free survival
PID	primary immunodeficiencies
PING	pushing Immunogenetics to the next generation
PMA	phorbol 12-myristate 13-acetate
pMHC	peptide-MHC
PMT	photomultiplier tubes
PTCy	post-transplant cyclophosphamide
PVR	polio virus receptor
R	recipient
RIC	reduced intensity conditioning
RSV	respiratory syncytial virus
Sars-Cov-2	severe acute respiratory syndrome coronavirus 2
SLO	secondary lymphoid organs
T	timepoint
TBI	total body irradiation
TCR	T cell receptor
TIGIT	T cell immunoreceptor with Ig and ITIM domains
TNF α	tumor necrosis factor α
TOF	time-of-flight
TRAIL	tumor necrosis factor-related apoptosis-inducing ligand
ULBP	UL16-binding proteins

Figure legend



ACKNOWLEDGEMENTS

There have been numerous people without whom this thesis couldn't have been possible and all of whom I would like to express my deepest gratitude.

First and foremost, I would like to deeply thank my principal supervisor, Prof. Jean Villard, for believing in me and sticking with me these past three years, providing invaluable mentorship and guidance in conducting research. I also thank him for the numerous opportunities to present my work at various conferences and to interact with other scientists through various collaborations around the globe.

I would like to express my gratitude to the members of my MD-PhD committee, Prof. De Seigneux, Prof. Chalandon and Prof. Poulton for valuable advice and input.

A huge thanks to all the members of Prof. Villard's research group: Florence Bettens, Zuleika Pythoud, Marie-Priscille Hervé and Stéphane Bühler, thank you for all the teaching, the practical help in experimental work, the support in solving bioinformatic issues and the interesting scientific and non-scientific discussions. It was a great pleasure working with you on a daily basis. Thanks, as well to Valérie Olivier for the final encouragements and sharing the same enthusiasm for the KIRs.

A big thanks to Prof. Norman and his research group, in particular to Ticiana Farias and Katie Kichula, for the warm welcome in Denver and for teaching me everything about the KIRs. It has been a blast working with you.

Many thanks to Miguel Garcia and Sènan d'Almeida from the flow cytometry core facility at the EPFL Campus in Lausanne for the very fun, insightful and fruitful collaborative work.

Many thanks to the core facilities at the Geneva University hospital, notably the flow cytometry core facility and the genomic platform for the help and guidance in the experiment acquisition.

Finally, I would like to deeply thank my family and friends for the continuous support and all the great adventures experienced together that kept me motivated during this thesis.

CHAPTER I General scientific background

1. Natural killer cell biology

Harbored with a broad array of germline-encoded activating and inhibitory receptors, Natural Killer (NK) cells are innate lymphocytes capable of recognizing and killing transformed cells, such as virally infected or cancerous cells, without prior sensitization. They are specifically trained in recognizing abnormal cells that downregulate major histocompatibility complex class I (MHC-I) expression through a missing-self mechanism mediated by their killer-cell immunoglobulin-like receptors (KIR). Additionally, they are sensitive to the upregulation of stress ligands through their activating receptors – a mechanism called induced-self. The net input of inhibitory and activating signals as a result of these engagements or missing interactions will determine the NK cell's fate and responsiveness. Upon target identification, NK cells eliminate cells based on different mechanisms involving cell to cell interactions and paracrine signaling with the formation of so-called immunological synapses to stabilize and render cell effector functions (Fig.1): The first pathway involves CD16 – a NK cell activating receptor – whose engagement triggers antibody-dependent cellular cytotoxicity (ADCC) [1-3]. Interaction of death receptors such as Fas-ligand (FasL) and TNF-related apoptosis-inducing ligand (TRAIL) with their ligands on the target cell represents another contact-dependent cytotoxic mechanism [4, 5].

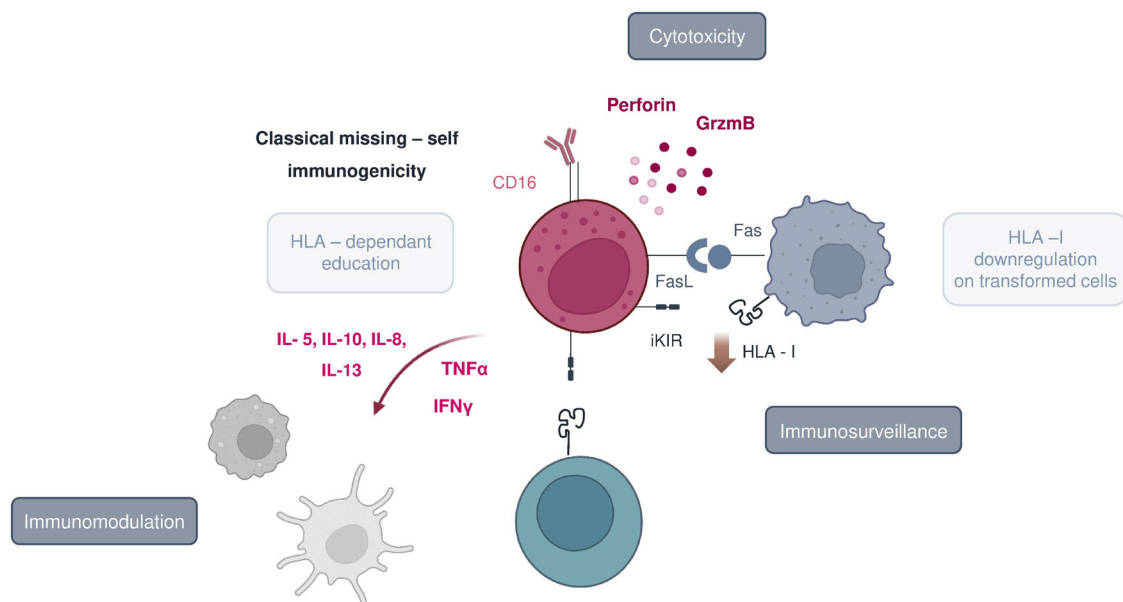


Figure 1: NK cell recognition, cytotoxic and immunomodulatory functions.

Finally, NK cell cytotoxicity involves the targeted release of perforin and granzyme containing lytic granules, similarly to cytotoxic T lymphocytes (CTL) [6]. Besides cytotoxic functionalities, NK cells are known to be sensitive to cytokines such as interleukin (IL)-2, IL-15, IL-18, IL-21 and in return, are also able to release a cascade of various immunomodulatory cytokines. Indeed, CD16 ligation promotes the release of pro-inflammatory cytokines (type II) such as interferon γ (IFN γ) and tumor necrosis factor α (TNF α) and induces the expression of IL-2R (CD25) on the cell surface [7, 8]. These key immunoregulatory factors shape the adaptive immune response by promoting the activation of dendritic cells (DC) and CTL. The secretion of cytokines in non-cell contact cytotoxicity is less well studied [9].

1.1 NK cell development and maturation

NK cell ontogenesis *in utero* predominantly takes place in the fetal liver and is then progressively taken over by the bone marrow (BM) at the end of gestation. While this anatomical site has for long been believed to be the primary site of hematopoiesis during adult life, recent findings point towards a dichotomic division of ontogenesis in which the latter steps of NK cell differentiation are carried over to secondary lymphoid organs (SLO) such as the spleen, tonsils and lymph nodes, where an enrichment in immature CD56^{bright} NK cells has been identified. Evidence of extra-medullary sites for NK cell development has been further corroborated by studies demonstrating the presence of CD34⁺ hematopoietic progenitors that originate in the BM and migrate towards SLO for further differentiation [10-13]. Whether NK cell ontogeny has an obligate pass towards SLO, whether SLO comprise a backup reservoir for immature NK cells that are released upon peripheral disturbances is still unresolved. Besides circulating in the blood stream at an approximate frequency of 5 to 15%, NK cells are well known to be widely dispersed throughout body tissues such as the lung, liver, kidneys, lymph nodes, thymus, and uterus. The developmental origin of tissue resident and circulating NK cells remains a major outstanding question in NK cell biology, especially given their distinctive phenotypic and transcriptional characteristics. The notion of peripheral tissue ontogeny has received growing fundament in recent years with the identification of CD34⁺ precursors bearing the capacity to differentiate into tissue resident NK cells in the liver, decidua, thymic and gut tissues [14, 15]. Recent data also suggests a pathway in which circulating NK cells are able to differentiate into long-lived resident NK cells upon recruitment to an infectious site [16]. The developmental steps of NK cell ontogenesis are currently a matter of debate among the scientific community. While the first originated model suggests linearity in the maturation based on the acquisition and downregulation of CD56 (Fig.2), recent studies have challenged this model, supporting a more branched developmental trajectory [17]. The adoption of omics-technologies has fine-tuned our understanding of NK cell development and their populational diversity in the peripheral circulation and tissues [11, 18, 19]. Recent detailed publications provide a great overview of current findings, perspectives and outstanding questions on this topic [20-22].

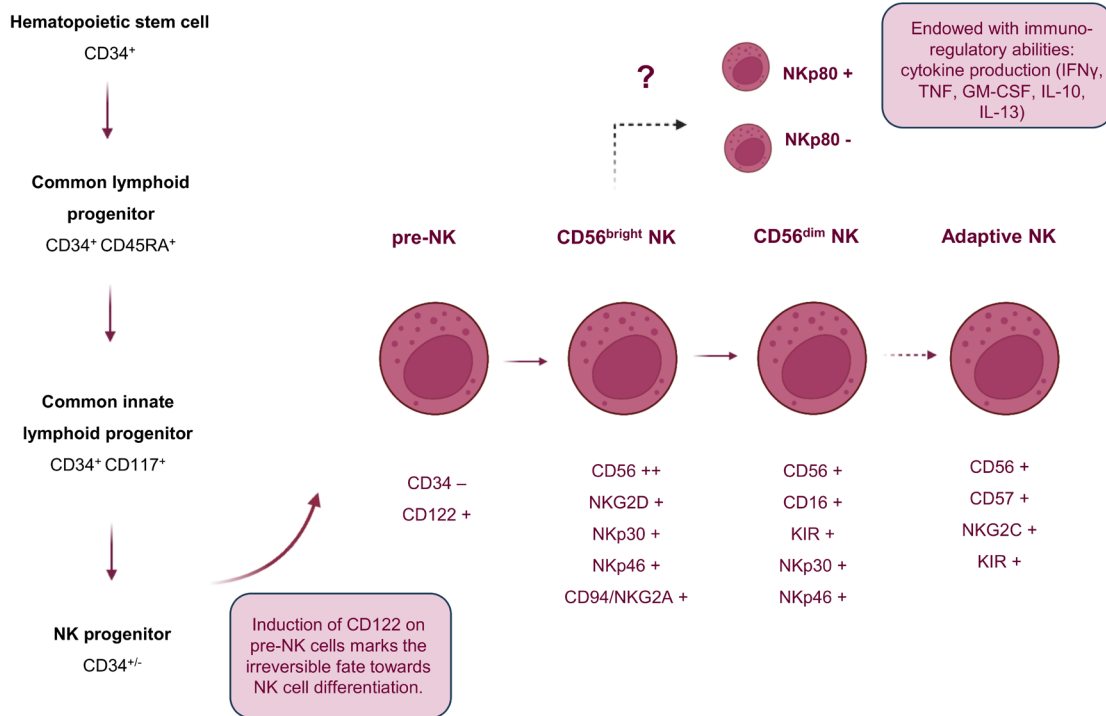


Figure 2: The commitment towards the NK specific lineage is characterized by the downregulation of CD34 and the acquisition of CD122, a common IL-2 receptor subunit β shared by IL-2 and IL-15 signaling pathways. Indeed, immature NK cells marked by this CD34 downregulation have lost their plasticity potential, as *in vitro* studies have shown their inability to differentiate into T cells or DCs in response to IL-15 [13, 20, 23, 24].

1.2 KIR immunogenetic system

The KIR gene complex – encoding a total of 15 receptors – is located in the human chromosomal region 19q13.42 at the level of the leucocyte receptor complex region. On a genetic level, these receptors are encoded by 13 different gene loci (KIR2DL1, KIR2DL2/L3, KIR2DL4, KIR2DL5A, KIR2DL5B, KIR2DS1, KIR2DS2, KIR2DS3, KIR2DS4, KIR2DS5, KIR3DL1/S1, KIR3DL2, KIR3DL3) and two pseudogenes (KIR2DP1, KIR3DP1), with several KIRs segregating as alleles on a single locus, such as KIR3DL1/KIR3DS1 and KIR2DL2/KIR2DL3. KIR gene length usually ranges between 4 kb (KIR3DP1) and 16 kb (KIR3DL2) with a large set of integrated introns separated by exons responsible to encode the cytoplasmic tail (exon 7-9), the stem region (exon 6), the extracellular immunoglobulin (Ig)-domains (exon 3-5) and the leader peptide (exon 1-2). Inherent to these receptors are their variable gene content, copy number variation (CNV), extensive allelic polymorphism and high degree of sequence homology (Fig.3). Each individual inherits KIRs in haplotypes, differing in their number of genes and gene content as a consequence of structural genetic recombination, gene duplication, deletion and fusion events. KIR haplotypes can be broadly classified into A and B and separated in a centromeric (cen) and telomeric (tel) part, each flanked by framework KIR genes: KIR3DL3, KIR3DP1, KIR3DL2 and KIR2DL4. While haplotype A has a consistent pool of seven – mostly inhibitory – KIR genes, haplotype B varies in terms

of its KIR gene content, balancing activating and inhibitory KIR genes. While a strong linkage disequilibrium persists between KIR genes such as KIR2DP1 and KIR2DL1, a recombination hotspot between the cen and tel parts further diversifies the haplotype variability. Additionally, they are subject to extensive single nucleotide variants – similar to human leucocyte antigen (HLA) genes – with at present (06.2024) 2238 different KIR alleles registered in the Immuno Polymorphism Database-KIR (IPD-KIR) database (<https://www.ebi.ac.uk/ipd/>). This database is constantly updated given the extensive application of high-throughput, high-resolution KIR genotyping in population genetics [25-28]. Structurally, KIRs can be divided into three distinct classes based on their number of extracellular Ig domains: (i) Type I KIR2D with two extracellular domains D1 and D2 (KIR2DP1, KIR2DL1-3, KIR2DS1-5); (ii) D0 and D2 extracellular bearing receptors represent type II KIR2D (KIR2DL4, KIR2DL5); (ii) KIR3D encoding three extracellular domains D0, D1 and D2 (KIR3DL1-3, KIR3DS1). Functionally, we further distinguish between inhibitory and activating KIRs based on their long (KIR2/3DL) or short (KIR2/3DS) cytoplasmic tail length containing either immunoreceptor tyrosine-based inhibitory motifs (ITIM) or immunoreceptor tyrosine-based activation motifs (ITAM), respectively [29].

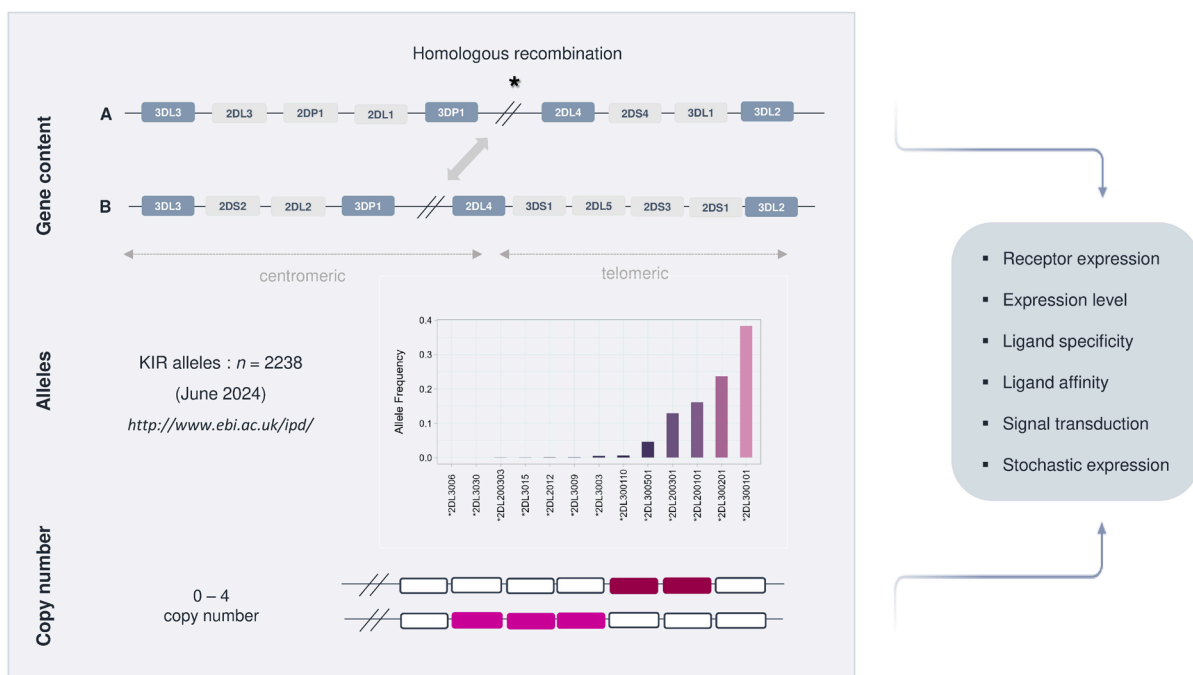


Figure 3: The KIR immunogenetic system displays high variation in terms of its gene content, copy number and alleles, affecting several functional aspects.

Table 1. KIR receptors and their HLA and non-HLA interacting ligands.

Receptor	Ligand	Status	Ref	Current implication in transplant and cancer immunology
KIR family				
KIR2DS1	Group 2 HLA-C	A	[30]	Protective effect of donor KIR2DS1 in an HLA-C1+ environment on relapse incidence in leukemia [31].
KIR2DS2	Group 2 HLA-C	A	[30]	
KIR2DS3	Unknown	A		
KIR2DS4	01:02, 02:02, 04:01, 05:01,14:02,16:01 HLA-A*11	A	[32]	
KIR2DS5 (*003, *004, *005, *006, *007, *008)	Group 2 HLA-C	A	[33]	
KIR3DS1	HLA-Bw4, HLA-F	A		
KIR2DL1	Group 2 HLA-C	I	[34]	
KIR2DL2/L3	Group 1 HLA-C	I	[34]	
KIR2DL4	HLA-G	I	[35]	
KIR2DL5	Unknown	I		
KIR3DL1	HLA-Bw4 (HLA-B, HLA-A loci)	I	[36, 37]	Protective effect of none inhibiting KIR3DL1 – HLA-Bw4 interactions in neuroblastoma and AML transplanted recipients [38].
KIR3DL2	HLA-A*11, *03	I	[39, 40]	
KIR3DL3	HHLA2	I	[41]	This ligand belongs to the B7 family and is expressed in various solid organ and hematological cancerous settings and is mainly expressed on CD8+ T cells [42, 43].

LIR family 1 and 2

LILRB1	HLA-A, -B, -C, -G, -F S100A9 UL18	I	[44]	A recent population genetic study unveiled an extensive haplotypic and allelic diversity of LILRB1 and 2 regions in a Brazilian population [44]. Current perspectives on its implication as an immune checkpoint target are reviewed here [45].
--------	--	---	------	--

CD159 (NKG2) family

* Dimerization of CD94 leads to the creation of inhibitory receptors that transmit inhibitory signals via two immunoreceptor tyrosine-based inhibition motifs (ITIM).

** via charged lysine residues in the transmembrane segment that allow them to interact with signal-transducing molecules such as DAP12 containing ITAM motifs.

*** Ligand-receptor interaction leads to the association with the adaptor molecule DAP10, followed by phosphorylation of tyrosine residues in the YINM domain of DAP10 enabling ligation to PI3K to Grb2-Vav1 for signal transduction.

NKG2A/CD94 *	HLA-E	I	[46]	Marker for NK cell immaturity and inversely correlated with KIRs acquisition. Educational pathway in -21 M/M genotype individuals [47]. Potentially promising target in immunotherapy and is greatly reviewed in the following publication [48].
NKG2C/CD94 **	HLA-E peptide-dependency (UL40, pp65)	A	[49]	Majorly present on adaptive “memory” NK cells that expands upon cytomegalovirus (CMV) infection. Major field of interest in disease outcomes, greatly reviewed here [50, 51].
NKG2D ***	MICA, MICB, ULBPs 1-6	I	[52]	Expressed on NK and CD8+ T cells. NKG2D plays a major role in the anti-tumor immunity as NKG2D ligand shedding is an important cancer immune escape mechanism [53].

NCR family

*mediated by adaptor molecules, TCR ζ and/or Fc ϵ RI- γ (for NKp30 and NKp46) or KARAP/DAP1266 (for NKp44), containing cytoplasmic ITAM motifs

NKp30 *	B7-H6, BAT3	A	[54]	Expression on resting mature NK cells.
---------	-------------	---	------	--

NKp46 *	Hemagglutinins, heparan sulfate proteoglycans	A	[54]	Expression on resting mature NK cells. Further details on the implication of NCRs in immunopathological settings are outlined in the following review [54].
NKp44 *	Hemagglutinins, heparan sulfate proteoglycans	A	[54]	Essentially upregulated upon cytokine priming by IL-2, IL-15, IL-1B on CD56 ^{bright} NK cells.
Inhibitory immune checkpoint inhibitors				
PD-1 (CD279)	PD-L1, PD-L2	I	[55]	Upregulation of PD-1 has been specifically reported by CMV on mature CD56 ^{dim} , CD57 ⁺ NK cells. PD-1 ⁺ NK cells have been commonly associated with an exhausted phenotype defined by a reduced functionality and cytotoxicity, enhanced apoptosis sensitivity and reduced proliferative capacities in PDL-1 ⁺ tumor microenvironments. The potential of PD1 ⁺ NK cells for targeted immunotherapy is reviewed here [56].
TIGIT	Nectin 2, PVR	I	[57]	
LAG-3	HLA-II, galectin 3	I	[58]	
CD96 (Tactile)	Nectin 1, PVR	I	[58]	
KLRG1	Cadherine (E, N,R)	I	[59]	

A; activating, AML; acute myeloid leukemia, CD159; C-type lectin superfamily, I; inhibitory, KIR; killer-cell immunoglobulin-like receptor, KLRG1; killer cell lectin-like receptor G1, LAG-3; lymphocyte activation gene 3, MICA; Major histocompatibility complex class I chain-related protein A, MICB; Major histocompatibility complex class I chain-related protein B, NCRs; natural cytotoxicity receptors, PD-1; programmed Death-1, PVR; polio virus receptor, TIGIT; T cell immunoreceptor with Ig and ITIM domains, ULBP; UL16-binding proteins

1.3 NK cell repertoire diversification

The combined application of high-resolution KIR genotyping and functional *in vitro* experiments have enhanced our comprehension of the biological translation of KIR allelic polymorphisms – of which the most important ones are schemed out in figure 4 [60-71]. Understanding the structural/functional behavior of KIR alleles is of great importance to emit accurate hypotheses in pathological settings.

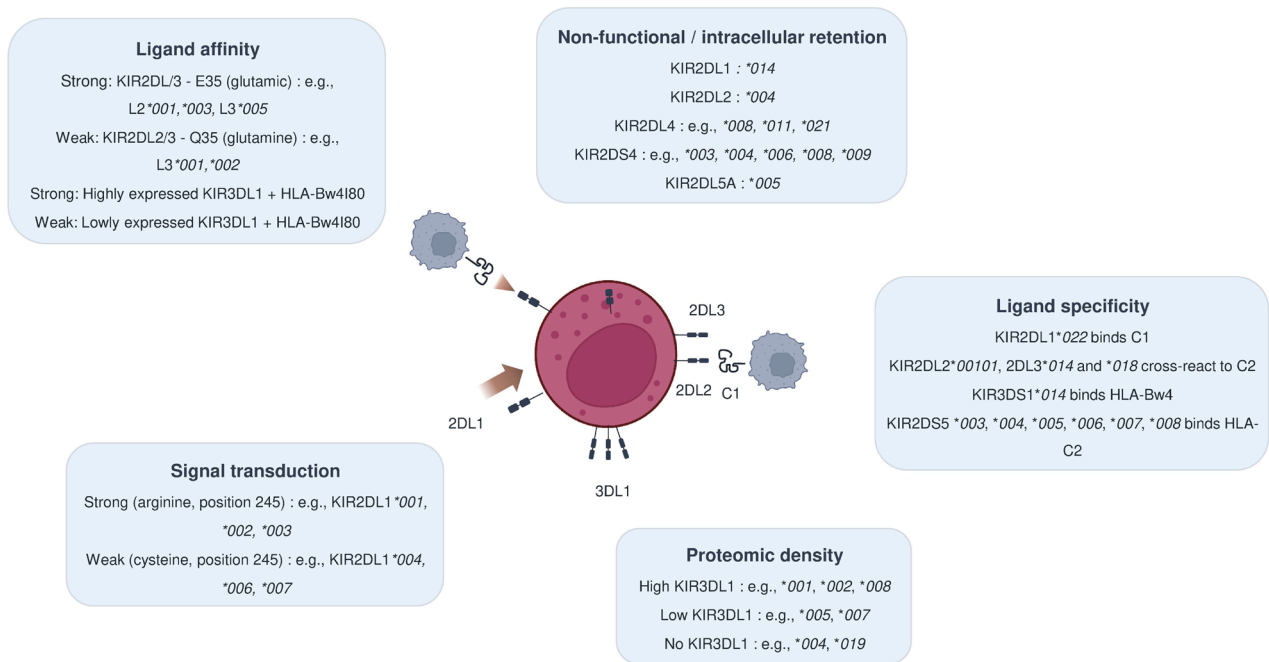


Figure 4: Allelic variation translates into different structural and functional KIR characteristics encompassing receptor expression, expression level, ligand specificity, signal transduction and ligand affinity.

Besides displaying extensive genetic diversity, KIRs are stochastically expressed on the NK cell surface, with the ability to express between zero and three KIRs, except for KIR2DL4 and KIR3DL3. [72]. Accounting for the KIRs stochastic expression and other NK-cell specific receptors, Horowitz *et al.* took advantage of high-resolution immunophenotyping and estimated that the peripheral repertoire of healthy individuals encompasses from 6'000 to 30'000 different NK cell clones [73].

It is thought that a bidirectional promoter termed “proximal switch” is the primary determinant for the stochastic expression, as antisense transcripts have been described in NK cells lacking the respective KIR [74-76]. Differential expression levels of KIRs, such as described for KIR3DL1 alleles, are thought to result from the strength between sense and antisense transcripts. Besides this proximal switch, the intermediate promoter is majorly responsible for regulating KIR expression level, as demonstrated for

KIR2DL1 alleles [77]. Additionally, epigenetic regulation such as DNA methylation (CpG) at the KIR proximal promoter correlates with KIR transcription, as enhanced DNA methylation or excessive demethylation have been demonstrated to silence KIR expression and repress NK cell cytotoxic abilities, respectively. Finally, various microRNAs targeting mRNA and inducing its degradation have been suggested to suppress KIRs expression, such as MiR-146a-5p downregulating KIR2DL1 and KIR2DL2 [78].

The probabilistic expression of KIRs surpasses the stochastic product rule, suggesting additional regulation by non-random events [79]. The HLA environment predicts peripheral non-self and self KIR repertoire, although results are controversial as to this regard [79, 80]. In addition, selective events such as infections might additionally skew the KIR repertoire. This repertoire diversification strategy enables to maximize the ability of sensing pathological settings.

1.4 KIR sequencing platform

The high degree of structural diversity, single nucleotide variation and sequence homology have greatly challenged the research community to genetically map the KIR gene complex at a high-resolution. KIR genotyping with the assessment of gene presence and absence performed by low-resolution techniques such as polymerase chain reaction using sequence specific primers (PCR-SSP) and PCR using sequence-specific oligonucleotide probes (SSOP) has for a long time been the gold standard [81]. Although informative, this reductionist view on the KIR immunogenetic system bears a high potential for biological misinterpretations in disease association studies. Additionally, missing scalability and throughput have thwarted further large-scale interrogations.

In recent years, there has been propelling advancements in sequencing technologies through next-generation sequencing (NGS), enabling a massive increase in the throughput and gene variation identification. Alongside these breakthrough technical developments, KIR system genotyping solutions providing allele level resolution have emerged, paving the way for an increased granularity in research studies. There are currently two different enrichment strategies, PCR amplification and hybridization capture. While the first approach uses primers for target amplification, the hybridization workflow makes use of oligonucleotide probes to pull down target sequences [82, 83]. Alignment and interpretation of NGS data is equally challenging especially in a highly homologous region susceptible to read misalignments.

A recent state-of-the-art sequencing workflow and bioinformatic pipeline called Pushing immunogenetics to the next generation (PING) have been developed for the high-resolution KIR genotyping: the wet-lab sequencing protocol is based on a capture hybridization methodology using more than 180'000 KIR-specific probes. For the subsequent genotype, CNV and allele calling, raw fastq files are passed through a first module to extract all KIR and HLA paired-end sequencing reads. Copy number determination is then executed by calculating a reads ratio between each KIR gene and KIR3DL3 – reference KIR present at two copies in each individual. Finally, allele calling is performed by multiple rounds of filtering and alignments, where all KIR reads determined by copy number variation

analysis are mapped against a set of KIR reference alleles. The final PING pipeline output includes accurate and high-fidelity (i) copy number determination (ii) KIR alleles at a 5-digit resolution for all KIR gene loci (iii) genotype determination [82].

1.5 NK cell education: modulation by genetic and non-genetic factors

Comparable to T lymphocytes education, NK cells undergo a process of education – based on ligand-receptor interactions – for the acquisition of functional competence and self-tolerance. The exact mechanism by which NK cells are educated is currently a major field of debate, with four hypothetical mechanisms being brought forward in recent years:

- (i) Licensing model: Interaction between self-MHC-I ligands and inhibitory receptors on the NK cell surface will induce education [84-86].
- (ii) Disarming model: This model is based on the assumption that NK cells are educated by nature and become anergic in the case of the absence of inhibitory receptor engaging a self-MHC-I ligand [87, 88].
- (iii) Rheostat model: NK cell education is seen here as a quantitative rather than a dichotomous trait. The interaction strength between an inhibitory receptor and its ligand will dictate the educational threshold [89, 90].
- (iv) Tuning model: The most recent model can be seen as a refinement of the previous model and states that the net input of inhibitory receptor-ligand interaction will render the education of NK cells [91-93].

Educated NK cells are thus seen as reactive cells, sensitive to changes in the MHC-I expression level on target cells. In contrast, uneducated NK cells who failed to reach the educational threshold will remain hyporesponsive and require enhanced activating signals to reach functional competence. Although the KIR/HLA axis has been proposed as the central regulator in the predictive mechanism of NK cell licensing, there is mounting evidence that other activating and inhibitory receptors are able to regulate the NK cell educational process:

- (i) KIR2DS1 – HLA-C2: KIR2DS1⁺ NK cells have been shown to be rendered hyporesponsive in individuals being homozygous for HLA-C2, assumed to be caused by a time- and expression-induced anergy [94].
- (ii) Natural killer group 2 (NKG2) type A – HLA-E: Engagement of NKG2A/HLA-E has been proposed to be the second most important receptor engagement with genotype specification predicting the educational involvement. Indeed, individuals bearing an HLA-B (-21 methionine (M)) genotype are associated with a higher stability of HLA-E molecules on the cell surface due to the encoded peptide in comparison to HLA-B (-21 threonine (T)) alleles. Consequently, carrying -21 M alleles will favor NK cell education towards CD94/NKG2A-HLA-E interaction, whereas individuals with -21T alleles are thought to have a skewed education towards KIR/HLA interactions [47]. This educational division based on the immunogenetic framework of each individual has been shown to influence disease and

therapy outcomes. As a way of example, patients with -21 M alleles were less likely to develop severe respiratory failure after severe acute respiratory syndrome coronavirus 2 (Sars-Cov-2) infection and showed improved outcomes after NK cell immunotherapy [95, 96].

- (iii) LILRB1 – HLA-G: Engagement of the LIR-1 receptor with HLA-G has been reported as an important mediator of a co-educational mechanism enhancing NK cell functional competence *ex vivo*, characterized by increased granzyme B (GrzmB) and DNAM-1 expression and ADCC mediated cytotoxicity [97].
- (iv) TIGIT – CD155: Another example of co-education is the non-classical TIGIT-CD155 interaction. CD155-deficient mice were subject to a functional impairment with reduced CD107 α degranulation and IFN γ production [98].

These aforementioned licensing mechanisms beg the question why a considerable proportion of NK cells in the peripheral blood are “left” uneducated – in contrast to T cell education, in which failing in recognizing self-MHC I/II complexes induces cell apoptosis? Are these backup cells that are able to undergo a second round of education depending on environmental signals? That unlicensed NK cells are able to achieve potent effector functions is exemplified in the case study of cytomegalovirus (CMV), where hyporesponsiveness is broken under the influence of pro-inflammatory cytokines [99, 100]. Further, uneducated Ly49⁺ NK cells were subject to a higher proliferation rate than educated NK cells in murine models of influenza infection concomitant with the upregulation of HLA-ligands on lung epithelial cells [101].

Additional genetic and non-genetic factors have been demonstrated or bear the potential to fine tune this process, out of which some are listed below:

- (i) Recent data suggests a KIR allele dependent hierarchy of educated NK cells: subsets displaying highly expressed KIR3DL1 receptors (e.g., KIR3DL1*008) were subject to a higher education with different types of HLA-Bw4 ligands than low expressed KIR3DL1⁺ NK cells (e.g., KIR3DL1*005), defined by a higher cytotoxicity against autologous human immunodeficiency virus (HIV)- infected CD4⁺ T cells [102]. Similarly, KIR2DL2 and KIR2DL3 alleles bearing a glutamic acid at position 35 instead of a glutamine displayed higher functional responses against HLA-C1 deficient target cells [68]. Deciphering the genetic polymorphisms would greatly inform and fine tune insights into current licensing prediction models (i.e., rheostat model).
- (ii) There is mounting evidence that NK cells are sensitive to changes in the immunopeptidome. Peptide sequence variation bears the potential to modulate the recognition and binding mode of KIRs, thus affecting the level of NK cell activation. As a way of example: HLA-Cw*01:02 (C1-allotype) binding peptide VAPWNSFAL induces a strong KIR2DL2/3 mediated inhibition in contrast to HLA-Cw*01:02 binding peptide VAPWNSDAL or IVDKSGRTL. AGDDAPRAV promotes binding of KIR2DL1 to HLA-C*08:02, a C1 allotype – KIR2DL1 being otherwise highly specific for HLA-C2 [103, 104]. This KIR mediated

peptide permissiveness represents an important immune escape strategy or, inversely, a mechanism of NK cells to target viral pathogens.

- (iii) Non-classical HLA-E allele polymorphisms are associated with differences in their expression level – HLA-E^G (01:01) being highly expressed and HLA-E^R (01:03) being lowly expressed on the cell surface [105]. There is a plethora of studies in cancer immunology that have linked the presence of high HLA-E expression levels with detrimental outcomes, likely due to an increased NK cell inhibition threshold [106]. Additionally, a recent study elegantly provided insights into the impact of signal leader peptides and their sequence polymorphism on the HLA-E expression and binding affinity, calibrating the HLA-E/NKG2A receptor-ligand engagement with a potential impact on the NK cell responsiveness [107].

Comprehension of the NK cell implication in physiological and disease settings is highly dependent on the way we consider NK cells being educated. Consequently, an incomplete capture of the molecular basis of this process may skew biological interpretations. Out of the aforementioned studies, we can however cautiously perceive NK cell education as a reversible and adaptable process, governed by genetic and non-genetic factors.

2. T cell biology

To meet the challenge of a universal immune cover, T cell immunity is equipped with a universal system that satisfies the following requirements: a precise and fast recognition system for rare, infinitely diverse, and evolutionary constantly mutating antigens. The antigen recognition system of T lymphocytes is mediated by the interaction between a T cell receptor (TCR) – that confers the T cell its specificity – and a peptide loaded onto a major histocompatibility complex (pMHC) class I or II on antigen presenting cells (Fig.5) [108-110].

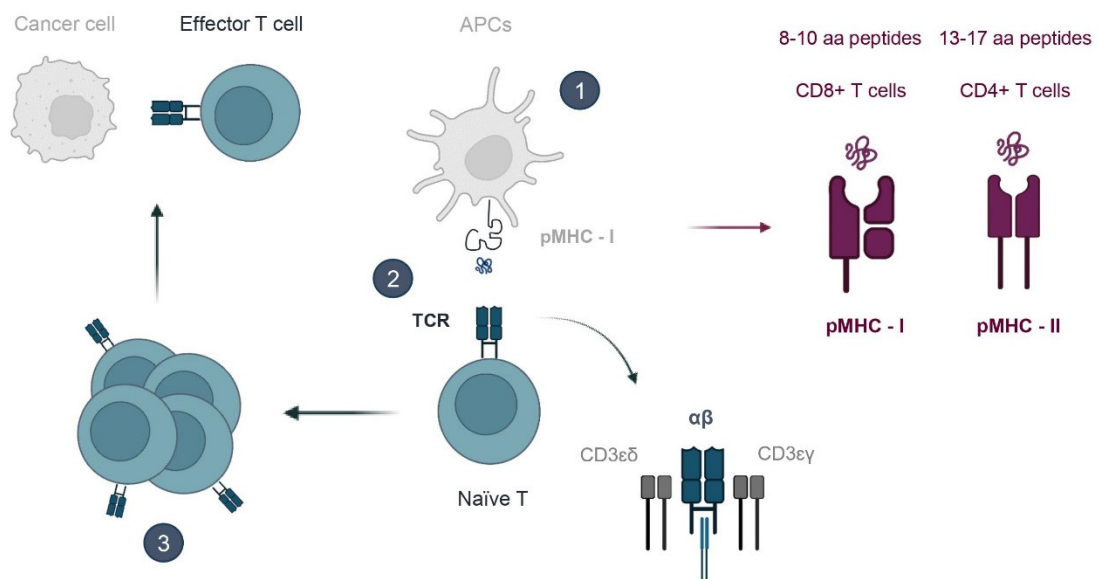


Figure 5: (1) T lymphocytes are constantly sensing the environment for antigens while circulating between SLO and the peripheral blood system. At the site of entry, antigen-presenting cells (APC) capture and intracellularly process them as follows: antigen breakage into smaller peptides, assemblage of MHC I or II molecules with their appropriate peptide within the lysosome and trafficking of the pMHC complex towards the cell surface to be recognized by surveying specific T lymphocytes in the SLO. (2) Activation of T lymphocytes results from two major receptor-ligand engagements: recognition by T lymphocytes of pMHC molecules through their $\alpha\beta$ TCR, a recombined receptor composed of two polypeptide chains that forms a multiprotein complex with the cluster of differentiation (CD) 3 chains and the CD4 /CD8 coreceptor [111-115]. It is thought that pMHC molecules can serially engage several TCRs, therefore progressively amplifying the magnitude of intracellular signals to reach a certain threshold for T cell activation (theory of engagement-dissociation-engagement). The second engagement involves co-stimulatory molecules that rely on the 44-kDa glycoprotein CD28 binding to B7-1 (CD80) and B7-2 (CD86) on APCs. These costimulatory signals are crucial for naïve T lymphocytes to achieve activation, while effector or memory T cells are mostly independent of such signals. (3) Successful activation of T lymphocytes initiates signaling cascades leading to cell differentiation into two phenotypic distinct cells: the majority are geared towards differentiation into effector T cell subsets enduring a 50'000-fold large expansion. This feature is crucial to mount an immediate, synergic and targeted immune response against the specific pathogen, as all progeny of this given T cell will carry the identical TCR. Resolution of the episode leaves peptide specific CD8+ T cells displaying an effector memory phenotype, allowing for a more rapid immune response in cases of recurrence. It is a matter of debate whether naïve T cells immediately differentiate into memory T cells or whether the decision of the fate is set at the activation stage [116].

2.1 TCR diversity generation

- (i) The main source by which T lymphocytes acquire specific TCRs to accommodate diverse immunogenic epitopes is through a process called somatic recombination. The molecular structure originates in the complementary determining region 3 (CDR3) region, composed of variable (V), diversity (D) and joining (J) genes. These genes – responsible for creating a three-dimensional pocket to recognize the antigen – are non-functional in their initial germline state and encompass a total of 76 TRBV, 2 TRBD and 14 TRBJ different segments at the TCR β locus and 54 TRAV, 61 TRAJ different segments at the TCR α locus. The first diversification event occurs through somatic random recombination of these V(D)J genes by recombinase activating enzymes (RAG) 1 and 2. This combinatorial diversity by simple multiplication can lead to up to 3.8×10^4 unique combinations. The largest contribution is made by junctional diversity, which relies on the stochastic insertions of N-nucleotides at the V-D or D-J junctional sites by the terminal deoxynucleotidyl transferase (TdT) enzyme. The product of these two sources of diversity confers the immune system its enormous pathogen recognition potential [117-119].
- (ii) The pMHC-TCR three-dimensional conformational interaction follows fixed “rules of engagement” based on a restriction triad claiming that conserved invariant structures and the highly variable domain of TCR complexes have a propensity to bind to their respective conserved and variable counterparts on the pMHC complex [120]. Several studies have to some extent discredited these rigid docking paradigms and have instead put forward the assumption of a binding promiscuity combined with flexible pMHC binding motifs [121].
- (iii) There are rough estimates that it is possible to generate a pool of approximately 20^9 peptides derived from the 20 proteinogenic amino acids available. Scanning up to 20^9 distinct peptides on an average of 1 recognition event per second would importantly delay the induction of an immune response. Further, it is thought that TCR genetic diversification events result in a theoretical repertoire of 10^{39} unique TCR β nucleotide sequences prior to thymic selection and 10^{12} distinct pMHC interactions. Given that the number of generated peptides largely overweights the number of distinct responsive TCRs, the concept of “crossreactivity” has been put forward as the ability of one TCR to recognize multiple pMHC complexes at once [120, 122]. Compelling arguments have indeed made this hypothesis a plausible explanation for the observation that a multitude of distinct peptides were able to activate T cells despite having undergone thymus-selection on different peptides [123].

Several studies have attempted to compute estimates about the sheer magnitude of the TCR repertoire since the advent of high-throughput sequencing technologies. This task still poses considerable challenges: (i) Sampling issues massively hamper efforts to accurately estimate the diversity of the TCR repertoire given the presence of billions of T lymphocytes circulating in the blood stream and throughout

tissue sites. (ii) Bulk TCR sequencing hinders the analysis of direct paired TCR $\alpha\beta$ chain analysis. (iii) While single-cell sequencing approaches allow for greater depth, their low sequencing throughput is several orders lower than bulk TCR sequencing impeding accurate estimates [124-129].

2.2 Central regulation of T cells

The diverse set of TCRs mounted on T cells are subject to core regulatory processes to ensure functionality and self-tolerance while avoiding autoreactivity to self. This cellular selection can be distinguished in positive and negative selection and occurs via several cellular players in the thymus. The early steps of T cell development are confined to the BM, from which T cell precursors migrate towards the thymus to complete their maturation and fate determination. There, T cells undergo selection based on the affinity model informed by the interaction strength between TCRs and self-pMHC complexes. To ensure tolerance towards a large spectrum of self-peptides, the AIRE transcription factor induces the expression of a range of tissue-specific proteins within thymic epithelial cells. TCRs that do not strongly engage pMHC complexes (missing-self tolerance) or with a too high affinity (autoreactivity) will be negatively selected and further die by apoptosis. Double positive thymocytes receiving sufficient $\alpha\beta$ TCR signals in response to self-pMHC complexes commit then to either the CD4 or CD8 lineage, thus becoming single positive cells. This selective process imparts T cells with a long-lasting tolerance towards self-peptides [130, 131].

2.3 Convergent recombination as a result of evolution

Out of these properties, we might expect that the likelihood to identify shared TCRs between individuals almost tends towards zero. However, thanks to the large-scale application of high-throughput TCR sequencing, sharedness in the T cell response with the high proportion of identical shared TCR sequences – termed as public – among individuals has become evident [132]. The emergence of public clones has sustained many questions related to their identity and role in adaptive immunity: Do they exhibit any evolutionary selective advantage or specific molecular features that confer them growth advantage? Intrinsic factors – mutually likely not exclusive – have been acknowledged to shape the formation of public repertoires:

- (i) The first hypothesis is based on the concept that public TCR owe specific molecular features, so called recombination biases, favoring the repeated occurrence of a particular recombination event. Some studies have highlighted that public TCRs resemble the near germline recombination with no to minimal random addition of nucleotides. The ease in producing less complicated TCR sequences would lead to a bias towards producing them at a higher frequency. Following recombination biases have been additionally proposed: preferential V(D)J usage, addition of specific random nucleotides such as single guanines and strings of guanines and extent of nucleotide deletion in germline sequences [133-135].

- (ii) The cross-comparison of individual immune repertoires has revealed the existence of differential nucleotide TCR sequences that merge towards identical amino acid sequences, leading to the next more plausible concept of convergent recombination. The current theory incorporates previous explanations but rationalizes TCR amino acid sequence generation as a result of three consecutive events: (a) Many recombination events converge towards the same nucleotide sequence (b) Multiple nucleotide sequences encode the same amino acid sequence (c) The production frequency of nucleotide sequences depends on the number of random nucleotide additions and different recombination events [134].
- (iii) Finally, the process of negative and positive selection within the thymus might confer a selective advantage to public TCRs by presenting favorable antigens or by persistent antigenic stimulation in the periphery [136].

2.4 T cell replenishment in the periphery

Compelling evidence points towards a TCR repertoire that is qualitatively and quantitatively controlled in a relative constant manner throughout life. Three core mechanisms seem to determine the T cell replenishment:

- (i) Thymic activity is the main contributor of new T cell clonotype generation during the time spanning the neonatal to early infant. *In utero* TCR formation and V(D)J diversification begin at the end of the first trimester with the migration of T cells to the thymus and the establishment of a first layer of zero insertion clones as a consequence of the absence of Tdt enzyme activity. The thymus reaches its peak activity during the time spanning one and eight years as a coping mechanism of the body against exposure to novel antigens and subsequently undergoes a progressive involvement [137, 138].
- (ii) To compensate for the lack of thymic activity, it has been postulated and further demonstrated that T cells undergo a self-regulatory process for maintaining the quantitative stability of their biological system through homeostatic proliferation. Homeostatic proliferation relies on the recruitment of T cells to SLO, where they encounter a network of fibroblastic reticulate cells, which are substantial for the production of IL-7 and antigen presentation – whereas the nature of the antigens involved in this mechanism remains to be elucidated. Studies revealed the tremendous importance of IL-7 as (a) transfer of T cells in lymphocyte-repleted hosts remains undivided and (b) exogenous supply in IL-7 is sufficient to induce homeostatic proliferation. This observation led to the establishment of a competition model between T cells for available resources, e.g., the amount of available IL-7 produced by stromal and fibroblastic cells [139].
- (iii) Spontaneous proliferation has been proposed as the other mechanism to recalibrate the repertoire equilibrium after perturbation and is triggered by the quality of the repertoire, i.e., presence of memory T cells rather than the overall TCR size. The TCR clonality is one of the leading drivers of spontaneous proliferation. Indeed, studies have shown that the transfer of TCR transgenic T cells between hosts with a sufficiently different clonal reservoir

induces spontaneous proliferation, while the transfer between similar clonal repertoires inhibits this response [140]. The second crucial factor controlling this proliferation is the T cell memory complexity. TCR transgenic T cell transfer experiments showed that infusing naïve T cells into mice with a preexisting low complexity memory T cell repertoire exhibits higher proliferation than into a high complexity environment [141, 142].

We can postulate a convergent replenishment system in which naïve and memory compartments are differentially regulated to ensure the global stability of the system and where so called “holes” are rapidly filled to prevent immune lacking.

2.5 Technical development to probe the TCR repertoire

Early days sequencing methods to probe the TCR repertoire landscape mainly consisted of flow cytometry- based techniques using anti-TRBV monoclonal antibodies and spectratyping generating data about the differential CDR3 lengths using V β primers [143-145]. At the gene level, TCR sequence identity was first exhibited by Sanger sequencing, with a limited output value given its restricted library size [146, 147]. These labor-intensive and time-consuming techniques have largely warranted further in-depth investigation due to their lack of accuracy, depth and throughput.

The emergence of high-throughput NGS was a groundbreaking achievement in the field of T-cell immunity, addressing the unmet need of a high-resolution technique while simultaneously capturing the broad range of a large set of different genetic TCR clonotypes and surveying for the presence of low frequent clonotypes. Additionally, it enables to target many samples at the same time for immune monitoring and cross comparison purposes [148-150]. The current gold standard enrichment strategies for TCR sequencing include (i) multiplex PCR and (ii) amplification of complementary DNA ends (5' RACE). The commonly adopted multiplex PCR library preparation involves forward primers targeting all V genes and a second set of reverse primers binding J or C regions. The DNA sequencing approach that has undoubtedly dominated in this field has been the Illumina platform, a short-read NGS method based on sequencing by synthesis [151].

2.6 Mathematical modeling in immunology

High-throughput TCR sequencing has led to a great enhancement in the depth and wealth of information retrievable from samples. This shift towards biological big data has required the adoption of more sophisticated analysis tools for data interpretation and has simultaneously opened new venues for a more mathematical approach to a system-level analysis.

One of the biggest challenges in the field of T-cell immunity relies in the identification of so-called orphan TCRs, without an identified cognate antigen. Classical approaches are based on antigen-pMHC multimer assays that have led to the emergence of online available databases assembling and referencing such experimentally verified known TCRs epitopes. These databases are continuously

updated, providing researchers with a great tool to infer such biological information on their acquired datasets [152, 153]. However, the lack of scalability majorly hampers its application.

In recent years, incremental effort has been put into developing bioinformatic tools to empower biological understanding of the TCR specificity. The first set of computational tools assumes that an antigenic specificity can bind the same pMHC complex when either sharing identical/ similar nucleotide/amino acid (aa) sequences or other physiochemical characteristics. These clustering strategies abandon the raw sequence information and rather focus on inferring specificity based on macroscopic features with the attempt to link a cluster enrichment to a specific setting. Known and unknown TCR-peptide interactions are able to co-cluster thus magnifying the clonal specificity structure of the dataset. Currently, many of these unsupervised analytical tools are out on the market and are based on simple CDR3 β alignment-based methods, CDR3 β similarity-weighted distances and *k*-mer based analysis, to name a few (Fig.6). Each of these computational methods shares similarities and differences in their analytical workflow and computational burden [154-156].

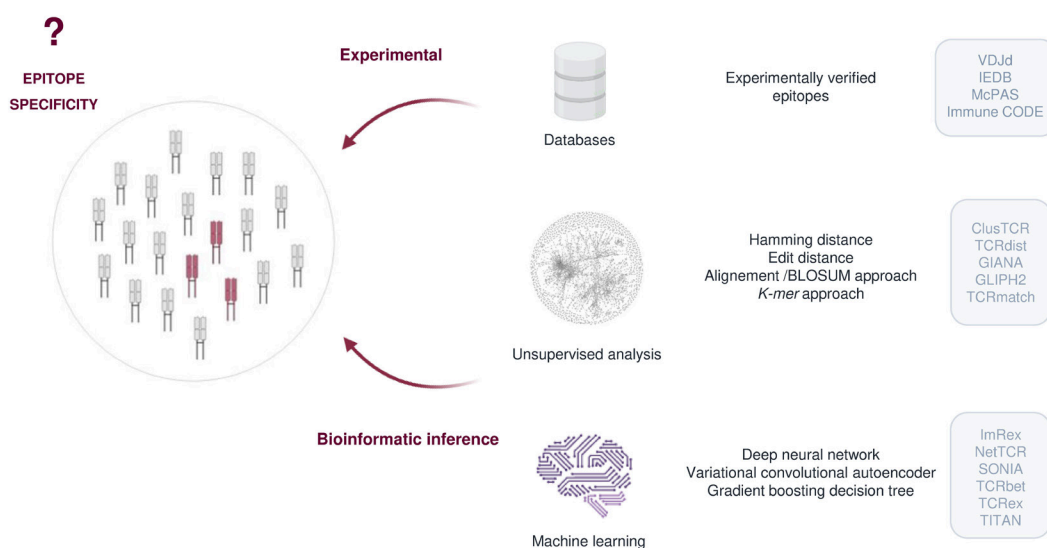


Figure 6: Different experimental and bioinformatic approaches for TCR epitope specificity identification.

The second set of computational tools utilizes artificial intelligence to train models on a set of TCR sequences with known specificities. Among the recently developed algorithms are DeepTCR (variational convolutional autoencoder) [157], SETE (gradient boosting decision tree) [158] and TcellMatch (multiple deep learning architectures) [159]. As the input data of these classification training models are dependent on the availability of validated labeled TCRs sequences, the discriminative accuracy is largely biased by overrepresented known TCR-peptide interactions and the lack of true negatives.

There are still many hurdles to overcome before gaining a comprehensive and more complete mapping of TCR-peptide interaction specificity. Further detailed information on the mathematical and bioinformatic background, current perspectives and challenges of computational biology in the TCR immunogenetic field is provided in the following publication [160].

3. Allogeneic stem cell transplantation as a therapeutic option for immune and hematological related disorders

Allogeneic hematopoietic stem cell transplantation (HSCT) is a well-established cellular therapy for patients with hematological malignancies, immune disorders and other hematological related diseases [161]. The utilization of HSCT commonly comes as a rescue treatment when previous less invasive and complicated options failed to reverse the disease. The potential curative effect of this procedure relies in the recreation of a new healthy hematopoietic system derived from donor-infused stem and immune cells combined with the eradication of residual tumoral cells through the exploitation of the so-called immunological graft-versus-tumor (GvT) effect [162]. To this end, the recipient's immune system needs to be depleted for the creation of an immunological space to accommodate the incoming donor transplant. Such a niche space is crucial to (i) hamper T cell mediated rejection and immuno-allogeneic events such as graft-versus-host disease (GVHD) (ii) mediate engraftment of the stem cell graft (iii) disease control by eliminating potential residual tumor cells. To this end, various conditioning regimens consisting of cytotoxic chemotherapy, irradiation and anti-T cell antibodies are currently applied depending on the recipient's age, fitness and disease state [163, 164].

The recreation of the immune system state thereafter is thought to be the contribution of a first wave of mature donor-derived immune cells followed by a second wave of *de novo* stem cell derived cells. Broadly, this *de novo* immune reconstitution follows three consecutive stages that differ in the immune cell composition, type and severity of adverse events [165-167].

- (i) The aplastic phase of a duration of one to three weeks post-transplant is defined by the profound deficiency of all immune cell lineages. This extreme immunosuppressive state greatly increases the patient's vulnerability to infectious agents.
- (ii) The following three months are characterized by the reconstitution of the innate immune system: the monocyte cell compartment is the first to recover, followed by the resurgence of granulocytes, predominantly neutrophils. Endowed with inherent abilities to identify and eliminate transformed cells, NK cells are the first lymphoid cells to reconstitute post-HSCT [168, 169] with a peak in immature NK cells during the first month and a subsequent reversion of the immature to mature NK cell ratio [170-177]. The timing of NK cell reconstitution has been shown essential, as several studies have reported improved overall survival, reduced transplant-related mortality and infectious risk in recipients with an early and enhanced NK cell restoration [178, 179]. During this critical time dominated by complications such as acute graft-versus-host disease (aGVHD), bacterial and viral infections, recipients rely essentially on the innate immune system defenses.
- (iii) The governing feature of the third phase that extends up until two years post-transplant is the reconstitution of the cellular and humoral immunity – key to achieve complete immunocompetence and establish immune cell memory. At first, donor-derived mature T cells undergo peripheral homeostatic proliferation, followed by a progressive diversification

of the peripheral pool by the emergence of thymus-derived naïve T cells [139]. At last, the B cell compartment arises with a gradual reconstitution that can take up to several years.

Although incremental advances have been made to increase the success rate, this therapeutic option remains a high-risk treatment associated with various life-threatening adverse events [161]. Besides immunosuppressive linked complications, immuno-allogeneic events such as GvHD – a multi-system disorder affecting most commonly the skin, gastrointestinal tract and liver – enhance morbidity and mortality rates post-HSCT. GvHD management has become increasingly successful in recent years with the development of post-transplant prophylaxis such as cyclophosphamide in HLA mismatched settings and the refinement of HLA related donor allocation algorithms considering, e.g., HLA-DPB1 alleles. Tackling this issue however comes at the expense of the beneficial GvT effect as both pathways share a common immuno-allogeneic origin that requires (i) presence of immunocompetent T and NK cells within the transplant graft (ii) a certain degree of genetic incompatibility between donor and recipient, inducing direct and indirect recognition pathways for major and minor histocompatibility antigens (Fig.7) [180].

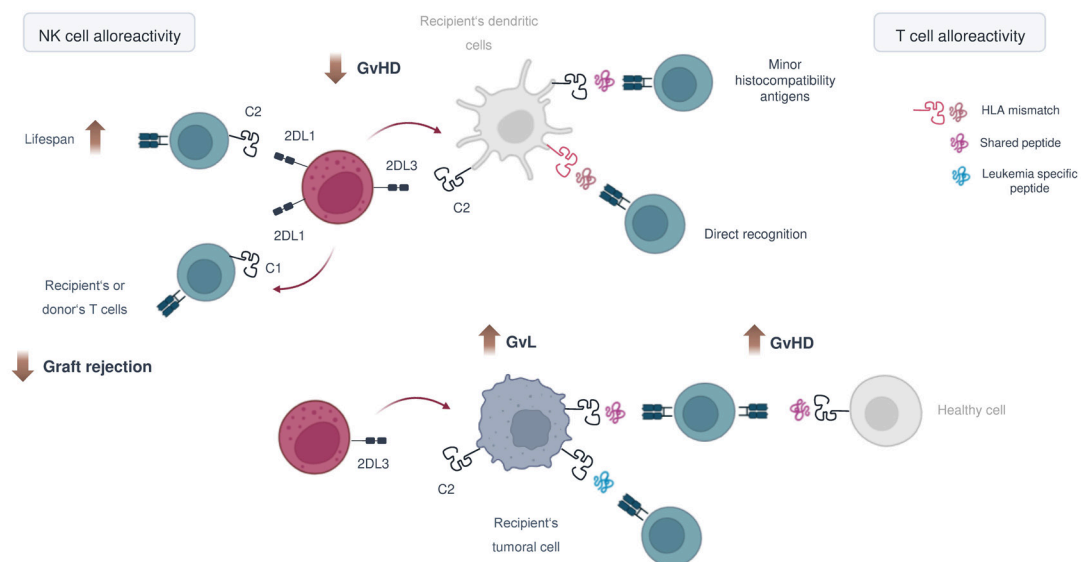


Figure 7: NK and T cell alloreactivity pathways in allogeneic HSCT. The main pathway by which T cells elicit alloreactivity is through direct recognition of HLA molecules and their cognate peptide in case of single HLA mismatches. Around 30% of HLA-matched HSCT exhibit GvHD as a consequence of the presentation of peptides derived from minor histocompatibility antigens on an HLA compatible molecule on APCs [180, 181]. NK cell alloreactivity is driven by a KIR/HLA mismatch that can be achieved in an HLA-mismatched setting through classical missing-self immunogenicity or in an HLA-matched setting through tolerance breakage by environmental priming. Thus, alloreactive NK cells are able to target recipient's leukemic blasts, reducing the disease relapse risk and recipient's DCs, thereby reducing the risk of T cell priming and GvHD induction. It is believed that NK cells are also able to kill T cells, however, recent data hints towards a KIR-HLA interaction increasing the lifespan of T cells that in turn might amplify the T cell alloreactivity [182].

3.1 Perspectives on therapeutic strategies to increase transplant outcomes

This duality greatly challenges the research community in the development of specific targeted therapeutic strategies attempting to boost GvT while abrogating GvHD [180, 183]. While a detailed review about the entire spectrum of therapeutic approaches is clearly out of frame for this introductory part, the following publications provide a precise overview on this topic [180, 184, 185].

Exploitation and refinement of the immunogenetic allocation system

The donor-recipient allocation system in allogeneic HSCT mainly relies on the HLA immunogenetic system, a key genetic barrier in transplantation. The current standard of care in HLA-matching is a 10 out of 10 (i.e., HLA -A, -B, -C, -DRB1, -DQB1) matching at a high-resolution in Europe and an 8 out of 8 matching in the United States and Japan (i.e., HLA -A, -B, -C, -DRB1). Several options are available when selecting such a suitable donor for a patient: HLA-identical siblings are favored with a 25% chance of being HLA-identical or a matched unrelated donor, which can be identified thanks to a worldwide registry encompassing a total of 40 million healthy volunteers [186]. In cases where matched donors are not available, mismatched unrelated donors with single HLA-I or II mismatches or haploidentical donors with one compatible HLA haplotype can be taken into consideration. The challenge of HSCT donor allocation is driven on the one side by the shortage of HLA-matched donors and on the other side – although less critical – the availability of several equivalent donors [187].

As a consequence, and in view of the high rate of transplant associated complications, research is undertaken to explore the impact of other immunogenetic determinants that could be used as additional stratification criteria to help clinicians in the donor allocation process.

This is exemplified by the case of HLA-DPB1: around 80% of recipients are transplanted across HLA-DPB1 allele mismatches, due to the low linkage disequilibrium and recombination hotspot between HLA-DPB1 and HLA-DRB1/DQB1. In recent years, there has been a fine-tuning in the evidence of HLA-DPB1 permissive and non-permissive mismatches on transplant outcomes, with a first model based on the allele expression level that informs the strength of an immune-allogeneic response [188] and the second T cell epitope structural (TCE) model states that different structural epitopes derived from HLA-DPB1 coding sequence polymorphisms result in various levels of immunogenicity. These encouraging results have partly led to their translation in the clinical routine [189].

Increase learnings from decisive non-genetic immune modulators and their impact on transplant outcomes

This immune system reconstitution is sensitive to immune challenges from the very beginning after transplantation. Adaptation and response to intrinsic and extrinsic factors bear the capacity to induce expansion and constrictions of specific cell subsets with the potential to induce long-lasting changes in immune phenotypes and functionality states [190, 191]. Gaining a comprehensive understanding of modulating factors on the immune state post-transplant and their effect on transplant outcomes could be beneficial to inform transplant risk predictors. A more detailed review on the nature of these factors is provided in the current state of research in chapter IV.

Exploitation of cellular – based strategies pre- and post-transplant

A well established and integrated practice is the process of *in vitro* T-cell depletion based on the rationale that donor T cells are the main effector cell in the GvHD process and is especially utilized in HLA-mismatched settings such as haplo-identical transplants. This process of graft manipulation has been sophisticated in recent years, with a more targeted depletion of key effector T cell subsets such as TCR $\alpha\beta$ + T cells [192] or naïve CD45RA+ T cells bearing the highest potential of inducing GvHD while maintaining less immunogenic T cell subsets for a certain level of immunocompetence [193, 194].

Another appealing strategy to magnify the immune state functionality is the adoptive transfer of NK cells – in a similar fashion to donor lymphocyte infusion – however with the value of being intrinsically tailored to elicit less GvHD. Several clinical trials have been conducted proving the feasibility of third-party NK cell infusions with encouraging results regarding an enhanced anti-leukemic effect [195-197]. Further optimization is on the way to determine the optimal infusion time and *in vitro* maturation polarization by cytokine stimulation through IL-2, IL-12, IL-18 and IL-15 [198, 199].

A few non-cell based immunotherapies have been developed with the aim to render the immune system development as early as possible after transplantation for it to be able to elicit a rapid GvT effect. To accelerate NK cell maturation and functionality, administration of immune molecules such as IL-15 superagonist (ALT-803) has been tested beneficial without major side effects [200].

Thesis aims

(A1) In this first chapter, we sort out to determine the predictive power of high-resolution KIR genotyping on post-transplant NK cell alloreactivity and its impact on transplant outcomes. To this end, we (C1) applied a state-of-the-art high-resolution sequencing and bioinformatic workflow to interrogate the KIR immunogenetic system at the allelic, CNV and genotype levels in a large-scale national HSCT transplanted cohort and (C2) proceeded to an associative analysis between donor-recipient KIR and HLA genotype determinants and various transplant outcomes (Fig.8).

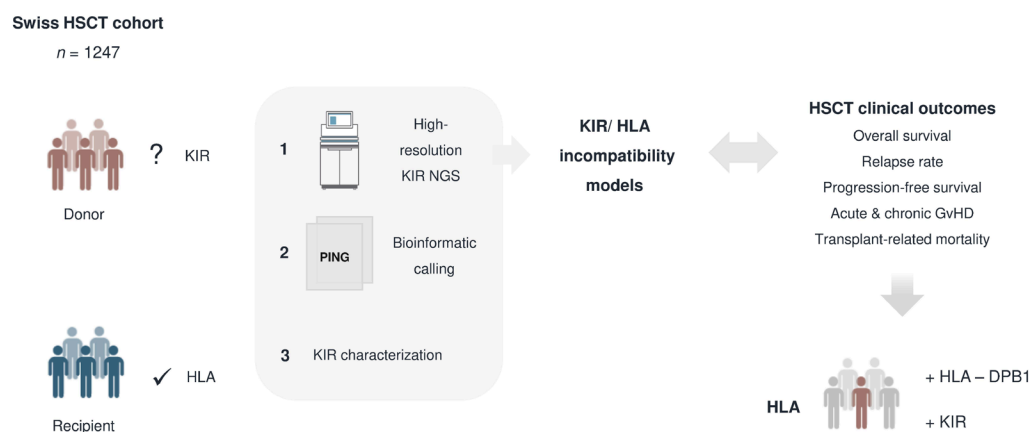


Figure 8: Graphical abstract of the first thesis aim.

(A2) In the second chapter, our aim was to complete a side-by-side comparison of two state-of-the-art technologies, time-of-flight mass cytometry and full spectrum flow cytometry, to deeply interrogate the repertoire of NK cells on a single-cell proteomic level. This was achieved by (C1) designing a high-dimensional single-cell proteomic panel, (C2) proceeding to a comparative assessment of the biological data output of mass cytometry and spectral flow cytometry by manual and semi-automated analysis and (C3) formulating key considerations in the decision-making process for the appropriate modality.

(A3) In the last chapter, we specifically set out to provide a deep and longitudinal profiling of the TCR and NK cell repertoire reconstitution alongside a time curve post-transplant and disentangle immune perturbations that govern immune development. To this end, we (C1) established a prospective Geneva-based HSCT transplanted cohort with a sampling scheme covering the first year post-transplant, (C2) leveraged TCR next-generation sequencing and high-resolution immunophenotyping for a system-level analysis and finally (C3) questioned the effect of viral exposure on the developing immune system post-transplant, with a specific focus on CMV – given its known profound immunomodulatory capacity (Fig.9).

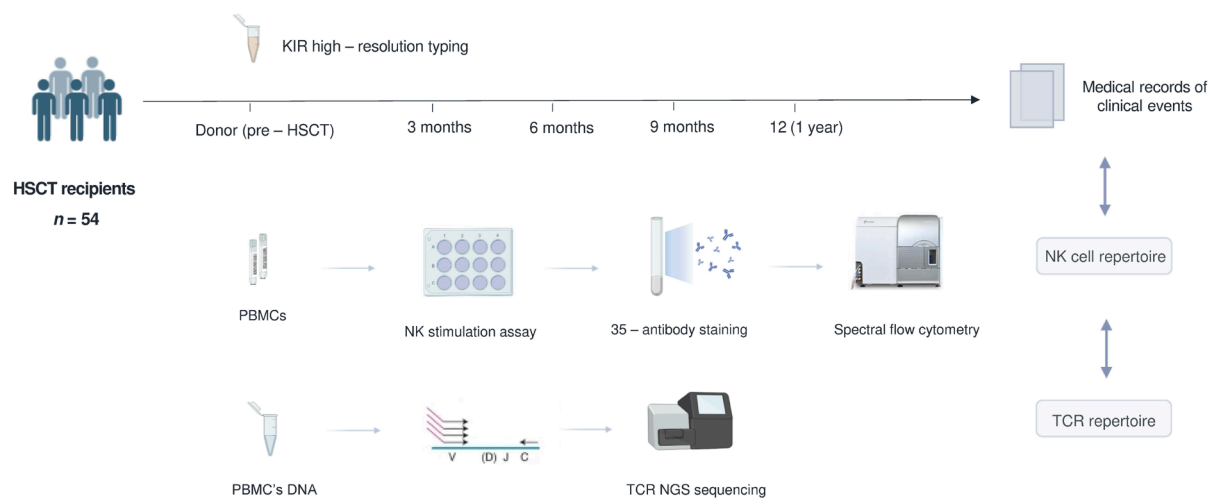


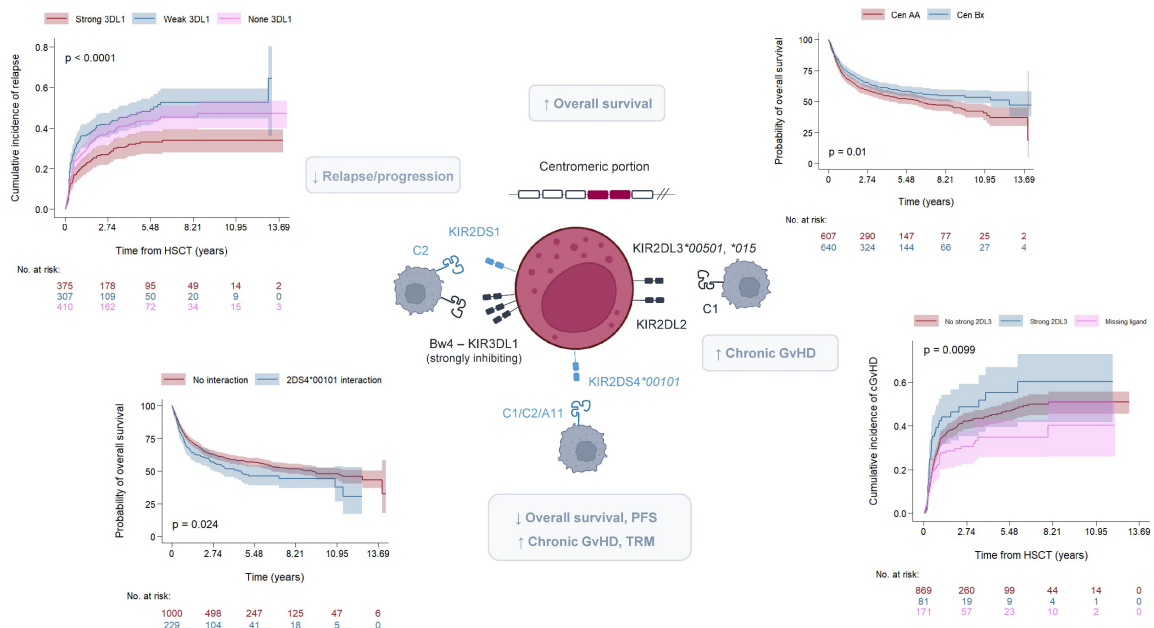
Figure 9: Graphical abstract of the third thesis aim.

CHAPTER II

High-resolution KIR genotyping and its implication in allogeneic hematopoietic stem cell transplantation

This chapter is based on following manuscript:

Schäfer, A., et al., *Integrating KIR high-resolution genotyping for predicting transplant outcomes in allogeneic hematopoietic stem cell transplantation*. 2024 (soon to be submitted).



1. Current state of research

The concept of NK cell alloreactivity in transplantation has raised from the assumption that alloreactive capabilities of NK cells can be triggered by a donor/recipient mismatch between inhibitory KIRs and their HLA-I ligands. The initial research work has impressively showcased this effect on reduced relapse rates following haploidentical transplantation and has become the foundation for years of extensive research [201].

Donor and recipient have indeed an inherent disposition of mismatching given the differential chromosomal segregation of KIR and HLA immunogenetic systems. Various biological principles rationalize the NK cell alloreactivity effect in different HLA incompatible and compatible transplant settings (Fig.10): HLA-haploidentical transplantation represents the most evident situation in which NK cell alloreactivity can be captured, at least on an immunogenetic theoretical level. KIR and HLA mismatch in this setting follows the classical NK cell missing-self principle, claiming sensitiveness of educated NK cells towards HLA-I downregulated abnormal cells. As a way of example, a C1/C2 donor bearing a KIR2DL1 would be alloreactive in a C1/C1 recipient. The question here isn't restricted to which NK cell subset elicits the most powerful anti-leukemic effect but additionally by which surrounding cell education is undertaken. It is commonly believed that donor hematopoietic cells such as DCs are majorly responsible for licensing NK cells as well as *cis* interactions within cells. Some evidence, however, suggests an additional contribution by the bone marrow stroma and non-hematopoietic cells from a recipient's origin. Yet, whether donor or recipient's hematopoietic and non-hematopoietic cells dominantly dictate the NK cell education is unclear – although decisive [169].

In a haploidentical setting, Vago *et al.* showed that NK cells have an abrogated functionality against leukemic targets at 1 month post-transplant compared to their donor counterparts. Single KIR⁺ NK cells (KIR2DL1 and KIR2DL3) with a predicted NK cell alloreactivity showed impaired cytotoxicity against HLA-I deficient cell lines, which partly recovered 2 years post-HSCT [174]. A recent study delineated the cytotoxic potency of alloreactive single KIR⁺ NK cells according to their immunogenic milieu during the first 6 months after a haploidentical transplantation: missing-self reactivity was prominent in KIR/HLA combinations in which donor and recipient expressed KIR ligands (e.g., KIR2DL1 from a C2/C2 donor transferred to a C1/C2 recipient), reaching the maximum in donor/recipient expressing all 3 KIR ligands (HLA-A,-B,-C). Notably, they observed a reduced functionality of predicted alloreactive NK cells in combination with a missing KIR ligand in the donor and/or in the recipient [202]. These observations are in accordance with the assumption of a linear correlation between the quantity of KIR/HLA interactions and the licensing level. Surprisingly, in another study, the functional competence of KIR2DL1⁺ NK cells was tuned downwards in an HLA-C2/C2 environment as compared to HLA-C1⁺ recipients and in turn correlated with worse clinical outcomes [203].

In the context of an HLA-matched transplantation but KIR mismatched, NK cells are thought to acquire tolerance to their environment strictly following immunogenetic rules, thus presumably abrogating any alloreactive potential. The concept of NK cell alloreactivity in this setting has however been adopted

from observations in autoimmune and infectious diseases in which the immunogenic potency of uneducated NK cells with inhibitory KIRs can be unleashed by the composite (i) acquisition of inhibitory KIRs and (ii) priming by environmental changes, e.g., pro-inflammatory cytokines. The NK cell functional competence is seen here as a continuous trait modulable by external changes. Yu *et al.* showed enhanced reactivity of NK cells for self and non-self HLA ligands (missing ligand in the host) 3 to 6 months after an HLA-matched T-cell depleted transplant [176]. Of note, NKG2A⁺ KIR⁻ NK cells were not responsible for the enhanced missing-self reaction observed. External validation, however, failed to replicate these findings – independent of the T cell depletion level – in which NKG2A⁻ NK cells expressing non-self KIR remained hyporesponsive compared to their educated counterparts at all time-points post-HSCT [154, 158].

Widely investigated, conflicting results have been gathered in the last 20 years that can be partly assigned to a large variety of genetic and non-genetic confounding parameters: KIR compatibility definition, HSCT protocol-related variables (underlying diagnosis, disease stage, conditioning and immunosuppressive regimes), donor and recipient related factors (age, gut microbiota), heterogeneity in cohort study sizes and the level of T cell depletion. While erasing each bias appears unrealistic, increasing standardisation across transplant, study and methodological protocols could greatly improve our attempts at linking genetics to a biological effect. Additionally, the genetic complexity of the KIR system encompassing gene, copy number and allelic variation combined with the lack of sequencing resolution depth and scalability have further challenged the interrogation of its effect and have limited the predictive value of KIR models to the exploration of KIR gene absence and presence.

The development of high-resolution and high-throughput KIR genotyping and increasing learnings about the functional aspects of KIR alleles have prompted the research community to reinterrogate the predictive power of the KIR immunogenetic system in transplantation. Several associative studies have started to test various KIR allele candidates, which are compiled in table 2.

Surely, one of the most promising research studies brought forward a model postulating that reduction in NK cell inhibition could unleash the NK cell anti-leukemic potency in oncological settings with HLA-I ligand expression maintenance following HLA-matched transplantation [204]. Detailed allelic sub-configuration analysis revealed a beneficial effect on the relapse incidence in weakly binding mode interactions (KIR3DL1^{high}/ HLA-Bw4 (80T) and KIR3DL1^{low}/ HLA-Bw4 (80I)) or with an abolished KIR3DL1 expression. The presence of highly inhibiting KIR3DL1-HLA-B (Bw4) interactions on the opposite was associated with a higher relapse rate. External validation in independent cohorts could, however, not replicate these results. The presence of non-expressed KIR3DL1 alleles (*019, *004) was reported to be detrimental on outcomes in HLA-matched transplantation [205, 206]. Pediatric recipients with a transplantation derived from a KIR2DL1-R²⁴⁵ alleles bearing donor – associated with a strong signaling transduction capacity – showed a higher disease-free progression and overall survival than recipients with a weak transduction capacity KIR2DL1-C²⁴⁵ alleles [207]. All these aforementioned studies focused solely on one specific KIR allele candidate. However, NK cell alloreactivity results from the additive effect of several KIR-HLA interactions, which substantiate the need to consider this system in its entirety – as far as biological pre-knowledge is available. In addition, effects of single KIR genes

might be biased due to linkage disequilibrium (e.g., absence of KIR2DS2 equates with the absence of KIR2DL2, the absence of KIR3DS1 equates with the presence of KIR3DL1).

The first study probing the implication of the entire KIR immunogenetic system and analyzing a large cohort of HLA-matched transplant recipients showed a trend towards a higher NK alloreactivity to transplant grafts from donors carrying ≥ 2 haplotype B segments, especially homozygous cen B02 donors. Involvement of highly praised KIR alleles such as KIR2DL1 and KIR3DL1 failed to be proven in the predictive model. These conclusions underpin previous results performed with low-resolution KIR genotyping, highlighting the protective effect of centromeric activating KIRs on disease relapse outcomes with KIR2DL2 and KIR2DS2 as potential significant NK cell alloreactivity drivers [208]. In the largest genetic association study conducted up to date with more than 5'000 transplanted recipients included, authors were unable to replicate any of the current proposed KIR models on transplant outcomes and were further unable to assess a link between KIR allele polymorphism and the level of post-transplant NK cell alloreactivity [209]. The discrepancies between theory and observed results might be due to the presence of clinical events and factors, the presence of other immune cells and the tumor microenvironment that might impede the NK cell alloreactivity post-transplant.

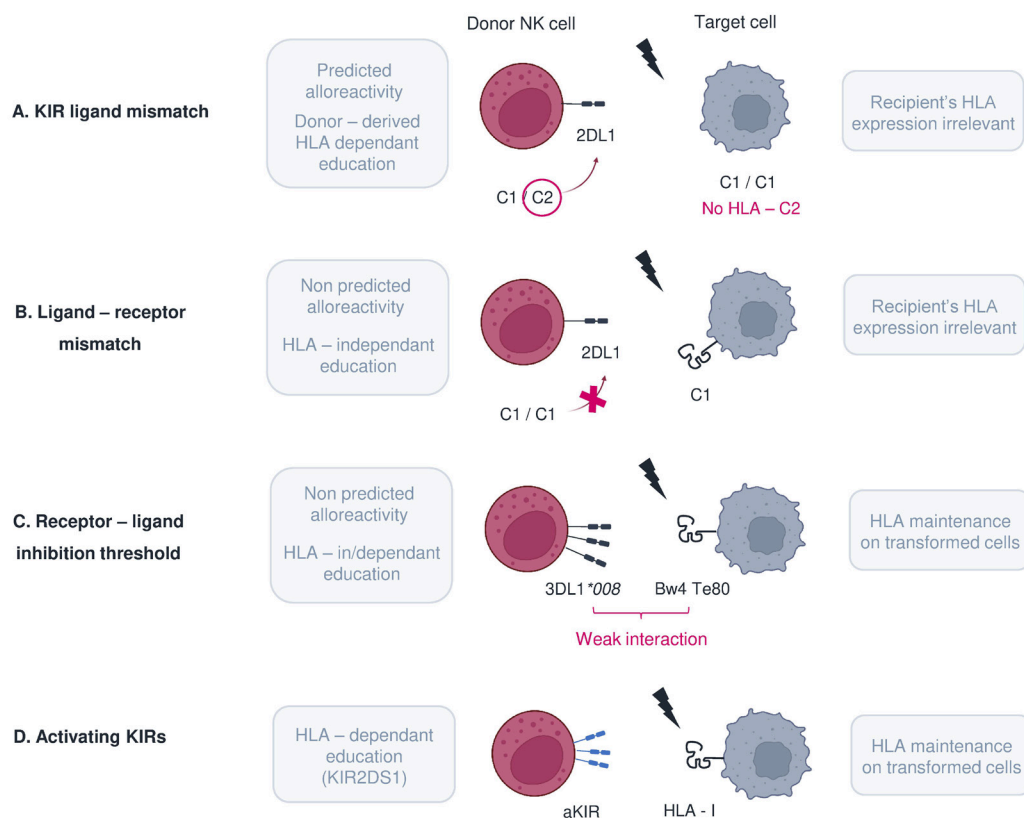


Figure 10: Different KIR and HLA models to rationalize NK cell alloreactivity post-transplant.

Table 2. KIR association studies in HSCT based on predictive high-resolution KIR and HLA genotyping.

Study	KIR typing	<i>n</i>	HLA	Disease	T depletion	Outcome
Boudreau et al. [204]	3DL1, 2DS1	1328	matched	AML	no	None and weak inhibiting KIR3DL1 and HLA-B (Bw4) interactions associated with reduced relapse rate and TRM; additive effect of KIR2DS1 ⁺ and HLA-C1/x interaction
Van der Ploeg et al. [210]	3DL1	604	matched	AML	mixed	Higher relapse rate in KIR3DL1 ⁺ (high/low) donors and HLA-A*24 recipients
Schetelig et al. [205]	3DL1, 2DS1	2222	matched	AML, MDS	mixed	External validation of Boudreau model: no beneficial effect observed on overall survival and relapse rate
Schetelig et al. [206]	3DL1,2DS1, haplotype	1704	matched	Secondary AML, MDS	mixed	No impact of KIR models on the risk of relapse and mortality
Schetelig et al. [209]	2DL1, 2DL2/L3, 3DL1, 2DS1, haplotypes	5017	matched	AML, MDS	mixed	No impact on relapse and TRM
Bari et al. [207]	2DL1	313	mixed	AML (pediatric)	mixed	Higher disease-free progression and overall survival with donors carrying KIR2DL1-R ²⁴⁵ alleles
Wright et al. [211]	2DL1	86	matched	AML	yes	Higher incidence of relapse and all-cause mortality with donors carrying KIR2DL1*003 alleles

Legrand et al. [212]	3DL1	186	haplo-identical	Mixed	no	High relapse rates in myeloid disease recipients with donors bearing non-expressed KIR3DL1 alleles (*004, *019)
Dubreuil et al. [213]	2DL1, 2DL23, haplotype	81	haplo-identical	Lymphoid and myeloid disease	no	Reduced incidence of relapse with donors carrying cen AA motif in myeloid disease recipients
Guethlein et al. [208]	2DS4, 3DL1, haplotype	890	matched	Lymphoid and myeloid disease	mixed	Reduced relapse incidence trend in donors with ≥ 2 B segments, especially homozygous cen B02 donors (potentially conferred by KIR2DL2 and/ or KIR2DS2)
Gowdavally et al. [214]	2DS4	2810	matched	Lymphoid and myeloid disease	mixed	Improved outcomes (OS, PFS, GvHD, TRM) in recipients with a lymphoid disease with donors lacking KIR2DS4 (deletion alleles or absent)
Shaffer et al. [215]	3DL1, 2DS1	527	matched	AML, MDS	mixed	D/R pairs with weak inhibiting KIR3DL1 – HLA-Bw4 interactions have reduced risk of relapse incidence

AML; acute myeloid leukemia, cen; centromeric, D; donor, GvHD; graft-versus-host disease, MDS; myelodysplastic syndrome, OS; overall survival, PFS; progression-free survival, R; recipient, TRM; transplant-related mortality

2. Manuscript I

Integrating KIR high-resolution genotyping for predicting transplant outcomes in allogeneic hematopoietic stem cell transplantation

Antonia Schäfer¹, Stéphane Buhler¹, Ticiana D. J. Farias², Katherine M. Kichula², Zuleika Calderin Sollet¹, Marie-Priscille Hervé¹, Sylvie Ferrari-Lacraz¹, Baptiste Micheli³, Stavroula Masouridi-Levrat⁴, Anne-Claire Mamez⁴, Helen Baldemero⁵, Jakob R. Passweg⁵, Dominik Schneidawind⁶, Vanessa Mesquita⁶, Tayfun Güngör⁷, Oliver Kürsteiner⁸, Cayathri Nair⁶, Jörg Halter⁵, Yves Chalandon⁴, Paul J. Norman², Jean Villard¹

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

²Department of Biomedical Informatics, University of Colorado School of Medicine, Aurora, CO, USA

³Genetic Medicine Division, Department of Diagnostic, Geneva University Hospitals, Geneva, Switzerland

⁴Division of Hematology, Department of Oncology, Geneva University Hospitals, Geneva, Switzerland

⁵Division of Hematology, Basel University Hospital, Basel, Switzerland

⁶Division of Medical Oncology and Hematology, University Hospital of Zurich, Zurich, Switzerland;

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

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

Corresponding author:

*Jean Villard, Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Geneva University Hospitals, Gabrielle-Perret-Gentil 4, 1211 Geneva 4, +41 22 372 93 94/ +41 79 553 34 09, jean.villard@hcuge.ch

Keywords

killer-cell immunoglobulin like receptor (KIR); hematopoietic stem cell transplantation (HSCT); natural killer cell; alloreactivity; high-resolution

ABSTRACT

The success of hematopoietic stem cell transplantation (HSCT) partly relies on the beneficial graft-versus-leukemia effect, mediated by alloreactive NK cells through their killer-cell Immunoglobulin-like receptor (KIR). Conflicting results have been reported regarding the impact of the KIR immunogenetic system on HSCT outcomes with a scarcity of data interrogating the effect of KIR allelic polymorphism. With the aim to fill this gap, donor KIR genes derived from a national cohort of 1247 HLA-matched transplanted donor/recipient pairs were determined at a high-resolution and tested in logistic regression models. D/R pairs bearing a KIR2DS4*00101 - HLA-C1/C2/A11 interaction showed a significant detrimental impact on progression-free survival (PFS), overall survival (OS), transplant-related mortality (TRM) and chronic GvHD (cGvHD) in multivariable analysis. Strong KIR2DL2/L3 – HLA-C1 and especially KIR2DL3*00501 and *015 interactions showed a significant increase in the incidence of cGvHD compared to missing ligand D/R pairs. Highly inhibitory KIR3DL1 – HLA-B and HLA-A (Bw4) interactions were associated with a reduced relapse incidence as compared to weak and none inhibiting interactions. Our study indicates that KIR high-resolution genotyping informs post-transplant outcomes with a seamlessly higher protection of educated NK cells.

INTRODUCTION

Allogeneic stem cell transplantation (HSCT) is currently the standard of care for hematological malignancies and other immunological disorders. Despite major advancements, HSCT remains a high-risk treatment and its success is significantly hampered by the occurrence of infectious, immunological complications and disease relapse [161]. Its therapeutic rationale partly relies on the beneficial graft-versus-leukemia (GvL) effect exerted by donor-derived alloreactive T and natural killer (NK) cells. NK cell immunosurveillance and killing capabilities are triggered by an HLA class I missing-self situation mediated by their germline-encoded killer cell immunoglobulin-like receptors (KIR) [216]. In a transplant setting, it has been initially hypothesized that a missing ligand situation could be mimicked by a donor and recipient KIR/HLA mismatch configuration, thus inducing alloreactivity against various target cells [201]. Subsequent research to dissect this effect on modulating transplant outcomes, especially the relapse propensity, has yielded highly conflicting results [81].

Transplant-related features have been recognized as potential confounding parameters in this association, such as *in vitro* t-cell depletion [217] or the type of the underlying diagnosis. The genetic complexity of the KIR system encompassing gene, copy number and allelic variation combined with the lack of sequencing resolution depth and scalability have further challenged the interrogation of its effect. Recent studies have allocated beneficial or detrimental effects to selected KIR allele candidates such as KIR3DL1 and KIR2DL1 on HSCT outcomes [204, 205, 207, 211, 212] and a scarcity of studies have been investigating the entirety of KIR loci simultaneously [208, 209]. In the largest genetic association study conducted up to date with more than 5'000 transplanted recipients included, authors were unable to replicate any of the proposed KIR models and were further unable to assess a link between KIR allele polymorphism and the level of post-transplant NK cell alloreactivity [209].

Given the importance of NK cells in antitumor and antiviral immunity and the pressing need to improve post-transplant outcomes, we sort out to revisit the predictive power of high-resolution KIR genotyping using a state-of-the-art KIR sequencing workflow in a large retrospective cohort of allogeneic HSCT transplanted recipients.

METHODS

Study cohort

For this multicenter retrospective study, 1247 patients who received a first allogeneic HSCT from an HLA-matched unrelated donor between January 2008 and April 2022 in one of the four transplant centers in Switzerland were selected. The study was approved by the local ethical committee for human studies of Geneva and the Geneva University Hospital (Commission Cantonale d’Ethique de la Recherche, CCER, CER 06-208 and 08-208) and performed according to the Declaration of Helsinki principles.

High-resolution KIR genotyping

All donors were genotyped at high-resolution for the KIR loci. To this end, a DNA probe-based capture method was used as described here [83]. Genomic DNA was directly extracted from whole blood samples and purified using the QIAGEN Blood and Tissue Kit according to the manufacturer’s instructions. DNA purity and concentration were assessed by Qubit. All samples were stored at 4°C until usage. For library preparation, 500 ng of genomic DNA was first fragmented by digestive enzymes (New England Biolabs, Boston, MA, USA), followed by barcode ligation with unique adaptors (IDT, Coralville, Iowa, USA). After post-ligation cleanup, dual size selection was performed with AMPure magnetic beads (Beckmann Coulter, Brea, California, USA) to acquire fragment sizes of 800 to 1200 bp length. In a second step, a pool of oligonucleotide probes specific for the KIR and HLA genomic regions was used for the targeted capture [83]. Final enriched libraries were normalized to a concentration of 12 pmol/l. Paired-end sequencing was performed using a NovaSeq instrument with a sequencing length of 2 x 250 bp (Illumina, San Diego, CA, USA).

High-resolution HLA typing

High-resolution HLA genotyping was performed on all recipients and donors using 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).

Computational and statistical analysis

PING bioinformatic pipeline

Pushing Immunogenetics to the Next Generation (PING) pipeline was applied for sequencing filtering, alignment, gene content and allelic genotype determination derived from the next generation sequencing fastq files as developed by Norman *et al.* and Marin *et al.* [82, 83]. High-resolution KIR genotype and copy number were determined for all KIR genes (KIR2DS1, 2DS2, 2DS3, 2DS4, 2DS3/S5, 3DL1/3DS1, 2DL1, 2DL2/L3, 2DL4, 2DL5A/B, 3DL2, 3DL3) and the two pseudogenes (KIR2DP1, 3DP1).

KIR alleles and haplotype assignment

Gene-level centromeric (cen) and telomeric (tel) haplotype assessment was done manually based on the KIR gene presence and copy number variation following current haplotype classification [218-220]. The centromeric portion is defined as KIR genes present in between the framework genes KIR3DL3 and KIR3DP1, while the telomeric part encompasses genes from KIR2DL4 to KIR3DL2. The presence of one or more of the following KIRs: 2DL2, 2DL5, 3DS1, 2DS1, 2DS2, 2DS3 and 2DS5 defines Bx haplotypes, while their absence marks an AA haplotype. KIR ambiguities were manually curated out. Haplotype, gene and allele frequency were set out by direct counting and the number observed was divided by $2N$ (alleles duplicated on a single haplotype were included as distinct loci, and gene absence was counted as a distinct allele).

KIR allotype assignment

KIR allotype refers to a distinct amino acid sequence and KIR alleles were grouped according to published known KIR allotypes: KIR3DL1 expression levels were classified into high, low and null as previously described [62, 63]. KIR2DS4 alleles were classified into the full-length version or the trunked variation [221]. KIR2DL1 alleles were segregated into strongly and weakly inhibiting based on their presence of an arginine or a cysteine group at position 245, respectively [71]. Functionally stronger KIR2DL2 and KIR2DL3 alleles were defined as alleles with a glutamic acid at position 35, whereas weak alleles were defined by the presence of a glutamine at position 35 [68] (Supplementary Tables S3 and S4).

KIR and HLA interactions

HLA-A, -B and -C alleles were categorized according to their relevant epitopes following known classifications using the Immuno Polymorphism Database [222]. KIR-HLA pairs were then summed up for each individual, with homozygous KIR or HLA alleles counted twice: HLA-C (C1 epitope) and HLA-B*46/ *73 with KIR2DL2 and KIR2DL3 [34], HLA-C (C2 epitope) with KIR2DL1 [34], HLA-C (C2 epitope) with KIR2DS1, HLA-C2 (*05:01, *02:02, *04:01), HLA-C1 (*16:01, *01:02, *14:02) and HLA-A*11 with KIR2DS4*001 [30, 32, 34], HLA-C1 (*16, *01:02) and HLA-A*11:01 with KIR2DS2 [30, 223, 224], HLA-A (Bw4 epitope) and HLA-B (Bw4 epitope) with KIR3DL1 [36, 37]. KIR2DS5*003, *004, *005, *006, *007, *008 with HLA-C2 [33], however, due to the very low number of interactions in our cohort, we did not account for this interaction. Assumption of KIR3DS1 with HLA-B (Bw4 epitope) as an interaction given its 97% sequence homology in its extracellular domain with the KIR3DL1 receptor without *in vitro* demonstration [225].

Statistical endpoints and analysis

Statistical endpoints analyzed were overall survival (OS: time from transplantation until death), progression-free survival (PFS; survival without evidence of active malignancy after transplantation), relapse and progression, and transplant-related mortality (TRM; time from transplantation until death without evidence of relapse). OS, TRM and PFS were censored at the last reported follow-up date. Further outcomes were the incidence of acute graft-versus-host disease (aGvHD) and chronic graft-versus-host disease (cGvHD).

A total of 18 variables were tested in univariable and multivariable analysis and treated as categorical variables except for copy number variation.

Survival functions for OS and PFS were estimated according to the Kaplan-Meier method starting from the baseline date to the event or the last follow-up date available and compared between groups using a log-rank test. Cumulative incidence rates were estimated for events with competing risks (i.e., TRM, relapse/progression, aGvHD and cGvHD) and compared between groups using a log-rank test.

All models were then tested in a multivariable analysis using cox proportional hazards regression. For the multivariable analysis, the following covariates were tested by forward in univariable analysis and by backward selection to eliminate the non-significant variables: recipient and donor age, disease type, HLA matching, Karnofsky performance status, EBMT risk score, disease status, conditioning regimen,

graft source, *in vitro* T-cell depletion, total body irradiation, comorbid conditions, donor/recipient cytomegalovirus matching, genre matching, transplantation center.

Descriptive results are presented as medians and interquartile ranges (IQR) for continuous variables and numbers and percentages for categorical variables. For hypothesis-testing analysis, all statistical tests were two-sided, with a threshold *p*-value of < 0.05 considered for statistical significance. All statistical and biological *in silico* analyses were computed using the statistical computing environment R, version 4.1.3 (R Core Team, Vienna).

RESULTS

Study cohort's characteristics

The median age of recipients at the time of transplantation was 51.77 (IQR: 33.19 – 62.03). The primary diagnoses included hematologic malignancies (acute myeloid leukemia (AML): 39.9%, myelodysplastic syndrome (MDS)/ myeloproliferative neoplasms (MPN): 20.4%, acute lymphoblastic leukemia (ALL): 13.5%), and non-malignant diseases (primary immunodeficiencies (PID): 4.57%). A majority (79.4%, $n = 990$) were a 10/10 HLA matched transplants and 19.5 % ($n = 243$) were a 9/10 HLA-matched. Myeloablative conditioning regimen was applied in 50.3% ($n = 627$) of transplants and 84.4% ($n = 1053$) of recipients received a peripheral blood stem cell graft. A minority of patients (7.78%, $n = 97$) received an *in vitro* T-cell depleted graft. The median follow-up time was 2.57 years (IQR: 0.88 – 5.21). Further recipient's and donor's demographic and transplant-related characteristics are compiled in Table 1 and 2.

D/R pairs with a centromeric Bx portion show a higher overall survival probability

To analyze the predictive power of the KIR immunogenetic system on transplant outcomes, donor KIR genes derived from a national cohort of 1247 HLA-matched transplanted donor/recipient (D/R) pairs were determined at a high-resolution. The detailed KIR allele characteristics are summarized in supplementary Tables S1 and S2. We tested individual KIR and HLA configurations on a total of six different transplant outcomes (OS, TRM, PFS, relapse/progression, aGvHD and cGvHD).

The KIR gene loci can be classified into group AA and Bx haplotypes, and further divided into centromeric and telomeric fragments (see Methods section). Recipients receiving a graft from a donor with a Bx haplotype had a significantly higher OS than D/R pairs with an AA haplotype in univariable analysis (Fig. 1A) and in cause-specific multivariable regression model (hazard ratio [HR] 0.71, 95% CI 0.59 – 0.87, $p < 0.001$). This effect was likely driven by the presence of a centromeric Bx portion which significantly increased the OS probability in univariable (Fig. 1B) and cause-specific multivariable regression model (HR 0.76, 95% CI 0.63 – 0.92, $p = 0.004$) whereas no effect of the telomeric portion on OS was noted (HR 0.96, 95% CI 0.79 – 1.16, $p = 0.685$). The effect of the centromeric Bx portion on OS was indirectly confirmed by the significant impact of the copy number variation of KIR2DL2 and KIR2DS2 (HR 0.82, 95% CI 0.7 – 0.95, $p = 0.008$). Segregating the centromeric portion into cA01, cB01 and cB02 as previously suggested [208], cA01/cB02 bearing D/R pairs showed a significantly increased

OS probability than all other centromeric motifs in multivariable analysis (HR 0.72, 95% CI 0.56 – 0.92, $p = 0.009$).

Presently, results are inconsistent as to whether the Bx protective effect is due to the presence of activating KIRs or the presence of inhibitor KIRs [226, 227], motivating us to further investigate the effect of centromeric KIR/HLA interactions. We did not observe any impact on transplant outcomes in D/R pairs with a KIR2DS2*00101 – HLA-C*16/01:02 and A*11:01 interaction (HR 1.15, 95% CI 0.9 – 1.46, $p = 0.253$). In line with the hypothesis that KIR2DL2 would be beneficial, we segregated the cohort according to whether D/R pairs have at least one inhibitory KIR2DL2 – HLA-C1 interaction compared to D/R pairs carrying only KIR2DL3 – HLA-C1 ligands or have a missing ligand status (C2/C2). D/R pairs bearing exclusively KIR2DL3 – HLA-C1 interactions had significantly lower overall survival probability in univariable (Fig. 1C) and cause-specific multivariable regression model (HR 1.38, 95% CI 1.13 – 1.69, $p = 0.002$) as compared to D/R pairs bearing at least one KIR2DL2 – HLA-C1 interaction.

KIR2DL1 and HLA-C2 interactions segregated according to whether they have a strong or weak signal transduction capacity had no effect on the OS probability (weak: HR 0.74, 95% CI 0.43 – 1.28, $p = 0.284$, strong: HR 1.06, 95% CI 0.87 – 1.3, $p = 0.56$).

KIR2DS4*001 represents the most frequent allotype and is the only encoding a functional activating receptor [221]. In our cohort, we observed that KIR2DS4*00101 alleles paired with its ligands from HLA-C1/C2/A11 epitopes and carried by 18.36 % ($n = 229$) of D/R pairs had a significantly reduced OS in univariable analysis (Fig. 1D) and cause-specific multivariable analysis (HR 1.26, 95% CI 1.0 – 1.59, $p = 0.047$). In addition, the interaction significantly negatively affected following transplant outcomes: TRM (HR 1.65, 95% CI 1.2 – 2.27, $p = 0.002$), PFS (HR 1.39, 95% CI 1.12 – 1.71, $p = 0.002$) when considered in cause-specific multivariable regression models. We did not observe an effect on the relapse/progression incidence.

Detailed results of the multivariable analysis for relapse incidence, OS, PFS, TRM, aGvHD and cGvHD are provided in the supplementary Table 5.

Inhibitory KIR2DL2 and KIR2DL3 – HLA-C1 interactions increase the likelihood of chronic GvHD

GvHD has been thought to be modulated by NK cells through an indirect pathway targeting recipient's dendritic cells or alloreactive T cells [228]. Analysis in univariable and cause-specific multivariable regression model revealed that D/R pairs having a missing ligand configuration had a significantly

reduced incidence of cGvHD as compared to recipients with inhibitory KIR2DL2 – HLA-C1 interactions (HR 0.69, 95% CI 0.5 – 0.94, $p = 0.018$).

Assuming that the inhibition threshold might play a role, we segregated KIR2DL2/L3 alleles according to their binding affinity with HLA-C1 as previously described [68]. Analysis in univariable (Fig. 2A) and cause-specific multivariable regression model revealed that – independently of the KIR gene type – the presence of two strongly or at least one strongly inhibitory allele significantly increased the risk of cGvHD as compared to D/Rs having only weakly binding KIR2DL2/L3 alleles (one strong: HR 1.36, 95% CI 1.05 – 1.76, $p = 0.02$, two strong: HR 1.27, 95% CI 1.0 – 1.62, $p = 0.049$) and the presence of one strong interaction significantly increased the incidence of aGvHD (HR 1.38, 95% CI 1.01 – 1.87, $p = 0.042$). Stratified according to the KIR gene type, especially the presence of D/R pairs bearing at least one strongly inhibitory KIR2DL3 – HLA-C1 interaction encompassing KIR2DL3*00501 and *015 alleles significantly increased the cumulative incidence of cGvHD in univariable (Fig. 2B) and cause specific multivariable analysis (HR 1.59, 95% CI 1.14 – 2.22, $p = 0.006$) as compared to D/R pairs without strong KIR2DL3 interactions.

Further, although not reaching the statistical significance, D/R pairs with KIR2DS1 alleles in a C2/C2 environment had a tendency towards a decreased cumulative incidence of cGvHD as compared to D/R pairs without a KIR2DS1 interaction (HR 0.63, 95% CI 0.37 – 1.05, $p = 0.07$).

KIR2DS4*00101 functional interactions were associated with a higher incidence of chronic GvHD in cause-specific multivariable analysis (HR 1.29, 95% CI 1.02 – 1.64, $p = 0.035$). No significant correlation was observed with acute GvHD (Supplementary Table 5).

KIR3DL1 and KIR2DS1 interactions are predictive of relapse/progression incidence

Incremental research has been conducted into examination of KIR and HLA configurations on the relapse propensity in AML transplanted recipients, especially with regards to KIR3DL1 and KIR2DS1 [31, 204].

We started by interrogating the effect of KIR3DL1 stratifying the entire cohort based on the previous known classification [102] without considering D/R pairs who possess HLA-A allotypes containing the Bw4 motif. D/R pairs bearing weak (HR 1.7, 95% CI 1.3 – 2.21, $p < 0.001$) and none inhibiting (HR 1.4, 95% CI 1.09 – 1.81, $p = 0.009$) KIR3DL1 – HLA-Bw4 interactions displayed a significantly increased incidence of relapse and progression compared to recipients bearing strongly inhibiting KIR3DL1 – HLA-

Bw4 interactions in cause-specific multivariable and univariable analysis (Fig. 3A). We next thought to restrict this analysis only taking into account recipients transplanted in the context of an AML diagnosis (table 1). In this subgroup, we confirmed that strongly inhibiting interactions confer protection against the relapse/progression rate as compared to weak and none inhibiting interactions in univariable (Fig. 3B) and cause-specific multivariable analysis (weak: HR 1.8, 95% CI 1.21 – 2.6, $p = 0.004$, none inhibiting: HR 1.7, 95% CI 1.16 – 2.4, $p = 0.006$).

There is evidence that HLA-A allotypes with Bw4 motifs are potent educators for KIR3DL1⁺ NK cells, except for HLA-A*25:01 and HLA-A*23:01 [229]. In consideration of these results, we integrated D/R pairs with HLA-A alleles encoding Bw4 epitopes and confirmed that – in the entire cohort and in the AML subcohort – they were not significantly different than strongly inhibiting KIR3DL1 – HLA-B Bw4 encoding D/R pairs (Cohort: HR 0.875, 95% CI 0.559 – 1.37, $p = 0.557$, AML: HR 0.59, 95% CI 0.28 – 1.3, $p = 0.168$).

Within the entire cohort, there was no effect of KIR2DS1 – HLA-C2 interactions on the relapse/progression incidence (HR 0.99, 0.79 – 1.23, $p = 0.911$). However, within the AML subcohort, D/R pairs with a KIR2DS1 – HLA-C2 interaction had a significantly lower relapse/progression incidence than D/R pairs lacking this interaction (HR 0.67, 95% CI 0.48 – 0.95, $p = 0.024$). We further refined D/R pairs segregating them according to whether they have a strong/weak or none inhibiting KIR3DL1 – HLA-Bw4 interaction combined with the absence or presence of a KIR2DS1 – HLA-C2 interaction. D/R pairs with a strong inhibitory KIR3DL1 – HLA-Bw4 and a KIR2DS1 – HLA-C2 interaction confer the highest protection against relapse compared to D/R pairs with a weak KIR3DL1 – HLA-Bw4 interaction and the lack of KIR2DS1 – HLA-C2 interaction (weak/no 2DS1: HR 3.1, 95% CI 1.53 – 6.3, $p = 0.002$, none/no 2DS1: HR 2.5, 95% CI 1.25 – 5.2, $p = 0.01$). Interestingly, the group with weak and none inhibiting KIR3DL1 – HLA-Bw4 interactions and the presence of a KIR2DS1 – HLA-C2 interaction did not have a significantly higher relapse/progression incidence (Weak/2DS1: HR 1.2, 95%CI 0.49 – 3.1, $p = 0.659$, None/2DS1: HR 2.0, 95% CI 0.91 – 4.4, $p = 0.086$) suggesting a protective effect of KIR2DS1 in these subgroups than the group with strong inhibiting KIR3DL1 – HLA-Bw4 interactions and the presence of a KIR2DS1 – HLA-C2.

Finally, it has to be noted that these KIR3DL1 and KIR2DS1 configurations did not aggravate the acute and chronic GvHD incidence in the entire cohort and in the AML subcohort (Supplementary Table S5).

DISCUSSION

In the following study, we aimed to challenge the immunogenetic hypothesis that specific allelic KIR/HLA configurations would enhance the NK cell alloreactivity propensity post-transplant, thus modulating transplant outcomes. Integrating current paradigms and extending to further KIR/HLA configurations, we could demonstrate that KIR high-resolution genotyping informs post-transplant NK cell alloreactivity mainly driven by KIR2DS4, KIR2DL2L3 and KIR3DL1 alleles.

We first revealed a detrimental effect of KIR2DS4*00101 interactions on almost all transplant outcomes in terms of a lower OS rate, lower PFS and higher risk of cGVHD and TRM while there was no effect on relapse/progression and aGVHD. These results partially overlap with recent studies demonstrating unfavorable transplant outcomes with KIR2DS4*00101, albeit none of these studies took into account the HLA ligands for KIR2DS4 [214, 230]. Poor outcomes were equally reported for patients carrying full-length KIR2DS4 in non-transplant settings such as HIV-1 and Sars-Cov-2 [231, 232], substantiating its potential deleterious effect. We could hypothesize that NK cell activation by KIR2DS4 might lead to a sustained induced inflammatory setting with the secretion of proinflammatory cytokines leading to detrimental paracrine side effects. It remains, however, elusive as to why this detrimental effect is only present with KIR2DS4 and lacking with other activating KIRs. From a genetic point of view, KIR2DS4 is the only activating KIR on the telomeric part from the A haplotype, which might hint towards either suppressing roles of co-receptors or a differential intrinsic mechanism.

We found that there is a strengthened protection for chronic GvHD development in KIR2DL2/L3⁺ D/R pairs with a missing ligand status and especially a higher risk conferred to recipients bearing strongly inhibiting KIR2DL3*00501 and *015 – HLA-C1 interactions. NK cell mediated alloreactivity has been shown to indirectly impact the development of chronic GvHD by targeting recipient's dendritic and T cells [228]. Thus, given the proinflammatory environment, NK cell tolerance could be broken and missing ligand situations be favorable in selectively killing target cells. In addition, there is compelling evidence that HLA-C expression plays a key role in modulating immune response. In stem cell transplanted settings, there is evidence that the highest expressed HLA-C allele (*14:02) was associated with the most striking risk of acute GvHD [233], which might substantiate that an increased inhibitory threshold hinders NK cell targeted cytotoxicity against dendritic and T cells.

The lacking effect of KIR2DL1⁺ D/Rs pairs with a missing ligand status (C1/C1) might be allocated to a reduced reconstitution pattern, reducing the NK cell alloreactivity potential [174]. Finally, another explanatory hypothesis that might account for this effect is based on a direct NK and T cell interaction. Indeed, a recent study showcased that the number of direct inhibitory KIR and HLA interactions between NK and T cells impacts the lifespan of T cells *in vitro*, with a higher number of inhibitory interactions leading to a prolonged longevity [182, 234]. Thus, we might hypothesize that NK cells could potentiate the level of T cell alloreactivity, which might in turn render GvHD development.

With regards to the anti-leukemic NK cell alloreactivity, out of our analysis, we might speculate that activating and inhibitory KIR do not contribute equally to leukemic control. We found that there is an increased protection in patients with highly inhibiting KIR3DL1 – HLA-Bw4 interactions, partially dissimilar to previous studies [204, 205, 215]. Indeed, none inhibiting KIR3DL1 – HLA-B (Bw4) interactions have been initially suggested to magnify anti-leukemic immunity through a reduced inhibition threshold and responsiveness acquired through a cytokine enriched environment [204]. These findings couldn't be replicated in a following prospective study, where only recipients with weak inhibiting interactions were subject to a decreased relapse incidence [215]. Breaking the tolerance might not occur outside of specific settings, such as infections and might thereby curtail the anti-leukemic effect.

The favorable association between inhibitory KIR3DL1 interactions and the relapse incidence in this study suggests that classical NK cell licensing might override missing ligand or none inhibiting configuration states in anti-tumor immunity. Recent data has demonstrated a higher frequency of HLA-A and HLA-B loss-of-function mutations but to a lesser extent HLA-C across multiple cancer types, which substantiates the implication of HLA-B specific KIR receptors in the sensing of transformed tumor cells [235]. Further, the expression of HLA-A and B ligands on cells is approximately 5 times higher than HLA-C [236]. Following the educational tuning model of NK cells [91], we might hypothesize that strongly inhibiting KIR3DL1⁺ NK cells might show an enhanced sensitivity for discrete changes in the HLA-Bw4 ligand expression. A recent study has shown similar results with a beneficial effect conferred by the presence of highly expressed KIR3DL1 alleles compared to non-expressed KIR3DL1*004 and *019 alleles in recipients following haploidentical transplantation [212].

Furthermore, we found a significant additive effect of KIR2DS1 – HLA-C2 interactions on the relapse incidence in AML recipients with a lacking effect within the entire cohort, suggesting that the tumor microenvironment might render the expression of activating KIR ligands.

Despite the large size of our cohort, we need to acknowledge some limitations of the present study. The lack of biological knowledge may lead to potential biological misinterpretation in the associative analysis. Missing biological information includes the ligand identity for KIR2DS3 and KIR2DL5 which might interfere with the analysis or the potential binding of activating KIRs such as KIR2DS4 to non-HLA ligands, which we cannot capture based on this analysis. Another major weakness is the fact that we ignore the HLA ligand distribution on target cells, a major component of NK cell licensing and responsiveness comprehension. Thus, our results do not preclude effects of KIR and HLA in subgroups of disease. The study design may have biased our results due to the retrospective nature and the heterogeneity of some transplant-related features hampering accurate comparability with other cohorts. Finally, multicentricity with differences in the post-transplant monitoring might have affected our results, although we have accounted for this effect in our multivariable analysis.

Notwithstanding these limitations, our results suggest that KIR high-resolution genotyping might be an additional immunogenetic stratification tool for use in clinical practice and might help clinicians in the donor allocation process. High-resolution sequencing across all loci might not be necessarily of great importance, but one could envision a smarter and more cost-efficient sequencing of selected KIR loci.

Authors' contributions

A.S, J.V. designed the study; A.S, M.P.H, Z.C.S, T.D.J.F, K.M.K performed experiments; B.M performed the bioinformatic gene and allele calling; A.S and S.B. performed statistical analysis; A.S and J.V. drafted the manuscript; all authors critically reviewed, edited the manuscript, and approved the final version.

Declaration of competing interest

All authors declare that there are no competing interests.

Acknowledgements

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

Data availability

Data will be made available on request.

MAIN FIGURES LEGENDS

Figure 1. (A) Kaplan-Meier estimates of the impact of AA (red) and Bx (blue) KIR haplotypes on the overall survival probability. Shaded bands represent the 95% confident interval. Non-adjusted *p*-values indicate the significance of the log-rank test. **(B)** Kaplan-Meier estimates of the impact of centromeric AA (red) and Bx (blue) KIR haplotypes on the overall survival probability. Shaded bands represent the 95% confident interval. Non-adjusted *p*-values indicate the significance of the log-rank test. **(C)** Kaplan-Meier estimates of the impact of the presence of ≥ 1 KIR2DL2 – HLA-C1 (red), KIR2DL3 – HLA-C1 (blue) and missing ligand (lila) interactions on the overall survival probability. Shaded bands represent the 95% confident interval. Non-adjusted *p*-values indicate the significance of the log-rank test. **(D)** Kaplan-Meier estimates of the impact of the presence (blue) and absence (red) of KIR2DS4*00101 interactions on the overall survival probability. Shaded bands represent the 95% confident interval. Non-adjusted *p*-values indicate the significance of the log-rank test.

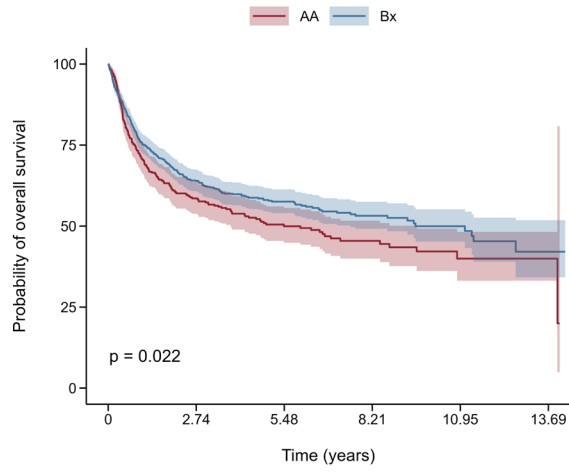
Figure 2. (A) Impact of two strongly (green), one strongly (blue), only weak (red) KIR2DL2/L3 – HLA-C1 interactions and missing ligand (lila) on the cumulative incidence of cGvHD. Shaded bands represent the 95% confident interval. Non-adjusted *p*-values indicate the significance of the log-rank test. **(B)** Impact of strongly inhibitory KIR2DL3 – HLA-C1 encompassing KIR2DL3*00501 and *015 alleles (blue), no strong KIR2DL3 – HLA-C1 interactions (red) and missing ligand (lila) on the cumulative incidence of cGvHD. Shaded bands represent the 95% confident interval. Non-adjusted *p*-values indicate the significance of the log-rank test.

Figure 3. (A) Impact of strongly inhibiting (red), weakly (blue) inhibiting and none (lila) inhibiting KIR3DL1 – HLA-B (Bw4) interactions on the cumulative incidence of cGvHD within the entire cohort ($n = 1247$). Shaded bands represent the 95% confident interval. Non-adjusted *p*-values indicate the significance of the log-rank test. **(B)** Impact of strongly inhibiting (red), weakly (blue) inhibiting and none (lila) inhibiting KIR3DL1 – HLA-B (Bw4) interactions on the

cumulative incidence of cGvHD within the AML subcohort ($n = 498$). Shaded bands represent the 95% confident interval. Non-adjusted p -values indicate the significance of the log-rank test.

1

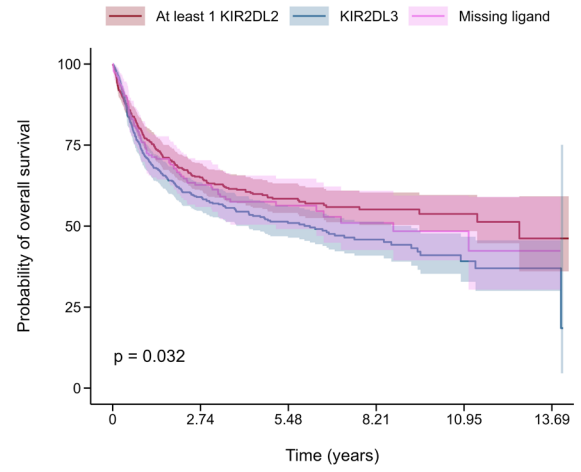
A



No. at risk:

406	188	96	49	18	2
841	426	195	94	34	4

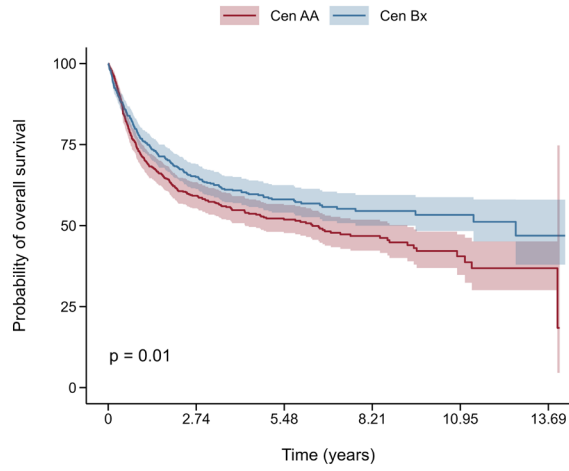
C



No. at risk:

549	277	128	60	23	3
512	243	120	61	21	2
186	94	43	22	8	1

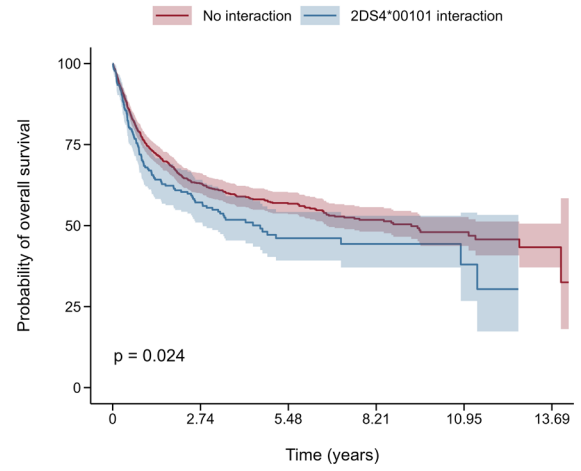
B



No. at risk:

607	290	147	77	25	2
640	324	144	66	27	4

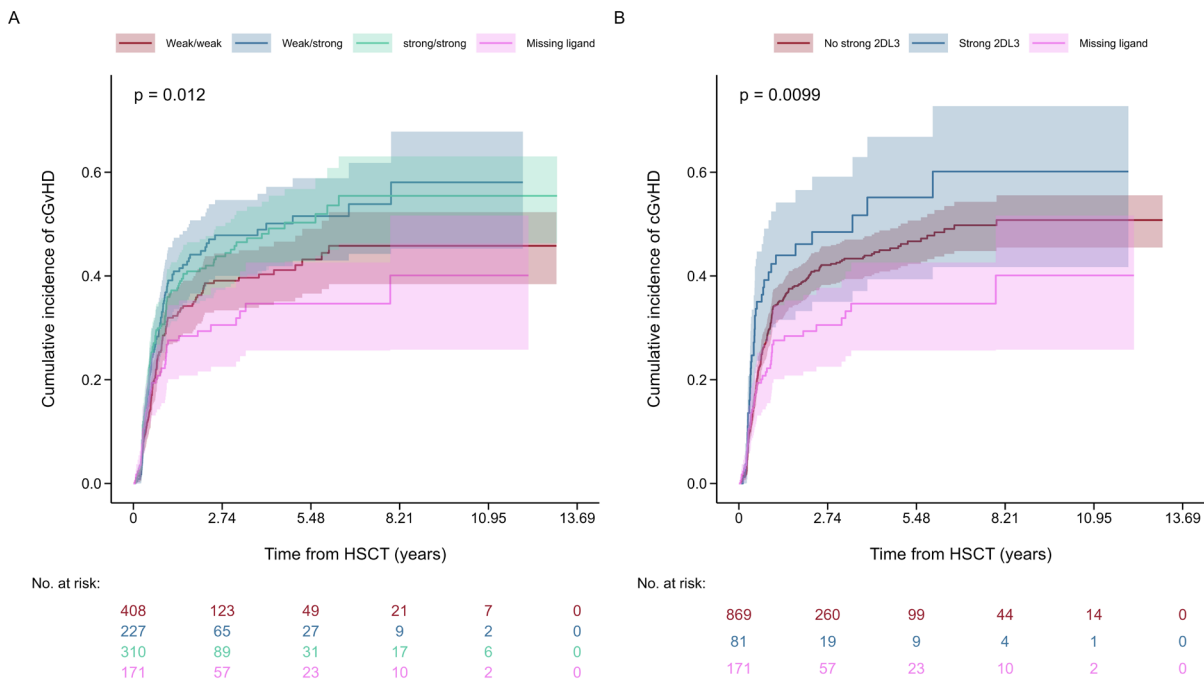
D



No. at risk:

1000	498	247	125	47	6
229	104	41	18	5	0

2



3

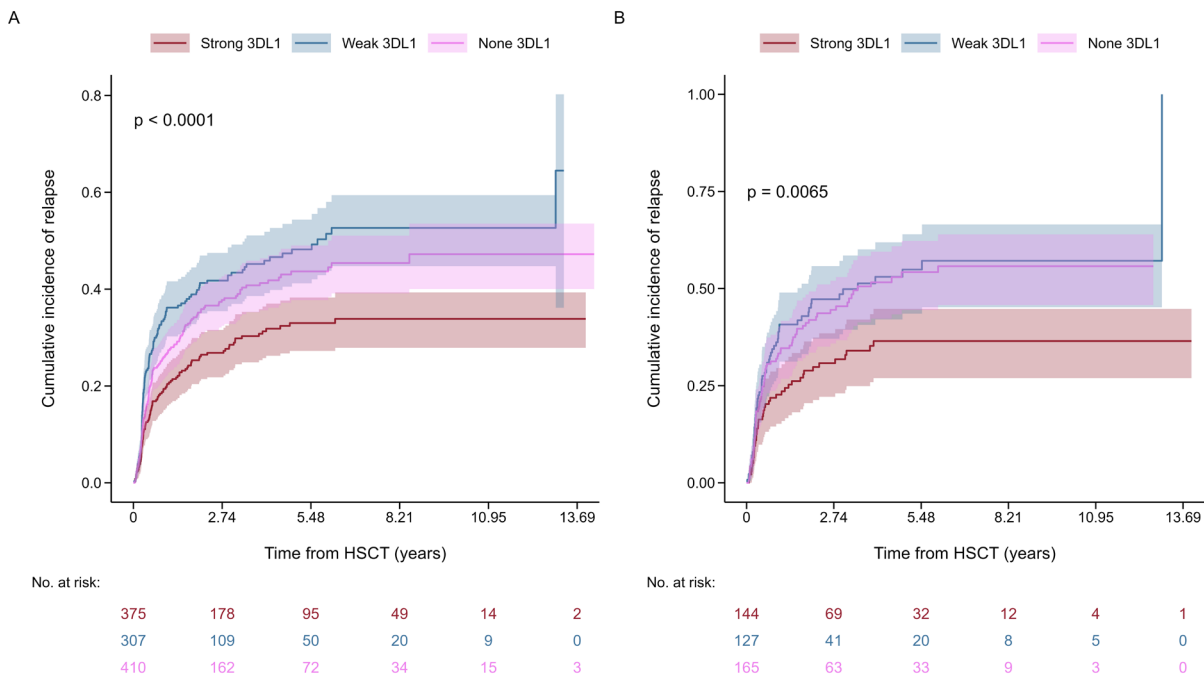


Table 1. Demographic and transplant related characteristics of the study cohort.

Parameter	Cohort (n = 1247)	AML (n = 498)
Recipient's age in yr (median, IQR)	51.77 (33.19 – 62.03)	54.6 (41.46 – 63.2)
Recipient's genre (M:F)	836 : 411	277 : 221
Donor's age in yr (median, IQR)	30.6 (25 – 39)	30.7 (24.2 – 39.1)
Donor's genre (M:F)	775 : 472	330 : 168
Year of transplantation, n (%)		
2008 – 2012	353 (28.3)	137 (27.5)
2013 – 2017	475 (38.1)	192 (38.5)
2018 – 2022	419 (33.6)	169 (33.9)
Transplant center, n (%)		
202	463 (37.1)	174 (34.9)
208	297 (23.8)	139 (27.9)
261	348 (27.9)	172 (34.5)
334	139 (11.1)	13 (2.6)
Underlying diagnosis, n (%)		
AML	498 (39.9)	498 (39.9)
ALL	168 (13.5)	-
CML	36 (2.9)	-
CLL	28 (2.25)	-
MDS	183 (14.7)	-
MPN	71 (5.7)	-
PCD	54 (4.3)	-
NHL	77 (6.17)	-
PID	57 (4.57)	-
BMF	35 (2.81)	-
Others	40 (3.2)	-
HLA-matching, n (%)		
10/10	990 (79.4)	400 (80.3)
9/10	243 (19.5)	95 (19.1)
Single mismatch at HLA-A	92 (7.37)	33 (6.6)
Single mismatch at HLA-B	33 (2.64)	8 (1.6)

Single mismatch at HLA-C	35 (2.8)	16 (3.2)
Single mismatch at HLA-DRB1	56 (4.49)	23 (4.6)
Single mismatch at HLA-DQB1	27 (2.16)	15 (3.01)
< 9/10	12 (0.96)	3 (0.6)
Missing information	2 (0.16)	-
Conditioning regimen, <i>n</i> (%)		
MAC	627 (50.3)	270 (54.2)
RIC	620 (49.7)	228 (45.8)
TBI		
No	887 (71.1)	379 (76.1)
Yes	358 (28.7)	118 (23.7)
Missing information	2 (0.16)	1 (0.2)
No T-cell depletion, <i>n</i> (%)	1150 (92.2)	454 (91.2)
Stem cell source, <i>n</i> (%)		
PBSC	1053 (84.4)	466 (93.6)
BM	186 (14.9)	31 (6.2)
CB	8 (0.64)	1 (0.2)
Disease state, <i>n</i> (%)		
Early	642 (51.5)	323 (64.9)
Intermediate	378 (30.3)	94 (18.9)
Late	227 (18.2)	81 (16.3)
Karnofsky Status, <i>n</i> (%)		
90 – 100%	970 (77.8)	405 (81.3)
≤ 80%	267 (21.4)	87 (17.5)
Missing information	10 (0.8)	6 (1.2)
EBMT risk score, <i>n</i> (%)		
0 – 1	62 (4.97)	14 (2.8)
2 – 3	539 (43.2)	291 (58.4)
4 – 5	523 (41.9)	166 (33.3)
6 – 7	123 (9.86)	27 (5.4)
CMV serostatus, <i>n</i> (%)		
D+ /R+	421 (33.8)	164 (32.9)
D- /R+	247 (19.8)	117 (23.5)
D+ /R-	129 (10.3)	57 (11.4)
D- /R-	436 (35)	160 (32.1)

Missing information	14 (1.12)	-
Genre matching (D/R), <i>n</i> (%)		
Male/Male	590 (47.3)	212 (42.6)
Female/Male	185 (14.8)	65 (13.1)
Male/Female	246 (19.7)	118 (23.7)
Female/Female	226 (18.1)	103 (20.7)

ALL; acute lymphoblastic leukemia, AML; acute myeloid leukemia, ATG; anti-thymocyte globulin, BM; bone marrow, BMF; bone marrow failure, CB; cord blood, CLL; chronic lymphoblastic leukemia, CML; chronic myeloid leukemia, CMV; cytomegalovirus, D; donor, IQR; interquartile range, MAC; myeloablative conditioning, MDS; myelodysplastic syndrome, MPN; myeloproliferative neoplasia, NHL; non-hodgkin lymphoma, PBSC; peripheral blood stem cells, PCD; plasma cell disorder, PID; primary immunodeficiency, RIC; reduced-intensity conditioning, TBI; total body irradiation

Table 2. Post-transplant clinical events and transplant outcomes of the study cohort.

Parameter	Cohort (<i>n</i> = 1247)	AML (<i>n</i> = 498)
Immunogenic post-transplant complications		
aGvHD, <i>n</i> (%)		
Yes	398 (31.6)	155 (66.3)
No	793 (63.6)	330 (31.1)
Missing information	56 (4.5)	13 (2.61)
aGvHD grade, <i>n</i> (%)		
0	516 (41.4)	225 (45.2)
1	277 (22.2)	105 (21.1)
2-4	398 (31.9)	155 (31.1)
Missing information	56 (4.5)	13 (2.61)
cGvHD, <i>n</i> (%)		
Yes	425 (34.1)	178 (35.7)
No	718 (57.6)	287 (57.6)
Missing information	104 (8.34)	33 (6.63)
Disease relapse, <i>n</i> (%)	425 (34.1)	199 (40)
Transplant-related mortality, <i>n</i> (%)	220 (17.6)	72 (14.5)
Overall survival, <i>n</i> (%)	708 (56.8)	230 (46.2)
Progression-free survival, <i>n</i> (%)	645 (51.7)	271 (54.4)
Median time of follow-up (IQR)	2.57 (0.88 – 5.21)	2.43 (0.87 – 5.03)

aGvHD; acute graft-versus-host disease, AML; acute myeloid leukemia, cGvHD; chronic graft-versus-host disease

Supplementary material

Antonia Schäfer¹, Stéphane Buhler¹, Ticiana D. J. Farias², Katherine M. Kichula², Zuleika Calderin Sollet¹, Marie-Priscille Hervé¹, Sylvie Ferrari-Lacraz¹, Baptiste Micheli³, Stavroula Masouridi-Levrat⁴, Anne-Claire Mamez⁴, Helen Baldemero⁵, Jakob R. Passweg⁵, Dominik Schneidawind⁶, Vanessa Mesquita⁶, Tayfun Güngör⁷, Oliver Kürsteiner⁸, Cayathri Nair⁶, Jörg Halter⁵, Yves Chalandon⁴, Paul J. Norman², Jean Villard¹

¹Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility,

Department of Diagnostic, Geneva University Hospitals, Geneva, Switzerland

²Department of Biomedical Informatics, University of Colorado School of Medicine, Aurora, CO, USA

³Genetic Medicine Division, Department of Diagnostic, Geneva University Hospitals, Geneva, Switzerland

⁴Division of Hematology, Department of Oncology, Geneva University Hospitals, Geneva, Switzerland

⁵Division of Hematology, Basel University Hospital, Basel, Switzerland

⁶Division of Medical Oncology and Hematology, University Hospital of Zurich, Zurich, Switzerland;

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

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

Corresponding author:

*Jean Villard, Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Geneva University Hospitals, Gabrielle-Perret-Gentil 4, 1211 Geneva 4, +41 22 372 93 94/ +41 79 553 34 09, jean.villard@hcuge.ch

Table of contents

Table S1. KIR alleles characteristics of activating KIRs

Table S2. KIR alleles characteristics of inhibitory KIRs

Table S3. Frequency of KIR2DS4, KIR2DL1, KIR2DL2L3 and KIR3DL1 allotypes according to their functional aspects

Table S4. Stratification of KIR2DS4, KIR2DL1, KIR2DL2L3 and KIR3DL1 alleles according to their functional aspects

Table S5. Multivariable analysis of KIRs variables on transplant outcomes

Table S1. KIR alleles characteristics of activating KIRs ($n = 1247$).

KIR2DS1	<i>n</i>	%	KIR2DS4	<i>n</i>	%
*null	1951	78.23	*null	551	22.1
*001	1	0.04	*00101	479	19.2
*00201	462	18.52	*00104	2	0.008
*00502	21	0.84	*00301	576	23.1
*006	31	1.24	*00401	71	2.846
*008	1	0.04	*00601	426	17.08
*011	2	0.08	*010	359	14.4
*unresolved	25	1.0	*016	1	0.004
KIR2DS2	<i>n</i>	%	*022	2	0.008
*null	1738	69.68	*unresolved	27	1.08
*00101	701	28.11	KIR3DS1	<i>n</i>	%
*00102	2	0.08	*null	1950	77.97
*00104	1	0.04	*01301	525	21
*00106	2	0.08	*014	2	0.008
*00107	1	0.04	*049N	19	0.76
*002	34	1.363	*1107	1	0.004
*02001	6	0.24	*unresolved	4	0.16
*005	1	0.04			
*007	1	0.04			
*008	1	0.04			
*unresolved	6	0.24			

Table S2. KIR alleles characteristics of inhibitory KIRs ($n = 1247$).

KIR2DL1	<i>n</i>	%	KIR2DL2L3			KIR3DL1		%
*00101	116	4.63	3*00101	957	38.7	*052	1	0.04
*00201	583	23.25	3*00110	14	0.56	*053	14	0.56
*00302	914	36.46	3*00201	591	23.6	*072	1	0.04
*00303	1	0.004	3*003	12	0.48	*089	1	0.04
*00401	281	11.21	3*00501	116	4.65	*110	1	0.04
*00402	12	0.48	3*006	1	0.04	* null	43	1.72
*007	34	1.36	3*009	3	0.12	* unresolved	50	2
*008	4	0.16	3*015	2	0.08			
*010	1	0.04	3*030	1	0.04			
*01201	1	0.04	*unresolved	45	1.8			
*01202	2	0.08	KIR3DL1	<i>n</i>	%			
*014	1	0.04	*00101	385	15.4			
*020	4	0.16	*00103	1	0.04			
*029	1	0.04	*00201	286	11.43			
*3201	8	0.32	*00401	360	14.4			
*03701	9	0.36	*00402	30	1.2			
*040	1	0.04	*00501	344	13.75			
*04301	5	0.2	*00701	64	2.56			
*044	4	0.16	*00801	149	5.96			
*049	1	0.04	*00901	18	0.72			
*05401	1	0.04	*01501	5	0.2			
*057	5	0.2	*01502	149	5.96			
*063	1	0.04	*01702	1	0.04			
*null	425	16.95	*019	17	0.68			
*unresolved	92	3.67	*02001	59	2.36			
KIR2DL2L3	<i>n</i>	%	*021	1	0.04			
2*00101	402	16.1	*02901	3	0.12			
2*00301	322	12.9	*03101	2	0.08			
2*00303	2	0.08	*033	3	0.12			
2*012	3	0.12	*039	1	0.04			
2*unresolved	23	0.92	*043	2	0.08			

Table S3. Frequency of KIR2DS4, KIR2DL1, KIR2DL2L3 and KIR3DL1 allotypes according to their functional aspects ($n=1247$).

KIR2DS4	<i>n</i>	%
<i>Expressed/ Expressed</i>	60	4.81
<i>Expressed</i>	361	28.9
<i>Not expressed</i>	808	64.8
Unresolved	18	1.44
KIR2DL1	<i>n</i>	%
<i>Strong/strong or strong</i>	854	68.48
<i>Strong/weak</i>	221	17.7
<i>Weak/weak or weak</i>	82	6.57
<i>unresolved</i>	51	4.1
KIR2DL2L3	<i>n</i>	%
<i>E³⁵ homozygous</i>	157	12.6
<i>E³⁵ / Q³⁵ heterozygous</i>	545	43.7
<i>Q³⁵ homozygous</i>	519	41.6
<i>unresolved</i>	26	2.1
KIR3DL1	<i>n</i>	%
<i>High/high or high</i>	644	51.64
<i>High/low</i>	167	13.4
<i>Low/low or low</i>	215	17.24
<i>Not expressed</i>	130	10.4
<i>unresolved</i>	30	2.41

E³⁵; glutamic acid at position 35, Q³⁵; glutamine at the position 35

Table S4. Stratification of KIR2DS4, KIR2DL1, KIR2DL2L3 and KIR3DL1 alleles according to their functional aspects ($n = 1247$).

KIR2DL1	Alleles	REF
Strong	<i>*00101, *00201, *00301, *00302, *00303, *00501, *008, *01201, *01202, *014, *010, *020, *040, *044, *049, *057, *063, *4301</i>	[71]
Weak	<i>*00401, *00402, *00701, *007, *05401, *029</i>	
Not expressed	<i>*03201</i>	
KIR2DL2L3		
2DL2 E ³⁵	<i>*00101, *00301, *00303, *012</i>	[68]
2DL3 E ³⁵	<i>*00501, *015</i>	
2DL3 Q ³⁵	<i>*00101, *01101, *00201, *003, *00501, *006, *00701, *009, *015, *030</i>	
KIR3DL1		
High	<i>*00101, *00103, *00201, *00801, *00901, *01501, *01502, *01702, *02001, *02901, *043, *052, *089</i>	[62, 63]
Low	<i>*00501, *00701, *033, *03101, *053</i>	
Not expressed	<i>*00401, *00402, *039, *019, *0072</i>	
KIR2DS4		
Expressed	<i>*00101, *00104</i>	[221]
Not expressed	<i>*00301, *00401, *00601, *010, *016, *022</i>	

E³⁵; glutamic acid at position 35, Q³⁵; glutamine at the position 35

Table S5. Multivariable analysis of the KIR variables tested on six different transplant outcomes ($n = 1247$).

		OS		TRM	
		<i>* Karnofsky score, disease state, disease type, presence of comorbidities</i>		<i>* Karnofsky score, disease type, transplant center, source of transplant</i>	
Variable	n	HR (95%-CI)	p	HR (95%-CI)	p
Haplotype					
AA	406	1		1	
Bx	841	0.71 (0.59 – 0.87)	< 0.001	0.73 (0.56 – 0.96)	0.026
Cen AA	607	1		1	
Cen Bx	640	0.76 (0.63 – 0.92)	0.004	0.76 (0.58 – 1.0)	0.051
Tel AA	742	1		1	
Tel Bx	505	0.96 (0.79 – 1.16)	0.685	0.93 (0.71 – 1.23)	0.627
Centromeric					
AA	587	1		1	
cA01/cB01	212	0.82 (0.63 - 1.07)	0.159	0.79 (0.54 – 1.17)	0.243
cA01/cB02	276	0.72 (0.56 – 0.92)	0.009	0.82 (0.58 – 1.16)	0.258
cB01/x	61	0.8 (0.51 – 1.26)	0.34	0.47 (0.2 – 1.06)	0.07
cB02/cB02	40	0.78 (0.44 – 1.38)	0.393	1.1 (0.53 – 2.27)	0.795
B content score					
Neutral	860	1		1	
Better	267	1.01 (0.8 – 1.27)	0.934	1.12 (0.81 – 1.54)	0.485
Best	120	0.83 (0.59 – 1.17)	0.289	0.7 (0.44 – 1.24)	0.249
KIR2DS1 – HLA-C2					
Absence	947	1		1	
Presence	300	0.95 (0.76 – 1.18)	0.648	1.03 (0.76 – 1.41)	0.847
KIR2DS1 – HLA -C2/C1					
Absence	947	1		1	
2DS1 – C1/x	233	0.93 (0.73 – 1.2)	0.59	0.98 (0.69 – 1.39)	0.902
2DS1 – C2/C2	67	1.00 (0.67 – 1.49)	0.998	1.21 (0.7 – 2.1)	0.488

KIR2DS2*00101 – HLA-C*16:01:02, A*11:01					
Absence	1035	1		1	
Presence	212	1.15 (0.9 – 1.46)	0.253	0.97 (0.68 – 1.39)	0.885
KIR2DS4*00101 – C1/C2/A11					
Absence	1000	1		1	
Presence	229	1.26 (1 – 1.59)	0.047	1.65 (1.2 – 2.27)	0.002
KIR3DS1					
Absence	779	1		1	
Presence	468	0.96 (0.79 – 1.17)	0.715	1.0 (0.76 – 1.31)	0.98
KIR3DS1 – Bw4 (HLA-B)					
Absence	946	1		1	
Presence	301	0.88 (0.71 – 1.1)	0.266	1.02 (0.75 – 1.4)	0.892
KIR2DL1 – HLA-C2					
Strong	671	1		1	
Weak	52	0.74 (0.43 – 1.28)	0.284	0.38 (0.14 – 1.03)	0.057
Missing ligand	429	1.06 (0.87 – 1.3)	0.56	0.83 (0.62 – 1.11)	0.204
KIR2DL2/L3 – C1					
≥ 1 KIR2DL2 – C1	549	1		1	
KIR2DL3 – C1	512	1.38 (1.13 – 1.69)	0.002	1.29 (0.96 – 1.73)	0.092
Missing ligand	186	1.0 (0.75 – 1.34)	0.995	1.2 (0.81 – 1.78)	0.36
KIR2DL2/L3 – C1					
Weak/weak	437	1		1	
Weak/strong	253	0.83 (0.65 – 1.07)	0.158	0.89 (0.62 – 1.29)	0.55
Strong/strong	342	0.72 (0.56 – 0.91)	0.007	0.68 (0.48 – 0.97)	0.033
Missing ligand	186	0.73 (0.55 – 0.99)	0.041	0.92 (0.62 – 0.97)	0.67
KIR2DL2/L3 – C1					
No strong 2DL3	949	1		1	
Strong 2DL3	88	0.95 (0.67 – 1.35)	0.77	0.97 (0.57 – 1.66)	0.923

Missing ligand	186	0.85 (0.65 – 1.12)	0.258	1.06 (0.73 – 1.52)	0.767
KIR3DL1 – HLA-B Bw4					
Strong inhibiting	375	1		1	
Weak inhibiting	307	1.23 (0.95 – 1.59)	0.112	1.06 (0.75 – 1.51)	0.734
None inhibiting	410	1.05 (0.82 – 1.35)	0.694	0.83 (0.59 – 1.17)	0.295
KIR3DL1 – HLA-B and A (Bw4)					
Strong inhibiting	375	1		1	
Weak inhibiting	307	1.23 (0.96 – 1.59)	0.105	1.06 (0.74 – 1.51)	0.748
None inhibiting	434	1.06 (0.83 – 1.35)	0.648	0.83 (0.59 – 1.16)	0.267
HLA-A*24 and A*32	100	1.27 (0.88 – 1.83)	0.209	1.1 (0.67 – 1.8)	0.696
KIR3DL1 and KIR2DS1					
Strong + 2DS1	85	1		1	
Weak + 2DS1	75	1.0 (0.59 – 1.69)	0.997	1.74 (0.85 – 3.57)	0.129
None + 2DS1	112	0.86 (0.51 – 1.43)	0.558	1.08 (0.52 – 2.26)	0.84
Strong without 2DS1	290	0.92 (0.6 – 1.42)	0.718	1.35 (0.74 – 2.48)	0.33
Weak without 2DS1	232	1.22 (0.6 – 1.42)	0.373	1.21 (0.64 – 2.29)	0.557
None without 2DS1	298	1.04 (0.68 – 1.58)	0.871	1.05 (0.56 – 1.95)	0.88

Variable	<i>n</i> = 1247	Relapse/progression		PFS	
		<i>*Age of recipient, disease type, disease state, risk score, source of transplant, transplant center</i>		<i>* Karnofsky score, disease state, disease type, presence of comorbidities, transplant center</i>	
		HR (95%-CI)	<i>p</i>	HR (95%-CI)	<i>p</i>
KIR haplotypes					
AA	406	1		1	
Bx	841	1.04 (0.843 – 1.27)	0.734	0.82 (0.69 – 0.98)	0.032
<i>Cen</i> AA	607	1		1	
<i>Cen</i> Bx	640	0.98 (0.808 – 1.19)	0.826	0.88 (0.74 – 1.04)	0.133
<i>Tel</i> AA	742	1		1	
<i>Tel</i> Bx	505	1.02 (0.841 – 1.24)	0.828	0.96 (0.8 – 1.14)	0.63
Centromeric					
AA	587	1		1	
cA01/cB01	212	1.23 (0.945 – 1.6)	0.123	1.01 (0.8 – 1.29)	0.867
cA01/cB02	276	0.88 (0.685 – 1.13)	0.318	0.83 (0.67 – 1.03)	0.093
cB01/x	61	0.95 (0.612 – 1.49)	0.834	0.72 (0.47 – 1.1)	0.126
cB02/cB02	40	1.0 (0.559 – 1.81)	0.988	1.02 (0.61 – 1.7)	0.947
B content score					
Neutral	860	1		1	
Better	267	0.99 (0.779 – 1.26)	0.945	1.09 (0.89 – 1.34)	0.407
Best	120	0.87 (0.621 – 1.23)	0.431	0.78 (0.57 – 1.06)	0.113
KIR2DS1 – HLA-C2					
Absence	947	1		1	
Presence	300	0.99 (0.79 – 1.23)	0.911	0.95 (0.78 – 1.16)	0.6
KIR2DS1 – C2/C1					
Absence	947	1		1	
2DS1 – C1/x	233	0.97 (0.754 – 1.24)	0.78	0.94 (0.76 – 1.18)	0.616
2DS1 – C2/C2	67	1.06 (0.7 – 1.62)	0.77	0.96 (0.67 – 1.39)	0.83

KIR2DS2 – HLA-C*16,01:02,A*11					
Absence	1035	1		1	
Presence	212	1.17 (0.92 – 1.49)	0.188	1.15 (0.92 – 1.43)	0.221
KIR2DS4*00101 – C1/C2/A11					
Absence	1000	1		1	
Presence	229	1.25 (0.982 – 1.59)	0.07	1.39 (1.12 – 1.71)	0.002
KIR3DS1					
Absence	779	1		1	
Presence	468	1.04 (0.857 – 1.27)	0.673	0.99 (0.83 – 1.18)	0.875
KIR3DS1 – HLA-B Bw4					
Absence	946	1		1	
Presence	301	1.0 (0.804 – 1.25)	0.972	0.97 (0.8 – 1.19)	0.796
KIR2DL1 – HLA-C2					
Strong	671	1		1	
Weak	52	0.78 (0.463 – 1.32)	0.361	0.67 (0.41 – 1.09)	0.108
Missing ligand	429	1.11 (0.907 – 1.37)	0.3	1.04 (0.86 – 1.25)	0.684
KIR2DL2/L3 – C1					
≥ 1 KIR2DL2 – C1	549	1		1	
KIR2DL3 – C1	512	1.08 (0.88 – 1.33)	0.456	1.19 (0.99 – 1.44)	0.059
Missing ligand	186	0.97 (0.722 – 1.3)	0.84	0.95 (0.73 – 1.23)	0.638
KIR2DL2/L3 – C1					
Weak/weak	437	1		1	
Weak/strong	253	0.97 (0.747 – 1.26)	0.808	0.91 (0.72 – 1.14)	0.408
Strong/strong	342	0.91 (0.711 – 1.16)	0.432	0.81 (0.65 – 1.0)	0.053
Missing ligand	186	0.89 (0.658 – 1.21)	0.455	0.79 (0.6 – 1.03)	0.079
KIR2DL2/L3 – C1					
No strong 2DL3	949	1		1	
Strong 2DL3	88	0.97 (0.67 – 1.42)	0.894	0.92 (0.66 – 1.27)	0.603
Missing ligand	186	0.93 (0.702 – 1.23)	0.605	0.86 (0.67 – 1.2)	0.233
KIR3DL1 – HLA-Bw4					

Strong inhibiting	375	1		1	
Weak inhibiting	307	1.7 (1.304 – 2.21)	< 0.001	1.44 (1.14 – 1.81)	0.002
None inhibiting	410	1.4 (1.090 – 1.81)	0.009	1.12 (0.89 – 1.4)	0.327
KIR3DL1 – HLA-Bw4					
Strong inhibiting	375	1		1	
Weak inhibiting	307	1.689 (1.299 – 2.2)	< 0.001	1.44 (1.14 – 1.81)	0.002
None inhibiting	434	1.387 (1.079 – 1.78)	0.011	1.11 (0.89 – 1.38)	0.366
HLA-A*24 - A*32	100	0.875 (0.559 – 1.37)	0.557	1.08 (0.77 – 1.52)	0.654
KIR3DL1 and KIR2DS1					
Strong + 2DS1	85	1		1	
Weak + 2DS1	75	1.26 (0.725 – 2.19)	0.412	1.25 (0.77 – 2.03)	0.358
None + 2DS1	112	1.37 (0.827 – 2.27)	0.221	1.07 (0.67 – 1.7)	0.778
Strong without 2DS1	290	0.98 (0.624 – 1.55)	0.94	1.04 (0.7 – 1.56)	0.836
Weak without 2DS1	232	1.82 (1.175 – 2.82)	0.007	1.58 (1.06 – 2.35)	0.025
None without 2DS1	298	1.04 (0.899 – 2.15)	0.138	1.19 (0.8 – 1.76)	0.395

Variable	<i>n</i> = 1247	Acute GvHD			Chronic GvHD	
		<i>*Disease type, risk score, conditioning regimen, t cell depletion, source of transplant, transplant center</i>			<i>*Age of recipient, disease type,transplant center</i>	
	<i>n</i>	HR (95%-CI)	<i>p</i>	<i>n</i>	HR (95%-CI)	<i>p</i>
KIR haplotypes						
AA	348	1		377	1	
Bx	750	0.95 (0.74 – 1.21)	0.654	766	1.12 (0.91 – 1.39)	0.276
<i>Cen</i> AA	527	1		563	1	
<i>Cen</i> Bx	571	1.09 (0.87 – 1.37)	0.446	580	1.19 (0.98 – 1.44)	0.079
<i>Te</i> / AA	643	1		682	1	
<i>Te</i> / Bx	455	0.9 (0.72 – 1.14)	0.391	461	1.03 (0.83 – 1.24)	0.802
Centromeric						
AA	508	1		545	1	
cA01/cB01	189	1.24 (0.91 – 1.69)	0.169	287	1.31 (1.0 – 1.7)	0.047
cA01/cB02	242	1.06 (0.79 – 1.43)	0.678	253	1.26 (0.99 – 1.6)	0.063
cB01/x	54	0.8 (0.43 – 1.48)	0.47	59	1.01 (0.64 – 1.6)	0.957
cB02/cB02	37	0.85 (0.43 – 1.68)	0.633	31	1.32 (0.79 – 2.21)	0.293
B content score						
Neutral	748	1		797	1	
Better	240	1.08 (0.82 – 1.43)	0.561	237	1.19 (0.95 – 1.5)	0.138
Best	110	0.78 (0.51 – 1.19)	0.247	109	0.95 (0.68 – 1.32)	0.742
KIR2DS1 – HLA-C2						
Absence	829	1		869	1	
Presence	269	1.01 (0.77 – 1.31)	0.959	274	1.07 (0.86 – 1.33)	0.553
KIR2DS1 – C2/C1						
Absence	829	1		869	1	
2DS1 – C1/x	209	0.99 (0.74 – 1.33)	0.947	213	1.21 (0.96 – 1.52)	0.112
2DS1 – C2/C2	60	1.07 (0.65 – 1.75)	0.8	61	0.63 (0.37 – 1.05)	0.077

KIR2DS2 – HLA-C*16,01:02,A*11						
Absence	912	1		946	1	
Presence	186	0.99 (0.74 – 1.34)	0.973	197	1.11 (0.87 – 1.42)	0.401
KIR2DS4*00101 – C1/C2/A11						
Absence	874	1		917	1	
Presence	208	1.23 (0.94 – 1.62)	0.134	208	1.29 (1.02 – 1.64)	0.035
KIR3DS1						
Absence	679	1		716	1	
Presence	419	0.9 (0.71 – 1.14)	0.38	427	1.06 (0.87 – 1.29)	0.54
KIR3DS1 – Bw4						
Absence	832	1		716	1	
Presence	266	0.96 (0.73 – 1.27)	0.792	427	1.03 (0.82 – 1.28)	0.842
KIR2DL1 – HLA-C2						
Strong	586	1		619	1	
Weak	48	0.68 (0.35 – 1.35)	0.272	49	1.13 (0.7 – 1.8)	0.62
Missing ligand	379	0.94 (0.73 – 1.2)	0.616	397	1.05 (0.85 – 1.29)	0.642
KIR2DL2/L3 – C1						
≥ 1 KIR2DL2 – C1	492	1		497	1	
KIR2DL3 – C1	447	0.92 (0.72 – 1.18)	0.535	475	0.84 (0.68 – 1.03)	0.087
Missing ligand	159	0.94 (0.67 – 1.34)	0.739	171	0.69 (0.5 – 0.94)	0.018
KIR2DL2/L3 – C1						
Weak/weak	380	1		408	1	
Weak/strong	227	1.38 (1.01 – 1.87)	0.042	227	1.36 (1.05 – 1.76)	0.02
Strong/strong	309	1.08 (0.81 – 1.46)	0.592	310	1.27 (1.0 – 1.62)	0.049
Missing ligand	159	1.06 (0.74 – 1.53)	0.747	171	0.87 (0.63 – 1.2)	0.399
KIR2DL2/L3 – C1						
No strong 2DL3	844	1		869	1	
Strong 2DL3	77	1.13 (0.73 – 1.74)	0.577	81	1.59 (1.14 – 2.22)	0.006
Missing ligand	159	0.95 (0.69 – 1.32)	0.776	171	0.77 (0.57 – 1.04)	0.089
KIR3DL1 – HLA-Bw4						
Strong inhibiting	330	1		342	1	
Weak inhibiting	276	1.06 (0.776 – 1.44)	0.722	279	1.11 (0.857 – 1.43)	0.442

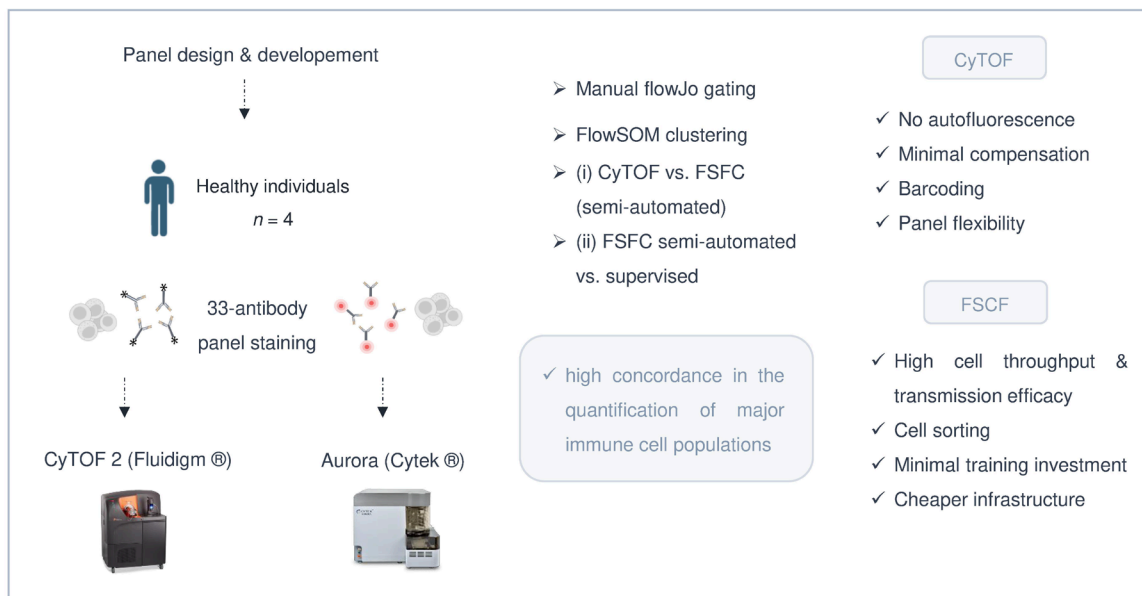
None inhibiting	355	1.08 (0.805 – 1.44)	0.616	382	0.99 (0.77 – 1.26)	0.945
KIR3DL1 – HLA-Bw4						
Strong inhibiting	330	1		342	1	
Weak inhibiting	276	1.05 (0.77 – 1.43)	0.769	279	1.1 (0.86 – 1.42)	0.447
None inhibiting	375	1.05 (0.79 – 1.4)	0.73	405	0.99 (0.78 – 1.25)	0.911
HLA-A*24 – A*32	90	1.01 (0.66 – 1.56)	0.951	88	0.92 (0.63 – 1.37)	0.695
KIR3DL1 and KIR2DS1						
Strong + 2DS1	78	1		79	1	
Weak + 2DS1	67	0.91 (0.467 – 1.76)	0.774	68	1.52 (0.921 – 2.51)	0.101
None + 2DS1	98	1.17 (0.661 – 2.08)	0.587	102	1.07 (0.647 – 1.75)	0.804
Strong without 2DS1	252	0.97 (0.591 – 1.58)	0.892	263	1.1 (0.716 – 1.68)	0.67
Weak without 2DS1	209	1.06 (0.646 – 1.75)	0.807	211	1.08 (0.698 – 1.68)	0.725
None without 2DS1	257	1.0 (0.613 – 1.64)	0.988	280	1.07 (0.696 – 1.64)	0.767

CHAPTER III

Comparative assessment of mass cytometry and full spectral flow cytometry

This chapter is based on following publication:

Schäfer, A., et al., *Comparative assessment of time-of-flight mass cytometry and spectral flow cytometry based on a 33-color antibody panel*. Journal of Immunological methods (2024).



1. Current state of research

The astonishing complexity of biological systems has driven the demand for accurate, highly performant and low-cost technologies to address biological questions in relation to health and disease. This challenge has brought propelling breakthroughs in the field of single-cell analysis with the emergence of so-called omics technologies covering transcriptomics, proteomics, metabolomics and lipidomics [237]. They have greatly enhanced our capacity to probe biological systems at a high-resolution with their ability to measure several parameters simultaneously in one assay. This technological shift has concomitantly paved the way towards a new wave of digitization. The advent of more sophisticated and analytical tools has been essential to extract and sort out valuable information out of biological big data and ease its interpretation. These high-precision and high-throughput technologies have ever since found a paramount importance in numerous research fields, including immunogenetics and have exponentially accelerated our knowledge of the characterization and functioning of such complex systems.

This section recapitulates the technical principle of the two single-cell proteomic platforms, cytometry by time-of-flight (CyTOF) and full spectrum flow cytometry (FSFC).

Using the unique properties of rare isotope metals combined with mass spectrometry, CyTOF has been the very first technology to fill the need for high parametric measurements on a single-cell proteomic level (Fig.11). The mass cytometer instrument is made of two main building blocks: inductively coupled plasma mass spectrometry and time-of-flight mass spectrometry. First, the sample composed of pre-stained cells is introduced into the CyTOF instrument in form of a liquid suspension. The sample is then aspirated to a concentric nebulizer responsible for generating single-cell droplets: after aerosolization, a shear force is applied to the sample by the nebulizer gas, breaking down the liquid into fine droplets. A heated spray chamber is positioned right after the nebulizer, which contains heated argon gas (~ 200°C) in charge of drying and evaporating the droplets. The single-cell aerosol droplets are at that timepoint dry enough to proceed to the ionization process taking place in the inductively coupled plasma (ICP) torch. Prior to this step, an aerosplitter located in between controls the entry flow rate of the gas stream into the torch. The vaporization, atomization and ionization processes – necessary for the subsequent mass analysis – are made possible by plasma, consisting of charged particles resulting in its electrically conductive state. In the mass cytometer, an electromagnetic coil is inductively heating the argon gas to an extremely elevated temperature (~ 7000 K) for the plasma formation. The argon gas serves now as an ionization source to fully decompose the sample into primary elements, free atoms, that are subsequently charged. The formed ion cloud is then passed through a high-pass ion optic aiming to remove low mass ions (< 80 kDa) such as carbon and oxygen. The remaining purified and high mass isotopic ions are then measured by the integrated time-of-flight (TOF) mass analyzer, where ions are accelerated by an electric field orthogonally towards the reflector and redirected towards the detector. The ion abundance can then be determined by the velocity of ions being directly proportional to the mass-to-charge ratio [238].

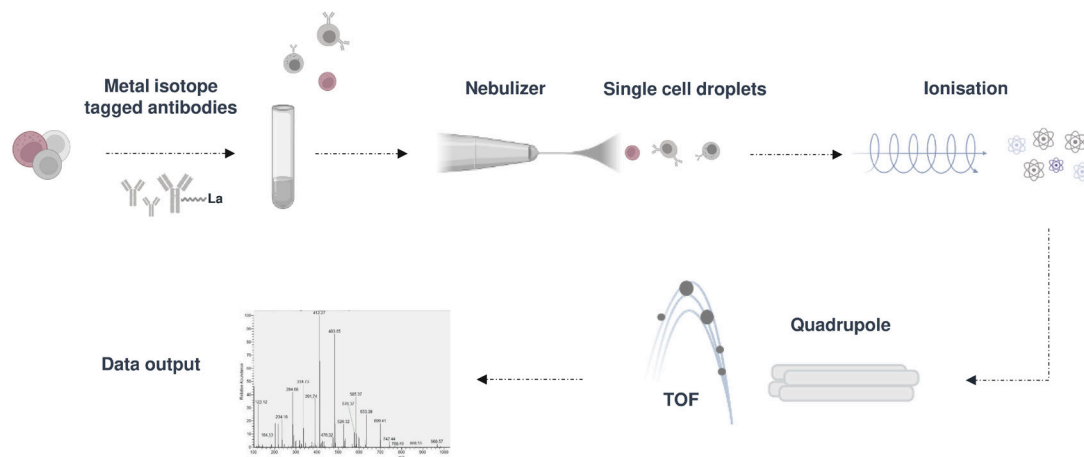


Figure 11: Overview of the principle of cytometry by time-of-flight.

The use of fluorochromes as reporters to measure single-cell profiles arised in the late 1960s in the Herzenberg laboratory at Stanford University. For several decades thereafter, this technology has been widely adopted by research and diagnostic laboratories, with the state-of-the-art acquisition being the simultaneous measurements of four to seventeen colors. The lack of computational power and the complexity in accurately resolving overlapping spectra have, importantly, hampered any further development. Firstly, introduced by Dr. Paul Robinson at the Purdue University in 2004, spectral cytometry has substantially revolutionized the field of flow cytometry. Several commercial instruments have been launched since then, with the most recent by Sony Biotechnology ® (ID7000) with 7 integrated lasers. The key innovation behind spectral flow cytometry is the combined use of high speed multianode photomultiplier tubes (PMT) connected to a dispersion device, such as prisms or spectrographs (Fig.12). As in conventional flow cytometry, a single-cell suspension with fluorescent-tagged cells is passed through a sample line to ensure constant velocity and linearity in the throughput and are then excited by laser beams detecting the fluorescent dyes. The emission fluorescence derived from the cells will be captured by the lasers, dispersed by the prisms and finally acquired with PMTs. The acquisition of full spectrum measurements by the 32 PMTs integrated in the instrument will then be subject to a process called spectral unmixing. This computational workflow consists in resolving the spectral signature or contribution of each fluorochrome from the baseline. Each signal derived from an event is supposed to be unique in terms of the spectral shape and intensity. Based on the acquired single-stained controls, the algorithm is then able to discriminate each spectrum derived from multiple dyes emitting in the 360 to 830 nm wavelength range [239-242].

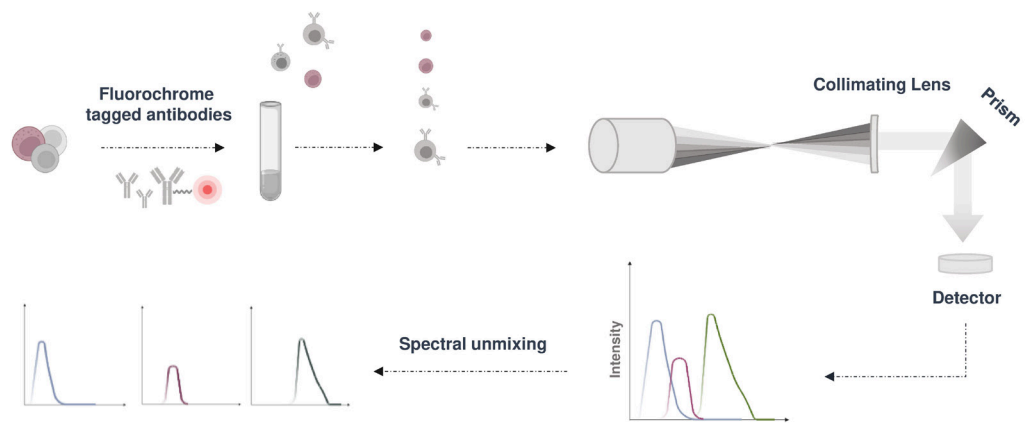


Figure 12: Overview of the principle of full spectral flow cytometry.

2. Manuscript II

Comparative assessment of cytometry by time-of-flight and full spectral flow cytometry based on a 33-color antibody panel

¹Antonia Schäfer, ²Sènan Mickael D'Almeida, ^{3,4}Julien Dorier, ^{3,4}Nicolas Guex,^{1*} Jean Villard, ²Miguel Garcia

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

²Flow Cytometry Core Facility, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

³Bioinformatics Competence Center, University of Lausanne, Lausanne, Switzerland

⁴Bioinformatics Competence Center, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Correspondence

*Jean Villard

Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Geneva University Hospitals, Geneva, Switzerland, jean.villard@hcuge.ch

Abstract

Mass cytometry and full spectrum flow cytometry have recently emerged as new promising single cell proteomic analysis tools that can be exploited to decipher the extensive diversity of immune cell repertoires and their implication in human diseases. In this study, we evaluated the performance of mass cytometry against full spectrum flow cytometry using an identical 33-color antibody panel on four healthy individuals. Our data revealed an overall high concordance in the quantification of major immune cell populations between the two platforms using a semi-automated clustering approach. We further showed a strong correlation of cluster assignment when comparing manual and automated clustering. Both comparisons revealed minor disagreements in the quantification and assignment of rare cell subpopulations. Our study showed that both single cell proteomic technologies generate highly overlapping results and substantiate that the choice of technology is not a primary factor for successful biological assessment of cell profiles but must be considered in a broader design framework of clinical studies.

Keywords

Mass cytometry; Full spectral flow cytometry; Clinical studies; High-dimensional analysis; Natural Killer cells

1. INTRODUCTION

Fundamental and clinical immunology research is highly dependent on the development and application of single-cell technologies to unravel the complexity of immune cell compartment. Over the past ten years, two innovative single cell proteomic platforms have emerged as new powerful research tools mitigating the number of markers that can be analyzed simultaneously by conventional flow-cytometry due to spectral overlap [242] and fulfilling the need for high dimensional analysis. Through the utilization of antibodies coupled to metal isotopes, cytometry by time-of-flight or mass cytometry (CyTOF), was the first pioneer technology to enable the simultaneous assessment of up to 50 markers [238, 243]. The recent advent of a next-generation flow cytometry technique, full spectrum flow cytometry (FSFC), also enables immunophenotyping on a high dimensional scale by the assessment of the full light spectrum emitted by each fluorochrome [241, 244-246] and a subsequent unmixing algorithm discriminating each unique fluorochrome by their specific spectrum [239, 240, 247].

These recent technical advancements in the field of single cell proteomics have provided researchers and clinicians with two similar tools to probe the immune system in greater depth. Their availability is a blessing and a curse at the same time, putting researchers and clinicians in front of a considerable challenge when it comes to choosing the appropriate technology. Given their differential technical nature, we aimed to proceed to a comparative study of the biological output of these two technologies – mass cytometry and full spectral flow cytometry – using an identical 33-color antibody panel. We further point out key considerations in the decision-making process for the appropriate modality.

2. MATERIALS AND METHODS

2.1 Sample preparation

Peripheral blood mononuclear cells (PBMCs) from four different healthy donors were isolated using Ficoll density gradient separation (Gibco). Briefly, whole blood was first diluted with the equal amount of phosphate-buffered saline (PBS) solution, layered onto the Ficoll-Paque (GE healthcare) and centrifuged at 400 x g for 20-25 minutes at 4°C. Subsequently, PBMCs were isolated from the interface, washed twice with PBS and cryopreserved in fetal calf serum (FCS) supplemented with 10 % dimethylsulfoxide (DMSO). Samples were then stored in liquid nitrogen until use. For the assay,

cryopreserved samples were rapidly thawed in a water-bath at 37°C and washed twice with warm RPMI-1640 (Gibco) supplemented with 10% FCS and-if required-with Pierce Universal Cell Nuclease. Thereafter, cells were counted, equally distributed in 5 ml polypropylene (PP) tubes with a total of 3-4 x 10⁶ cells per tube and rested overnight at 37°C in a medium consisting of RPMI 1640 supplemented with 10% FCS.

2.2 Full spectral flow cytometry

Samples were washed by pelleting with PBS at 400 x *g*. Viability staining was performed by adding 5 µl of a 1:500 diluted ViaDye Red viability staining solution (Cytek®) to the cells and incubated for 20 min at room temperature (RT) in the dark. Cells were subsequently washed with cell staining buffer (CSB). For Fc receptor blocking, cells were blocked with 10 µl of Fc-receptor blocking (Biosciences) solution for 10 min. Thereafter, 70 µl phenotyping antibody cocktail (Table S2; Fig. S1, S2) were added to each tube and samples were incubated for another 30 min at 4°C protected from the light. Following two washes with 2 ml of CSB, cells were fixed and permeabilized for the intracellular staining: To this end, cell were resuspended in 200 µl of Fix/Perm FoxP3 Solution (Foxp3/Transcription Factor Staining Buffer Set, eBiosciences, Thermo Fischer) and left for 20 min at RT in the dark. Cells were then washed twice by pelleting with 2 ml of FoxP3 permeabilization buffer at 800 x *g*. After fixation and permeabilization, 10 µL of intracellular antibody cocktail was added to the cells and incubated for another 30 min at 4°C protected from the light. Following the incubation time, cells were washed twice by pelleting with 2 ml of CSB, resuspended in a final volume of 400 µl CSB and finally filtered through a 35 µm nylon mesh filter for acquisition. Samples were acquired on a 5-Laser Aurora system (Cytek ®) using the SpectroFlo Software v2.2.0.2. The instrument was subject to daily quality control procedures using SpectroFlo® QC Beads (Lot 2004) as per the manufacturer recommendations.

2.3 Time of flight mass cytometry

Samples were washed by pelleting with PBS at 400 x *g*. Viability staining was performed by resuspending cells in cisplatin viability stain (Fluidigm) at a concentration of 25 µM for 1 min and quenched by adding 5 ml CSB. Cells were subsequently washed with CSB. For Fc receptor blocking, cells were blocked with 10 µl of Fc-receptor blocking (Biosciences) solution for 10 min. Thereafter, 70 µl phenotyping antibody cocktail (Table S1) were added to each tube and samples were incubated for another 30 min at 4°C. Following two washes with 2 ml of CSB, cells were fixed and permeabilized for

the intracellular staining: To this end, cells were resuspended in 200 µl of Fix/Perm FoxP3 Solution (Foxp3/Transcription Factor Staining Buffer Set, eBiosciences, Thermo Fischer) and left for 20 min at RT. Cells were then washed twice by pelleting with 2 ml of FoxP3 permeabilization buffer at 800 x *g*. Prior to the staining, samples were barcoded: each sample to be barcoded was resuspended in 800 µl of barcode permeabilization buffer, simultaneously, each barcode was resuspended in 100 µl and finally transferred to the samples. Thereafter, 10 µL of intracellular antibody cocktail was added to the cells and incubated for another 30 min at 4°C. Following the incubation time, cells were washed twice by pelleting with cell staining buffer, resuspended in 400 µl of intercalator solution containing 1.6% PFA and 0.5 µM iridium-intercalator (Fluidigm) and left at 4° C overnight. The following day, cells were washed twice by pelleting with 2 ml CSB, followed by two washes with cell acquisition solution. For the acquisition, cells were mixed with EQ calibration beads at a ratio of 1:5, further diluted at 1:10 with cell acquisition buffer and acquired on a CyTOF2.1 mass cytometry instrument (Standard Biotools).

2.4 Quantification and statistical analysis

The following preprocessing steps were applied to these data sets before further downstream analysis:

(i) Spectral unmixing for FSFC data was done using SpectroFlo Software v2.2.0.2 (Cytek ®). (ii) Initial data normalization including beads and debarcoding of the CYTOF data was done using the CyTOF software (Standard Biotools). (iii) Both datasets were then uploaded in FlowJo v.10.7.2 and subject to manual filtering steps on debris and time evolution scale. Thereafter, samples were pre-gated on single, alive, CD45+ immune cells and exported for downstream analysis.

For the semi-automated analysis, cleaned fcs files from FSFC and CyTOF experiments were loaded and preprocessed using R v4.2.2 [248] package flowCore v2.10.0 [249]. Data from all samples were concatenated. Raw marker expression values were transformed using inverse hyperbolic sine (asinh) transform with cofactor 5 for CyTOF dataset and 3'000 for FSFC dataset. For each marker, transformed expression values were normalized using a linear transformation to map the 1st and 99th percentiles to 0 and 1, respectively. Clustering was performed with FlowSOM v2.6.0 [250] with default parameters used except for the self-organizing map size (xdim=20, ydim=20) and the number of clusters for metaclustering (nClus=30). FlowSOM clustering was performed on the normalized expression values of the following markers: CD3, CD4, CD8a, CD19, CD14, CD56, CD16 and CD57. Clusters were manually assigned to cell populations by inspecting 1D and 2D distributions of marker expression values and grouped to a final of 15 distinct assigned cell populations: CD8+ T cells, CD4+ T cells, CD8+ CD4+ T

cells, CD4- CD8- T cells, CD4+ NKT cells, CD8+ NKT cells, CD4- CD8- NKT cells, CD4+ CD8+ NKT cells, CD19+ B cells, CD14+ monocytes, CD56++ NK cells, CD56+ CD16+ CD57- NK cells, CD56+ CD16- CD57- NK cells, CD56+ CD16- CD57+ NK cell, CD56+ CD16+ CD57+ NK cells (Fig. S3).

Uniform Manifold Approximation and Projection (UMAP) was performed with uwot v0.1.14 [16] using CD3, CD4, CD8a, CD19, CD14, CD56, CD16 and CD57 normalized expression values of 10e5 randomly sampled events. Remaining events were subsequently projected onto the UMAP embedding. ggplot2 v3.4.0 [251] was used for visualization.

For supervised analysis, cleaned fcs files were loaded in FlowJo 10.7.2 and the 15 cell populations of interest (see above) were manually gated. FlowJo workspace with manual gating information was then loaded into R using packages CytoML v2.10.0 [252] and flowWorkspace v4.10.0 [253]. The Matthews Correlation Coefficient (MCC) was used to compare cell populations obtained with semi-automated analysis and manual gating for the FSFC data. For each cell population, $MCC = (TP \cdot TN - FP \cdot FN) / \sqrt{((TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN))}$ where TP, TN, FP and FN are the number of true positive, true negative, false positive and false negative cell assignments to the selected population by clustering, arbitrarily considering manually gated cell assignment as the gold-standard (MCC is invariant under the exchange of gold-standard) [254].

3. RESULTS

We first designed and built a 33-color antibody panel to comprehensively quantify major immune and NK cell subpopulations based on extra- and intracellular markers (Tables S1, S2; Fig. S1, S2). This panel encompasses major NK cell receptor classes such as natural cytotoxicity receptors (NCRs), C-type lectin like receptors (NKG2) and activation markers such as perforin, granzyme B and CD107a [255-258]. To ensure that a relevant comparison of the technologies could be made, we developed identical panels for full spectral flow cytometry and mass cytometry encompassing the exact same antibodies. Due to the limited availability of antibodies, identical clones could only be allocated to the same antibody for part of the panel. The panel development and validation strategy was performed as follows: (i) We allocated antibodies to fluorochromes by associating low expressed targets such as KIRs to brighter fluorochromes and high expressed targets such as lineage markers to dim fluorochromes. (ii) We then proceeded to antibody titrations to ensure optimal staining patterns and minimize unspecific

binding by calculating the appropriate concentration. (iii) To accurately assess the positivity threshold of low expressed antigens and the specificity of the staining, we included a fluorescent-minus-one (FMO) control including following lineage markers: CD3, CD8a, CD4, CD19, CD14, CD56, CD16, CD57. (iv) We finally applied the full panel by staining non-stimulated and stimulated cryopreserved PBMCs derived from four healthy individuals.

Comparison of cell population abundances based on a semi-automated approach between CyTOF and FSFC

We conducted the comparative study on PBMCs samples derived from four healthy individuals. Aliquots from each individual were equally labeled with the CyTOF panel and the FSFC panel prior to machine acquisition. We firstly wondered about the consensus between the two platforms in distinguishing major canonical immune cell subsets selecting eight basic lineage markers: CD3, CD4, CD8a, CD19, CD14, CD56, CD16, and CD57. We analyzed each dataset by a supervised immune clustering labeling step using semi-automated approach based on unsupervised clustering with FlowSOM followed by a supervised assignment of clusters into 15 cell types (see materials and methods for details). With this approach, 95% of CyTOF events and 96% of FSFC events were assigned to a cell population and only the small remaining proportion couldn't be classified. Cell frequencies varied over a wide range, from 0.0039% to 52.1% per sample/per population. The frequency distribution of the cell clusters for each technique and sample is depicted in Figure 1A.

The proportion of each cell population derived from the two techniques was highly concordant as reflected by the Pearson's correlation coefficient of $r = 0.967$ ($p\text{-value} = 1.9 \times 10^{-38}$) stratified per cell population and per sample (Fig. 1B). As these correlative results can be biased by the presence of highly abundant cell populations, we next focused our analysis on low abundant cell populations comprising less than 10% of the total cell repertoire. Using the same approach, we revealed an overall high agreement ($r = 0.932$, $p\text{-value} = 4.6 \times 10^{-24}$) between these platforms despite decreasing cell population abundance, further enhancing the data alignment between the two techniques (Fig. 1C).

In addition to the quantitative comparison, Uniform Manifold Approximation and Projection (UMAP) was used for dimensional reduction and visualisation. Uniform Manifold Approximation and Projection plots of CyTOF (Fig. 1D) and FSFC (Fig. 1E) datasets show a remarkably similar distribution of cell subsets although not localized at the exact same localization. Some cluster populations were more delineated

when looking at the FSFC dataset such as the CD16⁻, CD57-NK cells and the CD4⁻, CD8⁻ NKT subset than in the CyTOF dataset (Fig.1D, 1E).

Comparison of cell population assignment between manual and automated approaches based on the FSFC dataset

There is an increasing tendency to shift towards more sophisticated automated or semi-automated methods to analyze and interpret high-dimensional datasets that comes at a cost of data accuracy. Herein, we assessed the concordance of cell population assignment and quantification between the manual gating and semi-automated analysis for the newly emerged spectral flow cytometry technique. We manually gated the cells to assign them to the 15 clusters of interest defined previously and imported the dataset in R environment for downstream analysis. The agreement between both approaches was highly significant, reaching a correlation of $r = 0.998$ ($p\text{-value} = 5.88 \times 10^{-77}$) for all cell populations (Fig. 2A) and $r = 0.969$ ($p\text{-value} = 9.17 \times 10^{-33}$) for low frequent cell populations (<10%) (Fig. 2B). In line, UMAP plots of semi-automated and manual gated cell populations show a highly similar projection in terms of distance and localization (Fig. 2C, 2D).

We further thought to separately investigate the overlap between the different cell populations obtained with semi-automated and with manual gating by applying the Matthew's correlation coefficient (MCC). Analysis across all samples revealed accurate assignment for highly abundant populations such as CD4⁺, CD8⁺, CD19⁺ and CD14⁺ cell populations with a median Matthew's Correlation Coefficient above 0.9. Suboptimal scorings were found for small populations such as NK cell subsets and more drastically in NKT populations with the lowest correlation coefficient reaching 0.45 (Fig. 2E). We thus searched to quantify and investigate the origin of the non-concordant assignment of these populations based on a contingency table (Fig. S4). Enhanced leakages of cells were found between CD4⁺/CD8⁺ T cells and CD4⁺/CD8⁺ NKT subsets, CD4⁻,CD8⁻ T cells and CD4⁻,CD8⁻ NKT cells and in-between CD56/CD16/CD57 combinatorial NK subsets. This reflects the non-dichotomous staining pattern of these subsets leading to false cluster allocation with the potential to impact the downstream statistical analysis and biological interpretation (Fig. S4).

4. DISCUSSION

The recently emerging mass cytometry and full spectrum flow cytometry techniques have greatly enhanced our capacity to probe biological systems at a high resolution and scalability with their capacity to measure several parameters simultaneously in one assay [73, 256, 257, 259-261]. In clinical studies, samples are often derived from precious, limited and non-replaceable sources, frequently yielding low cell numbers. Thus, a thoughtful examination of the appropriate analysis tool needs to be undertaken to maximize data output. In this proof-of-concept study, we set out a biological comparative analysis of the two state-of-the art single cell proteomic technologies, cytometry by time-of-flight and full spectrum flow cytometry using an identical panel of 33 markers. Our results first revealed a high agreement between the two datasets in the quantification of major canonical immune cell subsets using a semi-automated computational approach and which is in concordance with recent findings [262-266]. Rare subpopulations showed minor disagreements between both techniques that can be attributed to their low frequency abundance and the acquired cell number.

The increasing number of bioinformatics tools emerging alongside the increasing output of biological data substantiate the need for validation of cell phenotypes and accuracy in assignment. While distinct populations with a clear separation of negative and positive such as CD8⁺ T cells and CD19⁺ B cells could be accurately captured by both semi-automated analysis and manual gating in our analysis, populations that were more subject to spreading such as NKT cells showed major discrepancies in the assignment. Thus, despite their explorative and multi-dimensional analysis ability, clustering tools (FlowSOM) bear the potential to over-estimate or under-estimate cell populations even with supposedly basic markers such as CD56 and CD16 and need to be applied with cautions for non-bimodal markers.

The choice of technology is highly dependent on the study design and not primarily on the biological output, as shown to be highly overlapping in our analysis. Indeed, each of these technologies has its strengths and weaknesses regarding their technical performance. With a theoretical acquisition of 100 parameters simultaneously and the minimal spillover inherent to the use of metal isotopes, mass cytometry represents a great tool to investigate the breadth of cellular systems without autofluorescence or the need to compensate spectral overlap [267, 268]. The minimal overlap between metal isotopes offers furthermore a higher flexibility in the translation and customization of panels than for spectral flow cytometry and which can be of great benefit in various clinical settings. However substantial constraints exist with the mass cytometry technique: slow acquisition speed, low cell throughput and it remains

considerably expensive [269, 270]. Full spectral flow cytometry is a direct response to these drawbacks, enabling acquisition of a large amount of data in reduced runtimes and with a considerably higher cell throughput [260, 271]. Full spectral flow cytometry provides the greatest benefit in terms of time to cost efficiency which is valuable in clinical studies encompassing a large amount of samples. It allows simultaneously to rapidly extract as much data as possible, essential when detecting and monitoring the frequency of rare subpopulations. Finally, full-spectrum flow cytometry now offers the possibility to sort sub-populations of cells for subsequent and future applications. We summarized key considerations and questions in table 1 that need to be consciously interrogated when setting out a new clinical study. In conclusion, both tools can interchangeably be used to capture single cell data from a biological perspective as set out by our comparative analysis. Both modalities are valuable and complementary tools on the market for addressing research questions in immunophenotyping studies.

Authors' contributions

A.S., S.D.: performed CyTOF and FSFC analysis and measurements. A.S., S.D., J.D.: performed the bioinformatic analysis. A.S.: drafted and wrote the manuscript. All authors revised and approved the final version for publication.

Declaration of competing interest

All authors declare that there are no competing interests.

Acknowledgements

We thank Juliette Desfrancois from Cytex and Emilie Gregori from Fluidigm for her support with the design of the panels.

Data availability

Data will be made available on request.

Table 1 Comparative assessment of technical, logistical and financial aspects between CyTOF and FSFC platforms [238, 243, 269, 272, 273].

Theme	Parameter	CyTOF	FSFC	Study
Technical	Number of parameters	43	40	
	Throughput (cells/s)	500	10-15'000	How many samples are included? What is the timeframe of the study?
	<i>i.e., 60×10^6</i>	<i>~ 30 hours</i>	<i>~ 1, 5 hours</i>	Are rare subsets a special focus of the study?
	Cell transmission efficiency	30-60 %	> 95 %	What kind of samples are included? According to the study design is there a possibility of barcoding?
	Cell size/ complexity	No	Yes	
	Autofluorescence	No	Yes	Will there be several sample batches of acquisition? Pre-knowledge about the staining pattern of antibodies?
	Compensation	No	Yes	
	Sensitivity	300-400	< 40	
	Cell sorting	No	Yes	
	Panel flexibility	High	Medium	
Logistical	Panel extension	High	Limited	Are there specific clones that need to be considered?
	Antibody availability	Single vendor-in house conjugation	Many vendors-company customization	
	Training investment	High	Low	
	Barcoding	Yes	No	Is there a need of the samples to be freshly stained and acquired?
Financial	Antibody costs (100 tests, CHF)	550	300-600	
	Infrastructure and maintenance costs	2 32A plugs + Argon supply + Air extraction	MilliQ water + 1 12A plug	What is the overall financial budget for the study?
	Data acquisition costs ($1e^6$ cells)	15 CHF	1-2 CHF	

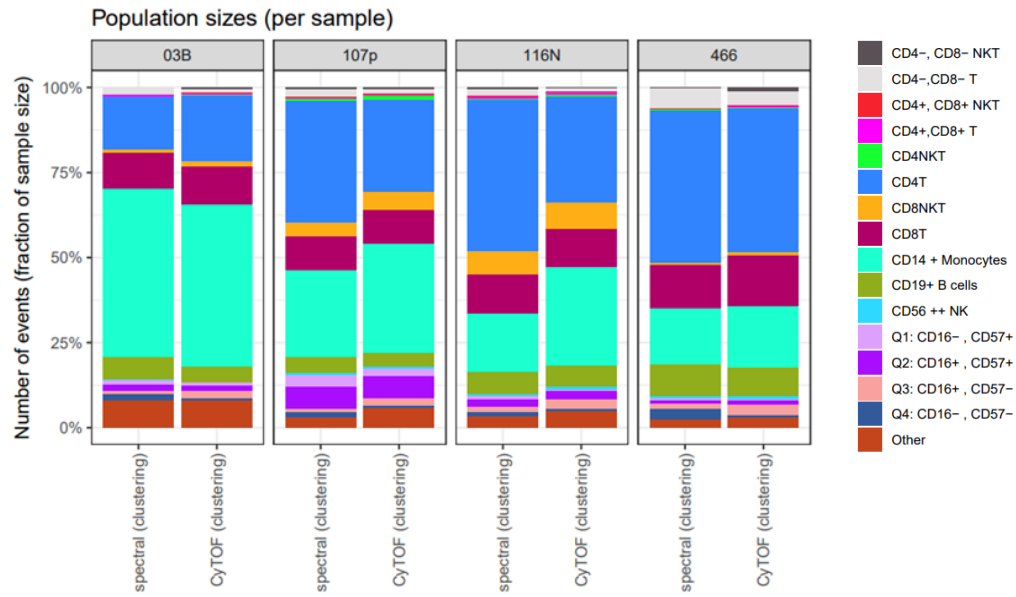
CyTOF: Cytometry by time-of-flight, FSFC: Full spectrum flow cytometry

MAIN PART FIGURE LEGENDS

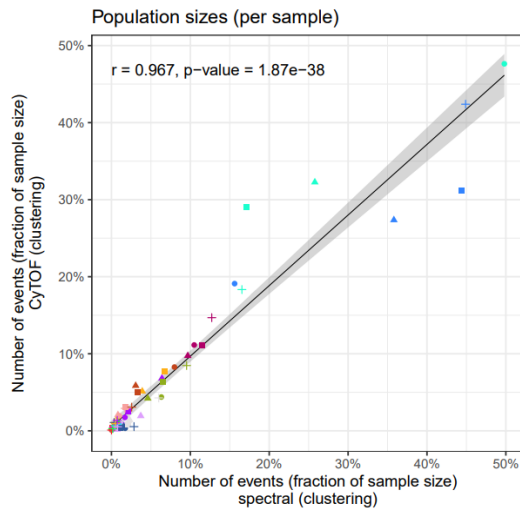
Figure 1. Comparison between CyTOF and FSFC dataset. **A.** Abundances of 15 cell populations identified with semi-automated analysis for the FCFS and CyTOF datasets across individual samples. **B.** Comparison of cell population (color) abundances obtained with semi-automated analysis for the CyTOF (y-axis) and FCFS (x-axis) data sets, per sample (symbol). Linear regression line is shown in black with the 95% confidence interval (grey). **C.** Comparison of cell population (color) abundances limited to populations with abundance below 10% obtained with semi-automated analysis for the CyTOF (y-axis) and FCFS (x-axis) data sets, per sample (symbol). **D.** UMAP for the CyTOF data set, color-coded by cell population obtained with semi-automated analysis. Events are color - coded according to the 15 cell populations obtained with FlowSOM clustering. **E.** Same as for D for the FSFC dataset.

Figure 2. Comparison between semi-supervised and supervised analysis in FSFC dataset. **A.** Comparison of cell population (color) abundances obtained with semi-automated analysis (y-axis) and manual gating (x-axis) for the FCFS data set, per sample (symbol). Correlation plots representative of the quantitative comparison (cluster frequencies) between FSFC and CyTOF for all 15 clusters identified by manual gating and FlowSOM clustering. Linear regression line is shown in black with the 95% confidence interval (grey). **B.** Comparison of cell population (color) abundances limited to populations with abundance below 10% obtained with semi-automated analysis (y-axis) and manual gating (x-axis) for the FCFS data set, per sample (symbol). **C.** UMAP for the FSFC data set. Events are color – coded according to the 15 cell populations obtained with manual gating. **D.** UMAP for the FSFC data set, color-coded by cell population obtained with semi-automated analysis. Events are color - coded according to the 15 cell populations obtained with FlowSOM clustering. **E.** Comparison of cell populations obtained with manual gating and semi-automated analysis using Matthews Correlation Coefficient (MCC).

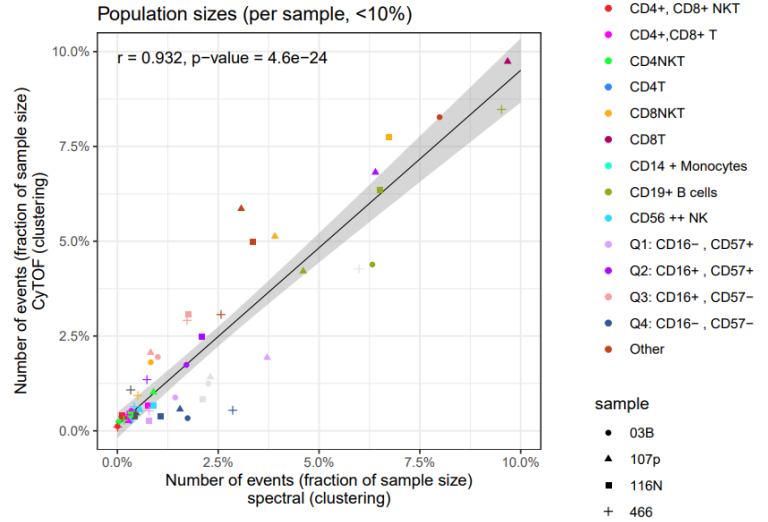
A



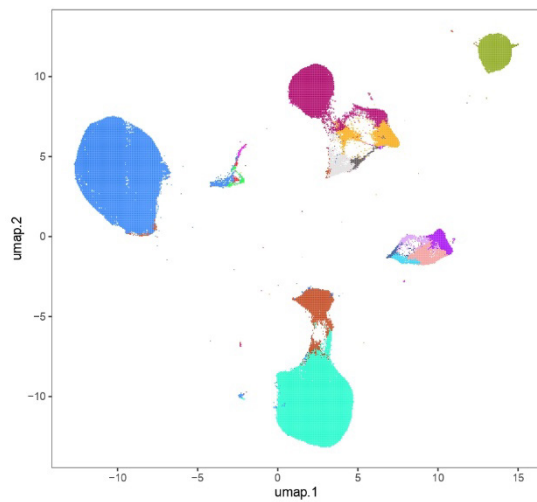
B



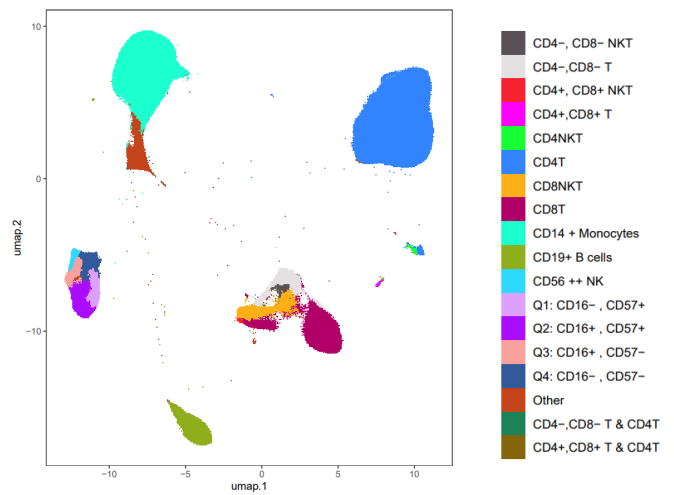
C

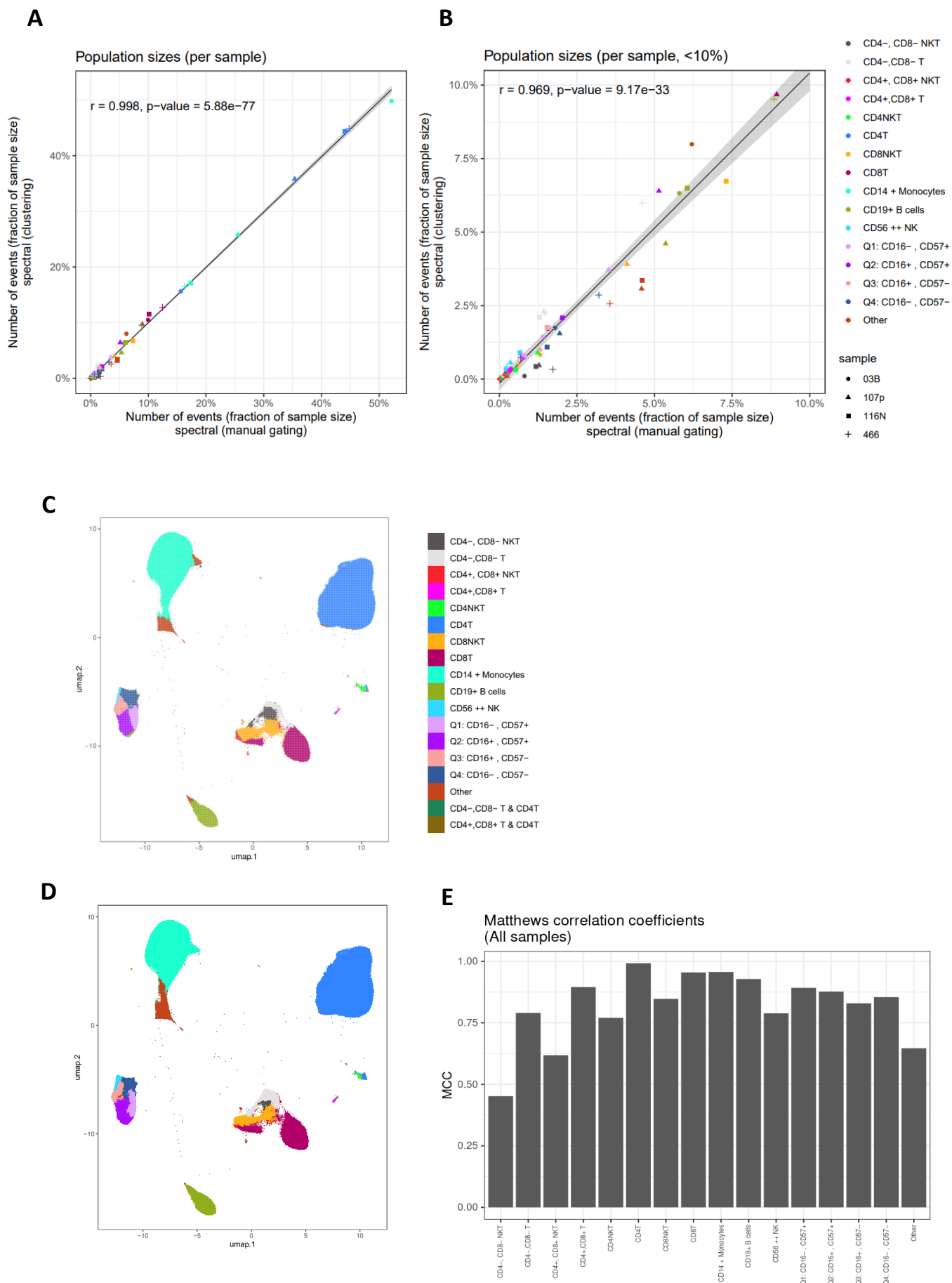


D



E





Supplementary material

Comparative assessment of cytometry by time-of-flight and full spectral flow cytometry based on a 33-color antibody panel

¹Antonia Schäfer, ², ^{*}Sènan Mickael D'Almeida, ^{3,4}Julien Dorier, ^{3,4}Nicolas Guex,¹Jean Villard, ²Miguel Garcia

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

²Flow Cytometry Core Facility, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

³Bioinformatics Competence Center, University of Lausanne, Lausanne, Switzerland

⁴Bioinformatics Competence Center, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

* Present address: Viollier AG, Allschwil, Switzerland

Supplementary table S1. List of antibodies used for the mass cytometry panel.

Target	Metal isotope	Antibody clone	Provider	ug /test	Staining step
CD45	089Y	HI30	Fluidigm	1	Surface
CD3	141Pr	UCHT1	Fluidigm	0.25	Surface
CD4	145Nd	RPA-T4	Fluidigm	1	Surface
CD8a	168Er	SK1	Fluidigm	1	Surface
CD19	142Nd	HIB19	Fluidigm	0.5	Surface
CD14	151Eu	M5E2	Fluidigm	1	Surface
CD56	149Sm	NCAM16.2	Fluidigm	0.25	Surface
CD16	209Bi	3G8	Fluidigm	0.25	Surface
CD57	176Yb	HCD57	Fluidigm	1	Surface
NKG2A	169Tm	Z199	Fluidigm	0.5	Surface
NKG2C	158Gd	REA205	Miltenyi	0.58	Surface
NKG2D	166Er	ON72	Fluidigm	1	Surface
NKp30	159Tb	Z25	Fluidigm	1	Surface
NKp44	150Nd	REA1163	Miltenyi	0.5	Surface
NKp46	162Dy	BAB281	Fluidigm	0.5	Surface
KIR2DL1	165Ho	HP-3E4	BD biosciences	0.76	Surface
KIR2DL1/S1	170Er	REA1010	Miltenyi	1.25	Surface
KIR2DL2/L3 /S2	173Yb	DX27	Fluidigm	0.25	Surface
KIR2DL3	153Eu	REA147	Miltenyi	1.2	Surface
KIR2DL4	147Sm	REA 768	Miltenyi	0.9	Surface
KIR2DL5	156Gd	REA955	Miltenyi	1.3	Surface
KIR2DS4	160Gd	REA860	Miltenyi	0.9	Surface
KIR3DL1	167Er	REA1005	Miltenyi	1.4	Surface
KIR3DL2	155Gd	539304	R&D Systems	1	Surface
KIR3DL3	144Nd	1136B	R&D Systems	1	Surface
CD107a /APC	163Dy	APC003	Fluidigm	1	Intra – cellular
TNFa	152Sm	Mab11	Fluidigm	1	Intra – cellular
PD-1	174Yb	EH12.2H7	Fluidigm	1	Surface
TIGIT	154Sm	MBSA43	Fluidigm	1	Surface
TRAIL	161Dy	REA1113	Miltenyi	1.2	Surface
Fas-L	164Dy	REA1056	Miltenyi	1.4	Surface
Granzyme B	171Yb	GB11	Fluidigm	1	Intra – cellular
Perforin	175Lu	B-D48	Fluidigm	1	Intra – cellular

Supplementary table S2. List of antibodies used for the full spectrum flow cytometry panel.

Target	Fluorochrome	Antibody clone	Provider	ug/test	Staining step
CD45	cFluor V547	HI30	Cytek custom	0.2	Surface
CD3	APC/Fire 810	SK7	Biolegend	0.1	Surface
CD4	PE/Fire 640	SK3	Biolegend	0.1	Surface
CD8a	BV570	RPA-T8	Biolegend	0.4	Surface
CD19	BV750	H1B19	Biolegend	0.4	Surface
CD14	BUV805	M5E2	BD biosciences	0.2	Surface
CD56	BV785	5.1H11	Biolegend	0.05	Surface
CD16	BUV496	3G8	BD biosciences	0.05	Surface
CD57	BV605	QA17A04	Biolegend	0.4	Surface
NKG2A	PE-Cy5	S19004C	Biolegend	0.4	Surface
NKG2C	BUV737	134591	BD biosciences	0.2	Surface
NKG2D	cFluor BYG710	1D11	Cytek custom	2.5	Surface
NKp30	BUV563	P30-15	BD biosciences	1	Surface
NKp44	BUV615	P44-8	BD biosciences	1	Surface
NKp46	BV711	9E2	BD biosciences	1	Surface
KIR2DL1	PE-Vio 770	REA284	Miltenyi	1	Surface
KIR2DL1/ S1	APC-Vio 770	REA1010	Miltenyi	1	Surface
KIR2DL2/ L3 /S2	BB700	CH-L	BD biosciences	0.4	Surface
KIR2DL3	APC	REA147	Miltenyi	2.5	Surface
KIR2DL4	PE	181703	R&D Systems	2.5	Surface
KIR2DL5	BV421	UPR1	BD biosciences	1	Surface
KIR2DS4	PE – Vio 615	REA860	Miltenyi	2.5	Surface
KIR3DL1	VioGreen	REA1005	Miltenyi	1	Surface
KIR3DL2	Alexa Fluor 647	539304	R&D Systems	2.5	Surface
KIR3DL3	Alexa Fluor 700	1136B	R&D Systems	2.5	Surface
CD107a	BUV395	H4A3	BD biosciences	0.05	Intra – cellular
TNFa	BV650	Mab11	Biolegend	1	Intra – cellular
PD-1	BUV661	EH12.1	BD biosciences	0.4	Surface
TIGIT	cFluor BYG750	A15153G	Cytek custom	1	Surface
TRAIL	VioBright B515	REA1113	Miltenyi	2.5	Surface
Fas-L	BV480	NOK1	BD biosciences	1	Surface
Granzyme B	Pacific Blue	GB11	Biolegend	0.2	Intra – cellular
Perforin	cFluor R685	DG9	Cytek custom	1	Intra – cellular
KIR3DL1 / S1	FITC	REA168	Miltenyi	1	Surface
Viability	ViaDye Red		Cytek		

Supplementary figure S1.

Similarity index matrix of fluorochrom pairs: The matrix shows the index of similarity between fluorochromes that assesses the degree of overlap between fluorochrome spectra. Indexes of fluorochrome pairs vary between 0 – being the state of no similarity – to 1 indicating a complete overlapping spectrum. The overall complexity index of our panel was 43.73.

Similarity™ Indices

Configuration: 5L 16UV-16V-14B-10YG-8R

	BUV395	BUV496	BUV563	BUV615	BUV661	BUV737	BUV805	BV421	BV480	BV570	BV605	BV650	BV711	BV750	BV785	Pacific Blue	VioGreen	cFluor V547	VioB515	FITC	BB700	PE	PE-Vio 615	PE-Fire 640	PE-Cy5	cFluor BYG710	cFluor BYG750	PE-Vio770	APC	Alexa Fluor 647	cFluor R685	Alexa Fluor 700	ViaDye Red	APC-Vio 770	APC-Fire 810			
BUV395	1	0.25	0.05	0.03	0.04	0.03	0.11	0.05	0.02	0.01	0.01	0	0	0	0	0.01	0.01	0.01	0.01	0.01	0	0.01	0.01	0	0	0	0	0	0	0	0	0	0	0				
BUV496	0.25	1	0.29	0.08	0.02	0.01	0.03	0.09	0.38	0.09	0.05	0.02	0	0	0	0.13	0.29	0.22	0.07	0.09	0	0.03	0.01	0	0	0	0	0	0	0	0.01	0.01	0	0				
BUV563	0.05	0.29	1	0.4	0.06	0.01	0.01	0.01	0.07	0.37	0.2	0.04	0.01	0	0	0.01	0.14	0.14	0.04	0.06	0.02	0.51	0.24	0.08	0.07	0.1	0.09	0.01	0.03	0.01	0.01	0	0	0				
BUV615	0.03	0.08	0.4	1	0.35	0.1	0.04	0	0.04	0.29	0.51	0.26	0.07	0.03	0.02	0	0.11	0.15	0	0.01	0.08	0.25	0.58	0.36	0.32	0.2	0.13	0.04	0.17	0.05	0.04	0.04	0.03	0.02	0			
BUV661	0.04	0.02	0.06	0.35	1	0.37	0.1	0	0.01	0.06	0.17	0.4	0.25	0.09	0.05	0	0.02	0.04	0	0	0.33	0.03	0.21	0.38	0.42	0.22	0.12	0.05	0.78	0.71	0.64	0.42	0.35	0.16	0.1			
BUV737	0.03	0.01	0.01	0.1	0.37	1	0.39	0	0	0.02	0.05	0.13	0.41	0.37	0.25	0	0.01	0.01	0	0	0.3	0	0.04	0.1	0.1	0.16	0.22	0.14	0.24	0.22	0.28	0.51	0.5	0.33	0.21			
BUV805	0.11	0.03	0.01	0.04	0.1	0.39	1	0.01	0	0	0.01	0.02	0.09	0.14	0.21	0	0	0	0	0	0.05	0	0	0.01	0.01	0.02	0.04	0.07	0.02	0.02	0.03	0.05	0.11	0.19	0.24			
BV421	0.05	0.09	0.01	0	0	0	0.01	1	0.28	0.15	0.06	0.09	0.09	0.05	0.07	0.78	0.11	0.07	0	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
BV480	0.02	0.38	0.07	0.04	0.01	0	0	0.28	1	0.26	0.16	0.06	0.03	0.02	0.02	0.58	0.74	0.59	0.06	0.1	0.02	0.05	0.02	0	0	0.01	0.01	0	0	0	0	0.01	0	0	0			
BV570	0.01	0.09	0.37	0.29	0.06	0.02	0	0.15	0.26	1	0.7	0.25	0.07	0.04	0.03	0.15	0.62	0.75	0.01	0.03	0.08	0.52	0.29	0.11	0.09	0.1	0.09	0.01	0.05	0.01	0.01	0	0.01	0.01	0			
BV605	0.01	0.05	0.2	0.51	0.17	0.05	0.01	0.06	0.16	0.7	1	0.53	0.17	0.1	0.07	0.06	0.45	0.59	0	0.01	0.16	0.28	0.45	0.25	0.21	0.14	0.1	0.03	0.14	0.03	0.03	0.03	0.02	0.02	0.01			
BV650	0	0.02	0.04	0.26	0.4	0.13	0.02	0.09	0.06	0.25	0.53	1	0.45	0.23	0.15	0.08	0.16	0.25	0	0	0.37	0.05	0.25	0.31	0.29	0.15	0.09	0.04	0.37	0.18	0.19	0.16	0.09	0.06	0.05			
BV711	0	0	0.01	0.07	0.25	0.41	0.09	0.09	0.03	0.07	0.17	0.45	1	0.69	0.49	0.07	0.05	0.07	0	0	0.53	0.01	0.08	0.14	0.14	0.21	0.18	0.11	0.23	0.18	0.26	0.44	0.26	0.19	0.12			
BV750	0	0	0	0.03	0.09	0.37	0.14	0.05	0.02	0.04	0.1	0.23	0.69	1	0.82	0.04	0.03	0.04	0	0	0.3	0	0.03	0.05	0.05	0.08	0.16	0.14	0.07	0.02	0.05	0.17	0.09	0.19	0.12			
BV785	0	0	0	0.02	0.05	0.25	0.21	0.07	0.02	0.03	0.07	0.15	0.49	0.82	1	0.06	0.02	0.03	0	0	0.21	0	0.02	0.03	0.03	0.05	0.12	0.17	0.04	0.01	0.03	0.1	0.07	0.22	0.18			
Pacific Blue	0.01	0.13	0.01	0	0	0	0	0.78	0.58	0.15	0.06	0.08	0.07	0.04	0.06	1	0.26	0.17	0	0.02	0	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0			
VioGreen	0.01	0.29	0.14	0.11	0.02	0.01	0	0.11	0.74	0.62	0.45	0.16	0.05	0.03	0.02	0.26	1	0.96	0.01	0.06	0.05	0.14	0.07	0.02	0.01	0.02	0.02	0	0.02	0	0	0.01	0	0	0			
cFluor V547	0.01	0.22	0.14	0.15	0.04	0.01	0	0.07	0.59	0.75	0.59	0.25	0.07	0.04	0.03	0.17	0.96	1	0.01	0.05	0.08	0.16	0.1	0.03	0.02	0.02	0.02	0	0.03	0	0.01	0.01	0	0	0			
VioB515	0.01	0.07	0.04	0	0	0	0	0	0.06	0.01	0	0	0	0	0	0	0.01	0.01	1	0.97	0.01	0.05	0.02	0.01	0.01	0.01	0.01	0	0	0	0	0	0	0	0			
FITC	0.01	0.09	0.06	0.01	0	0	0	0.01	0.1	0.03	0.01	0	0	0	0	0.02	0.06	0.05	0.97	1	0.02	0.09	0.04	0.01	0.01	0.01	0.01	0	0	0	0	0	0	0	0			
BB700	0	0	0.02	0.08	0.33	0.3	0.05	0	0.02	0.08	0.16	0.37	0.53	0.3	0.21	0	0.05	0.08	0.01	0.02	1	0.06	0.31	0.42	0.48	0.47	0.28	0.17	0.33	0.33	0.41	0.38	0.3	0.16	0.1			
PE	0.01	0.03	0.51	0.25	0.03	0	0	0	0.05	0.52	0.28	0.05	0.01	0	0	0.01	0.14	0.16	0.05	0.09	0.06	1	0.43	0.15	0.14	0.18	0.17	0.02	0.04	0.01	0.01	0	0.01	0	0			
PE-Vio 615	0.01	0.01	0.24	0.58	0.21	0.04	0	0	0.02	0.29	0.45	0.25	0.08	0.03	0.02	0	0.07	0.1	0.02	0.04	0.31	0.43	1	0.75	0.68	0.44	0.25	0.09	0.25	0.09	0.06	0.06	0.03	0.02				
PE-Fire 640	0	0	0.08	0.36	0.38	0.1	0.01	0	0	0.11	0.25	0.31	0.14	0.05	0.03	0	0.02	0.03	0.01	0.01	0.42	0.15	0.75	1	0.94	0.6	0.35	0.15	0.48	0.25	0.2	0.18	0.17	0.08	0.05			
PE-Cy5	0	0	0.07	0.32	0.42	0.1	0.01	0	0	0.09	0.21	0.29	0.14	0.05	0.03	0	0.01	0.02	0.01	0.01	0.48	0.14	0.68	0.94	1	0.61	0.34	0.14	0.52	0.32	0.27	0.2	0.19	0.08	0.06			
cFluor BYG710	0	0	0.1	0.2	0.22	0.16	0.02	0	0.01	0.1	0.14	0.15	0.21	0.08	0.05	0	0.02	0.02	0.01	0.01	0.47	0.18	0.44	0.6	0.61	1	0.65	0.26	0.24	0.15	0.19	0.28	0.23	0.1	0.06			
cFluor BYG750	0	0	0.09	0.13	0.12	0.22	0.04	0	0.01	0.09	0.1	0.09	0.18	0.16	0.12	0	0.02	0.02	0.01	0.01	0.28	0.17	0.25	0.35	0.34	0.65	1	0.72	0.14	0.08	0.1	0.21	0.28	0.23	0.13			
PE-Vio770	0	0	0.01	0.04	0.05	0.14	0.07	0	0	0.01	0.03	0.04	0.11	0.14	0.17	0	0	0	0	0	0.17	0.02	0.09	0.15	0.14	0.26	0.72	1	0.05	0.03	0.04	0.09	0.21	0.3	0.2			
APC	0	0	0.03	0.17	0.78	0.24	0.02	0	0	0.05	0.14	0.37	0.23	0.07	0.04	0	0.02	0.03	0	0	0.33	0.04	0.25	0.48	0.52	0.24	0.14	0.05	1	0.9	0.75	0.49	0.44	0.2	0.13			
Alexa Fluor 647	0	0	0.01	0.05	0.71	0.22	0.02	0	0	0.01	0.03	0.18	0.18	0.02	0.01	0	0	0	0	0	0.33	0.01	0.09	0.25	0.32	0.15	0.08	0.03	0.9	1	0.92	0.53	0.47	0.2	0.12			
cFluor R685	0	0	0.01	0.04	0.64	0.28	0.03	0	0	0.01	0.03	0.19	0.26	0.05	0.03	0	0	0.01	0	0	0.41	0.01	0.06	0.2	0.27	0.19	0.1	0.04	0.75	0.92	1	0.66	0.58	0.27	0.16			
Alexa Fluor 700	0	0.01	0	0.04	0.42	0.51	0.05	0	0.01	0	0.03	0.16	0.44	0.17	0.1	0	0.01	0.01	0	0	0.38	0	0.06	0.18	0.2	0.28	0.21	0.09	0.49	0.53	0.66	1	0.81	0.37	0.22			
ViaDye Red	0	0.01	0.01	0.03	0.35	0.5	0.11	0	0	0.01	0.02	0.09	0.26	0.09	0.07	0	0	0	0	0	0.3	0.01	0.06	0.17	0.19	0.23	0.28	0.21	0.44	0.47	0.58	0.81	1	0.79	0.48			
APC-Vio 770	0	0	0	0.02	0.16	0.33	0.19	0	0	0.01	0.02	0.06	0.19	0.19	0.22	0	0	0	0	0	0.16	0	0.03	0.08	0.08	0.1	0.23	0.3	0.2	0.2	0.27	0.37	0.79	1	0.68			
APC-Fire 810	0	0	0	0.01	0.1	0.21	0.24	0	0	0	0.01	0.05	0.12	0.12	0.18	0	0	0	0	0	0.1	0	0.02	0.05	0.06	0.06	0.13	0.2	0.13	0.12	0.16	0.22	0.48	0.68	1			
Complexity™ Index: 43.73																																						

Complexity™ Index: 43.73

SSM	APC-Fire 810	BUV395	BUV496	BUV563	BUV615	BUV737	BUV805	BUV421	Pacific Blue	BUV480	VoGreen	cFluor V547	BUV570	BUV605	BUV650	BUV711	BUV750	BUV785	VoBright 515	FTTC	BB700	PE	PE-V6615	PE-Fire 640	PE-Cy5	cFluor BUV710	PE-V6770	cFluor BUV781	APC	Alexa Fluor 647	cFluor R685	Alexa Fluor 700	APC-V6770	SUM	
APC-Fire 810	0.42	0.54	0.40	0.27	0.29	0.30	1.96	0	0.36	0	1.47	2.91	1.02	0.89	0.59	0.44	0.64	1.93	0	0	0.54	0.68	0	1.11	1.09	0.66	0.65	0	0.65	2.62	1.39	0.59	2.57	27.01	
BUV395	0	1.51	0.51	0.33	0	0	0	0.49	0.34	0.48	6.49	5.88	1.00	0.78	0	0	0	0.19	0	0	0	0.71	0	0	0	0	0.47	0	0	0	0	0	0	0	19.17
BUV496	0.46	0.63	1.67	0.87	0	0.24	0	0.41	0	0	10.28	9.58	1.74	1.42	0.10	0.30	0.44	0.14	3.69	2.85	0	1.12	0.44	0.34	0.48	0.17	0	0.38	0	0.78	0.44	0.26	0.46	39.68	
BUV563	0	2.27	3.61	1.57	0.73	0.38	0.38	0.75	0.44	0.96	12.10	12.13	3.06	1.58	0.58	0	0.33	0.34	5.84	6.40	0.46	3.99	1.74	1.48	0.67	0.61	0.21	0.28	1.13	1.70	0.79	0.27	0	66.80	
BUV615	0.34	2.52	1.73	2.01	2.00	1.15	1.14	0.92	0.65	1.69	17.17	17.37	3.75	4.22	1.08	0.92	0.73	0.48	0.62	0.25	1.18	1.47	4.13	4.59	2.61	1.80	0.78	1.11	3.78	4.16	2.45	0.95	0.36	90.10	
BUV661	0.79	5.96	3.69	0.55	1.51	2.33	1.84	0.99	0.76	1.50	13.83	12.82	2.32	1.37	2.95	1.35	1.12	0.85	0	0.25	2.73	0.70	1.12	4.13	4.73	2.23	0.84	1.37	17.58	6.90	1.88	1.48	107.55		
BUV737	1.26	2.58	1.52	0.27	0.47	0.80	3.92	0.49	0.44	0.60	4.59	4.38	0.76	0.65	0.78	1.32	2.50	1.73	0.78	0.44	2.61	0.42	0.47	0.85	1.06	1.70	0.96	1.74	1.25	2.26	2.40	6.77	1.91	54.71	
BUV805	1.68	0.70	0.74	0.43	0	0.36	1.04	0.25	0.45	0.28	1.95	2.12	0	0.33	0	0.42	0.25	1.24	0.38	0	0.31	0	0.38	0	0.50	0	0.33	0.33	0.84	1.30	0.69	0.50	1.36	19.16	
BUV421	0	0.80	0.93	0.64	0.31	0.13	0	0.27	4.36	2.24	9.12	8.00	1.34	1.05	0.16	0.24	0.21	0	0.24	0.40	0	0.51	0.21	0.46	0	0	0	0	0.20	0.41	0.26	0	0	32.50	
Pacific Blue	0.24	0	0.73	0.57	0.31	0.31	0	0.11	0.49	2.00	9.02	7.40	1.67	0.77	0.20	0	0	0	0	0	0	0.55	0	0	0	0	0.47	0.42	0.17	0	0	0.35	0.13	0.13	25.92
BUV480	0	0.73	4.01	1.49	0.80	0.26	0.22	0	3.34	30.81	21.91	18.49	3.13	2.58	0.77	0.37	0.31	0.23	3.32	2.35	0.04	1.49	0.45	0.36	0.25	0.18	0	0	0.49	0.28	0.22	0.16	0.19	99.25	
VoGreen	0	0.20	2.27	2.05	0.86	0.41	0.20	0.14	0.65	0.73	1.13	11.53	3.37	2.53	1.17	0.49	0.85	0.69	1.08	0.65	0.07	1.65	0.49	0	0	0	0.40	0	0	0.45	0	0	0	0	34.05
cFluor V547	0.30	0.14	0.86	1.48	0.75	0.35	0.37	0.36	0.61	0	0.62	7.35	3.87	2.72	1.33	0.91	0.91	0.63	1.73	0.64	0	1.55	0	0.32	0.46	0.22	0	0.68	0.88	1.22	0.82	0.59	0.29	32.96	
BUV570	0.18	0.32	0.96	1.56	1.02	0.60	0.39	0.31	2.86	1.81	14.78	15.32	3.36	1.36	1.36	0.72	0.75	0.58	1.52	1.59	0.55	4.18	1.15	1.96	1.19	0.91	0	0.27	1.35	1.77	1.00	0.40	0.22	66.90	
BUV605	0.15	0.07	0.80	0.99	2.06	0.97	0.58	0.52	1.06	0.65	10.14	10.22	1.92	1.96	1.18	1.26	1.03	0.63	0.63	0.54	0.88	0.80	1.45	2.39	1.80	1.34	0.41	0.55	1.87	2.45	1.41	0.55	0.24	53.70	
BUV650	0.31	0.25	0.40	0.17	0.38	1.77	0.92	0.75	1.61	0.95	5.39	4.38	0.34	1.31	1.31	1.56	1.71	1.70	0.81	0.83	0.81	0.30	0.63	2.45	1.48	1.14	0.31	0.70	2.32	3.63	1.92	0.88	0.42	43.30	
BUV711	0.77	0	0.57	0	0.35	0.44	1.44	0.98	0	0.37	0	1.18	0.74	0.56	0.61	2.68	2.59	4.08	12.35	0	0	0.85	0.31	0	0.95	0.63	1.56	0.61	0.62	1.16	2.79	2.38	1.37	1.26	27.78
BUV750	0.79	0	0	0.31	0.35	0.34	2.11	1.59	0.69	0.53	0.39	2.72	2.40	0.67	0.11	0.24	1.38	0	4.08	0.10	0.10	0.44	0.28	0	0	0.43	0.56	0.42	0.55	0.70	0.87	0.77	0.98	1.09	26.01
BUV785	0.92	0.21	0	0	0.17	0	1.01	2.53	0.47	0.46	0.20	1.54	0.33	0.41	0.17	0.39	2.01	0	0	0.43	0.17	0.20	0	0.40	0.06	0.25	0.55	0.42	0.55	0.70	0.87	0.77	0.98	1.09	26.01
VoBright 515	0.13	0.08	0.69	0.62	0.19	0.13	0.03	0	0.39	0.22	0.53	6.01	1.86	0.68	0.24	0.04	0.20	0.13	10.62	10.62	0.29	1.39	0.73	0.28	0.30	0.20	0	0	0.71	0.76	0.42	0.10	0	39.43	
FTTC	0	0	0.77	0.59	0.29	0.34	0.19	0	0.31	0	0.59	7.78	2.27	0.93	0.23	0	0	0.06	12.35	4.39	0.29	2.20	2.32	2.70	2.65	3.38	1.41	1.57	5.61	8.21	9.35	2.01	1.11	132.38	
BB700	0.78	0.26	1.18	1.74	1.04	1.35	1.53	0.85	0.85	0.65	2.06	23.48	24.13	4.05	1.97	4.43	4.95	3.08	3.08	4.03	4.39	0.50	2.20	2.32	2.70	2.65	3.38	1.41	1.57	5.61	8.21	9.35	2.01	1.11	132.38
PE	0.14	0	0.71	1.52	0.43	0.31	0.15	0	0.56	0	1.15	13.61	3.37	3.44	1.67	0.57	0.27	0.23	0	1.42	1.43	0.50	2.20	2.32	2.70	2.65	3.38	1.41	1.57	5.61	8.21	9.35	2.01	1.11	49.93
PE-V6615	0.29	0	0.58	1.18	1.71	0.84	0.57	0.32	0.53	0.29	1.13	11.48	11.21	2.65	1.91	1.19	1.04	0.67	0.56	1.27	1.04	2.05	2.96	1.07	1.32	3.27	2.66	0.94	1.48	4.46	5.74	3.29	0.93	0.26	73.45
PE-Fire 640	0.57	0	0	0.60	0.67	0.70	0.55	0.44	0	0.19	4.05	2.98	0.77	0.95	1.85	1.34	0.65	0.55	0.85	0.80	2.57	1.07	1.32	3.73	3.73	3.37	1.29	1.74	6.00	7.80	4.25	1.44	0.60	53.84	
PE-Cy5	0.47	0	0.26	0.32	0.51	0.91	0.65	0.43	0.09	0	0.34	4.19	4.11	0.75	0.89	2.12	1.75	0.86	0.71	0.42	0.44	3.82	0.97	1.06	6.40	3.99	3.99	1.42	2.03	6.65	9.69	5.51	1.79	0.80	64.33
cFluor BUV710	0.45	0	0	0.41	0.23	0.37	0.72	0.47	0.23	0	0.42	4.67	4.56	1.25	0.85	0.98	1.25	0.89	0.66	0.55	0.26	3.58	3.05	0.58	1.92	1.90	3.99	1.41	2.03	1.67	2.75	2.07	0.92	0.71	41.80
PE-V6770	1.48	0.10	0	0.13	0.19	0.39	0.68	1.27	0.11	0	0	2.15	1.94	0.55	0.34	0.23	0.68	0.62	2.10	0.51	0.38	1.20	0.85	0.41	1.03	0.91	0.94	2.37	0.77	0.80	0.49	0.30	1.70	25.61	
cFluor BUV781	0.89	0	0	0.43	0.21	0.22	1.03	0.85	0.13	0.21	0.47	4.43	4.02	0.98	0.63	0.48	0.87	1.92	1.40	0.78	0.65	1.54	1.97	0.59	1.57	1.16	2.90	2.69	1.3	1.01	1.54	1.01	0.76	1.14	38.45
APC	0.82	0	0	0.24	0.28	1.00	0.66	0.58	0.21	0.18	0.49	2.99	2.55	0	0.76	3.13	0.98	1.00	0.66	0.12	0	1.34	0.95	0.82	4.40	2.92	1.73	0.80	1.13	8.13	4.61	5.44	1.76	1.25	46.49
Alexa Fluor 647	0.55	0	0.39	0	0.44	0.43	0.22	0.25	0.21	0.65	0	3.19	3.94	0.78	0.60	1.15	0.58	0	0.38	0.91	0.53	0.89	0.66	0	2.09	1.34	1.32	0.58	0.79	3.87	6.00	5.44	1.81	1.27	35.24
cFluor R685	0.65	0	0	0.53	0.17	0.53	0.65	0.40	0	0	0.11	2.36	1.42	0	0	0.88	0.80	0.57	0.58	1.27	0.79	0.82	0	0	1.53	1.42	1.42	0.83	0.94	2.96</					

Supplementary figure S2.

Spillover spreading matrix (SSM) for the 33-color panel. The spillover was calculated based on single-stained controls on beads. The SSM values in each cell reflect the spread of spillover of one fluorochrome into another.

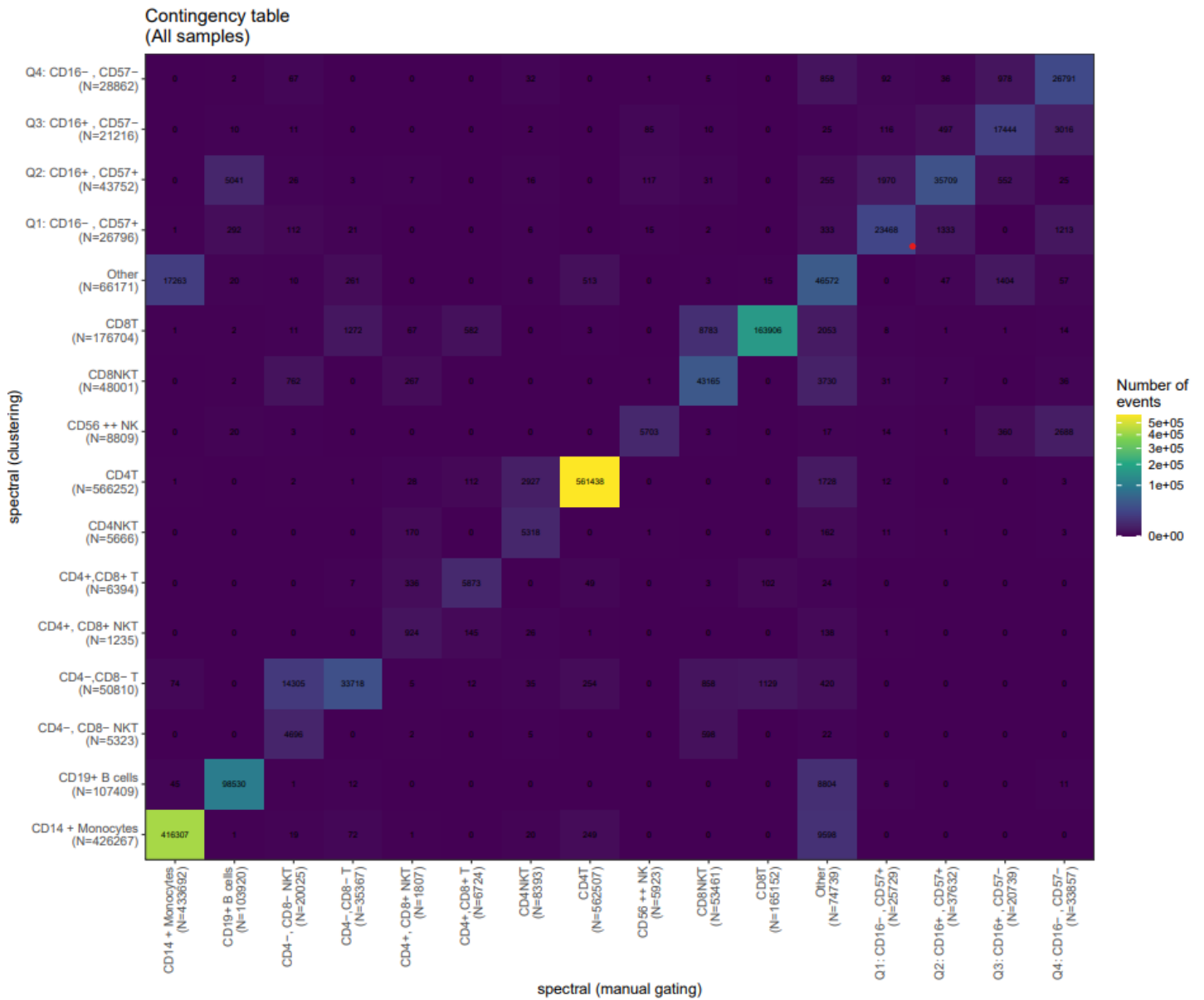
Supplementary figure S3.

Manual gating strategy to identify the 15 immune cell populations throughout the hierarchy.

```
> CD3neg
  >> CD14 + Monocytes
  >> CD56 ++ NK
  >> NK
    >>> Q1: CD16- , CD57+
    >>> Q2: CD16+ , CD57+
    >>> Q3: CD16+ , CD57-
    >>> Q4: CD16- , CD57-
> CD3pos
  >> NKT cells
    >>> CD4+, CD8+ NKT
    >>> CD4-, CD8- NKT
    >>> CD4+ NKT
    >>> CD8+ NKT
  >> Tcells
    >>> CD4+,CD8+ T
    >>> CD4-,CD8- T
    >>> CD4+ T
    >>> CD8+ T
> CD19+ B cells
```

Supplementary figure S4.

Contingency table showing the raw distribution of spectral FlowSOM (y-axis) clustering subsets into the spectral manual gated subsets (x-axis).



Supplementary information

1.1 Panel design

The central key in designing a high-parameter panel relies in the ability to distinguish all levels of antigen expression while maximizing the signal resolution [274].

The antibody-fluorochrome allocation strategy for this study followed general rules: (i) The first rule adopted was to pair lower density antigens, i.e., KIR receptors, with higher sensitivity channels (153-176 Da range) /brighter fluorochromes. (ii) Secondly, fluorochrome pairs with an increased similarity index were associated with markers expressed on different cell types to avoid loss in marker resolution and prioritize the use of bright fluorochromes whose signals are not severely impacted by the spread of other fluorochromes for KIR antigens. (iii) As far as possible, similarity indices of fluorochrome pairs above 0.95 were avoided to secure accurate unmixing results.

To assess the compatibility of fluorochrome combinations based on their spectrum emission signatures, the computation of a similarity index matrix can be interpreted, reflecting the overall complexity of an antibody panel. Indices of fluorochrome pairs vary between 0 – being the state of no similarity – to 1 indicating a complete overlapping spectrum. The overall complexity index of our panel was 43.73, given a range between 42 to 50 for a well-designed 30 to 40 antibody color panel. The first requirement of this panel was the integration of all available antibodies targeting KIRs. Indeed, KIR are highly homologous between each other, leading to the lack of single KIR receptor-targeting antibodies. To overcome this issue, Beziat *et al.* set out a detailed analysis on the identification of KIR cell subsets based on the utilization of combinatorial antibodies such as for KIR2DL1 with KIR2DL1/S1 and KIR2DL3 with KIR2DL2/L3/S2 [29]. Expression of PD-1 was reported to be efficiently captured by one clone, PD.1.3.1.3. and to a lesser extent by two other competing clones (MIH4 and EH12). This specific clone lacks widespread commercialization and was thus not available in a fluorochrome that could fit in our panel, which led us to keep EH12 to at least assess PD-1 positivity [275].

1.2 Single-stained controls

Compensation controls were performed using single-color staining on compensation beads for dimly expressed markers and on PBMCs for highly expressed markers. For single-stained controls on beads, one drop of beads solution was diluted at a ratio of 1:1 with PBS. All antibodies were micro-centrifuged

at 1200 x *g* for 10 seconds to avoid any aggregates. For each single-stained control, 1 μ l of antibody solution was incubated with 50 μ l of diluted beads for 20 minutes at 4°C. After incubation, beads were washed twice by pelleting with 200 μ l CSB at 600 x *g*. For single-stained controls on cells, approximately $1-3 \times 10^6$ cells were stained with the appropriate concentration of each antibody solution and incubated for 20 min at 4°C protected from the light. Cells were subsequently washed twice by pelleting with 2 ml CSB and resuspended in 300 μ l CSB for acquisition.

1.3 CyTOF barcoding

The process of barcoding is advantageous in many regards: (i) increased technical accuracy by minimizing inter-sampling pipetting errors and variations in incubation times; (ii) reduced reagent consumption; (iii) reduced bench time work; (iv) minimization of sample cross-contamination; (v) improved sample throughput; (vi) reduced cell doublets [276]. In the present study, we took advantage of one of the current three available barcoding approaches, suitable for fixed and permeabilized cells (Cell-ID® 20-Plex Pd barcoding kit (Fluidigm, San Francisco, CA, USA)). In essence, this “intracellular” barcoding concept utilizes a combination of different chelated palladium (Pd) isotopes combined with isothiocyanobenzyl-EDTA reagents. The chemical principle of the linkage can be explained as follows: an acrylic acid polymer containing a repeat of fixed polymer units is functionalized with multiples copies of NHS (N-Hydroxysuccinimide) reactive groups. Within these groups is embodied a DOTA (1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid) chelator or alternatively a DTPA (diethylene triamine pentaacetic acid) chelator that allows chelation of trivalent metal lanthanides (Ln^{3+}). Incorporated chelated Ln and other metals have a dissociation constant of 10^{-16} resulting in a nearly impossible leaching or exchange between metal isotopes of different antibodies. The polymer contains a terminal maleimide group that allows the conjugation to the antibody, building a stable thioether linkage with reduced disulfide groups in the hinge region of the Ig heavy chain.

The number of unique combinations equals to $n! / k! (n-k)!$ with n being the number of barcode reagents. Thus, the use of a combinatorial conjugation of 3 out of 6 palladium isotope based barcoding scheme allows the generation of 20 different barcodes.

1.4 CyTOF heavy metal conjugation of antibodies

Fourteen antibodies targeting low abundance antigens had to be in-house *de novo* conjugated with polymer-metal isotope tags. For this purpose, the MaxPar® X8 antibody-labeling kit provided by

Fluidigm was used: metal chelation was performed by adding Ln metal solutions (final 0.05 M) to MaxPar chelating polymers in 95 μ l L-buffer followed by a 40 min incubation at 37 °C. Metal-loaded polymers were washed twice with L-buffer using a 3 kDa amicon filter at 12,000 x *g* for 30 minutes. Simultaneously, antibody buffer exchange was performed by washing 100 mg of antibody with R-buffer using a 50 kDa amicon filter. Antibodies were then reduced with 100 μ l of 4 mM TCEP tris(2-carboxyethyl)phosphine hydrochloride) for 30 min at 37°C and washed twice with 300 μ l C-buffer. Partially reduced antibodies and metal-loaded polymers were incubated together for 60-120 min at 3°C. Thereafter, conjugated antibodies were washed four times with 400 mL W-buffer and collected following two centrifugations with 50 mL W-buffer into an inverted column. The protein content was assessed by NanoDrop IgG sample type measurement to ensure a correct conjugation. For storage purposes, antibody stabilization buffer (Candor Bioscience, Wangen, Germany) was added to achieve a final solution with a ratio of 50% buffer, 50% antibody. In addition, in-house conjugated antibodies were tested on beads on the CyTOF instrument.

1.5 Antibody titration

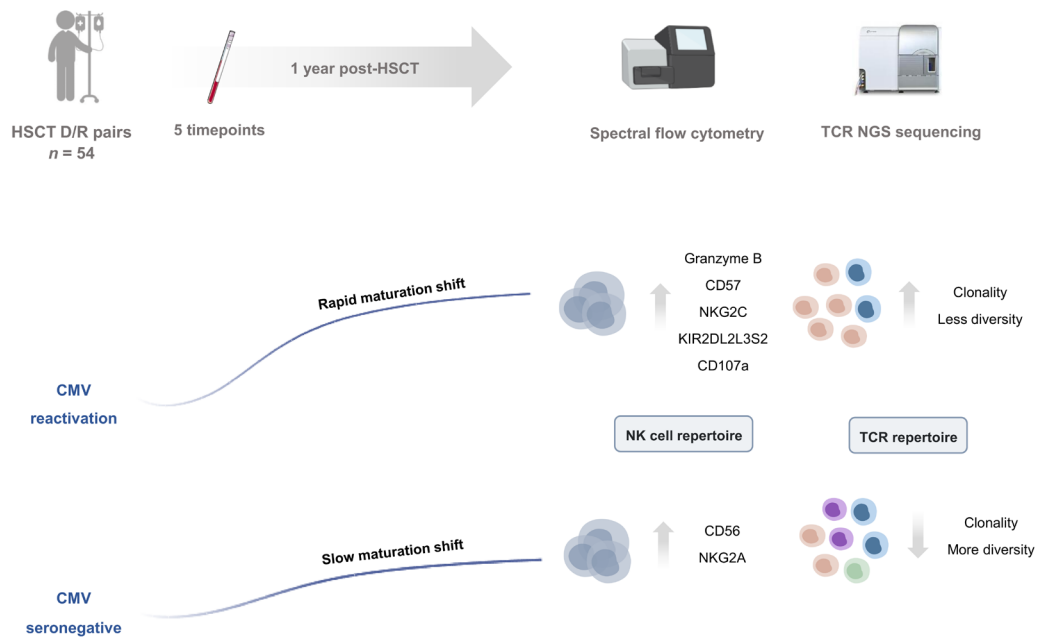
The procedure of titrating antibodies ensures the correct functionality of antibodies and determines optimal levels of staining, minimizing non-specific signals that might occur at increasing concentrations. For each designed panel, antibodies were co-titrated on cryopreserved PBMCs, following three consecutive stages: at first, major lineage markers were titrated and served as FMO and gating markers for the subsequent titrations. Thereafter, the remaining extracellular receptors (NCR, NKG2, KIR) were titrated and finally the activation and intracellular markers. Antibodies for mass cytometry were co-titrated at two-fold dilutions with four different concentrations, while for FSFC six different concentrations were defined. Approximately $1-3 \times 10^6$ PBMCs were stained for each concentration. For the third titration stage, PBMCs were stimulated with a chemical activation stimulus phorbol 12-myristate 13-acetate (PMA)/ionomycin. The final titer of each antibody was determined on both the visual optimal separation between positive and negative populations and the computation of the separation index when possible.

CHAPTER IV

Deciphering the immune reconstitution dynamic post-transplant and its relation according to clinical events

This chapter is based on following publication:

Schäfer A., et al. *Natural killer and T cell repertoire reconstitution is established early after allogeneic hematopoietic stem cell transplantation and is profoundly imprinted by CMV reactivation.* Blood advances (2024).



1. Current state of research

There is an increasing body of evidence indicating that our immune system state is subject to heritable and non-heritable influences. It is believed that up to 40 % of the immune system phenotype variations are a consequence of one's individual genetic framework [277]. Thus, a large part of the immune state is modulated by external non-genetic factors such as age, infections and lifestyle with the ability to leave long-lasting phenotypic imprints [278]. Given the transplant setting's complexity, isolating cause-effect relationships between the developing immune system and extrinsic events is challenging. Several studies have, however, delineated factors capable of interfering with the recreating immune system with the potential to curtail or amplify selected immune compartments.

1.1 Extrinsic NK cell reconstitution modulators

1. Immunosuppressive drugs

NK cell proliferation inhibition has been demonstrated for several immunosuppressive drugs encompassing methylprednisolone, rapamycin, cyclosporine A and mycophenolate mofetil in a dose-dependent manner and an incremental effect when combined. Methylprednisolone had the most potent inhibitory effect on cytotoxic abilities and cytokine production of NK cells compared to all others [279]. Cyclosporin A has additionally been shown to suppress *in vitro* proliferation of NK cells, especially CD56^{dim} CD16⁺ KIR⁺ NK cells, resulting in a relative increase in the number of immature CD56^{bright} CD16⁻ KIR⁻ NK cells [280]. A recent study showed an altered NK cell reconstitution defined by a reduced absolute NK cell count paired with fewer CD56^{bright} and NKG2A⁺ CD56^{dim} NK cells in the late time course after transplantation of sirolimus-treated recipients compared to conventional immunosuppression. This observation was attributed to a reduced *in vitro* proliferation of NK cells unrecoverable by IL-15. In contrast, T cell reconstitution was shown to be mostly spared out by the addition of mTOR inhibitors [191].

2. GvHD prophylaxis

Mounting evidence indicates a deleterious effect of the GvHD prophylaxis cyclophosphamide on the quantitative and qualitative NK and T cell reconstitution. Russo *et al.* were the first to reveal an initial deep ablation of circulating mature NK cells by post-transplant cyclophosphamide (PTCy) – probably stemming from the donor's graft. This was followed by a delayed rebound of reconstituting NK cells displaying an immature phenotype (CD56^{bright}, NKG2A⁺, CD62L⁺) concomitantly with an enrichment in IL-15 [190]. A following report confirmed these results, revealing an increased proportion of immature NK cells (CD56^{bright}, CD16⁻) during the first 6 months compared to HLA-matched transplants receiving a conventional tacrolimus/methotrexate regimen. Comparing their anti-leukemic effect *in vitro*, NK cells at

day 30 in PTCy-treated recipients experienced a reduced functional competence than their donor counterparts [281]. A following report interrogated the differential recovery of NK and T cells between combined PTCy and anti-thymocyte globulin (ATG) vs. only PTCy treated recipients, demonstrating that PTCy treated recipients were enriched in immature NK cells (NKp46⁺, 2B4⁻) while a higher amount of mature NK cells were present in PTCy and ATG treated recipients. Further, loss of alloreactive NK cell subsets expressing KIR2DL2/3 has been linked to PTCy treatment, potentially impeding a GvL effect [282]. Taken together, these results suggest that ATG induces an inflammatory setting favoring the maturation of re-emerging NK cells compared to PTCy alone. The accelerated maturation in PTCy and ATG treated recipients could additionally be attributed to the *in vivo* T cell depleting action, increasing the availability of homeostatic resources for NK cells. Indeed, there are numerous studies demonstrating an NK cell recovery impairment in relation to T cells as a consequence of a dual competition for identical homeostatic resources such as IL-15 and IL-2 [217, 283, 284].

3. Chronic viral infections

The human virome has been recognized as one of the decisive shapers of immune system states. First evidence has been drawn from animal models revealing a significant linkage between CMV and the proliferation of murine NK cells acquiring a memory phenotype by expressing Ly49H receptors [285]. Following those reports, studies have shown that CMV infection imprints NK cell subpopulation in humans by inducing the expansion of adaptive highly differentiated NK cells co-expressing CD57 and NKG2C [49, 286].

Following transplantation, several studies reported a priming of NKG2C⁺ CD57⁺ CD56^{dim} adaptive NK cells in CMV-reactivating recipients independently of the clinical symptomatology. This expansion occurs several days post-reactivation with an incremental increase of up to 11-fold during the time-course of several months post-transplant [287-292]. NKG2C expression has not exclusively been confined to mature NK cells but has been transiently associated with immature CD56^{bright} cells – reflecting imprinting effects on NK cell ontogenesis [289]. High-resolution immunophenotyping studies have brought forward additional insights into their phenotypic characterization: NKG2C⁺ NK cells were shown to be associated with the expression of T-bet, perforin, granzyme A and B and the downregulation of NKp46, CD62L and NKG2A. CMV reactivation has further been linked to the presence of a subset CD57⁻ CD56^{dim} NKG2A⁺ displaying an activating effector profile with the expression of CD8, NKp44, NKG2D with equal cytotoxic and immunomodulatory capacities as adaptive NK cells [293, 294].

In the light of such results, we are tempted to assume that the induced inflammatory setting of CMV reactivation may led NK cells to break tolerance and become educated for non-self KIRs. Data on the modulation of NK cell alloreactivity are scarce and not consolidated. While comparable cytotoxicity of self and non-self KIR⁺ NK cells was observed, Foley *et al.* failed to find evidence for this hypothesis and concluded that non-self KIR⁺ NK cells remain tolerant to their immunogenetic milieu upon CMV reactivation [292, 295]. Further, CMV reactivation has been shown to render immunomodulatory capacities of NK cells by producing elevated levels of TNF α and IFN γ . Indeed, single-cell proteomics *in vitro* assays and single-cell transcriptional data showed that adaptive NK cells are potent producers of IFN γ – especially for self-HLA ligands, overriding non-adaptive NK cells post-HSCT [288, 289].

Whether other viral infections own the ability to skew the NK cell repertoire during immune reconstitution remains poorly investigated. In non-transplant settings, several studies have reported an expansion of adaptive NK cells following other viral infections such as hantavirus [296], chikungunya [297], HIV-1 [298] and Sars-Cov-2 [299] while data are lacking for Epstein-Barr virus (EBV) [300], hepatitis C virus (HCV) [301] and herpes simplex virus (HSV) 2 [302]. These observations likely reflect the co-replication of CMV alongside these infectious episodes, although further studies need to be conducted to warrant this hypothesis. A recent study has showcased a differential NK cell reconstitution pattern alongside human pegivirus-1 presence after allogeneic HSCT. Viraemic recipients displayed a less differentiated phenotype reflected by a reduced CD16, CD57 and granzyme A/B expression as compared to non-viraemic recipients at 6 months post-transplant [303].

1.2 TCR reconstitution in allogeneic HSCT

Individual composition and diversity of TCR CDR3 β repertoire results from the convergence of genetic and non-genetic determinants: a well-known feature of our adaptive immune system is its decreased response alongside aging, leading to an increased susceptibility of older individuals to pathogens. Britanova *et al.* were the first to take advantage of TCR sequencing to track age-related contractions, demonstrating a linear TCR diversity decrease with age in a cohort of 39 healthy individuals aged from 6 to 90 [304]. Yoshida *et al.* followed TCR repertoire changes in a time window spanning 20 years, showing that the CD4⁺ T cell compartment is mostly unaffected while aging was particularly imprinting the CD8⁺ T cell compartment [305]. It is believed that the TCR richness declines by a factor of 2 to 5 during the lifetime spanning young and old adulthood, presumably without evidence of leaving “holes” in the immune repertoire. HLA polymorphism has been recognized as a potential genetic modulator founded on the hypothesis that HLA diversity enhances the immunopeptidome diversity, leading to a potential greater TCR repertoire diversity, although research studies are contradictory as to this regard [306, 307]. Non-genetic influences have equally been shown to impose constraints on the TCR repertoire, with the most striking effect being attributed to chronic viral infections, especially CMV [308, 309]. Vaccination episodes have further been proven to disrupt TCR repertoire, inducing a reorganization of immunodominant TCR clones [310]. Interestingly, the TCR repertoire of preterm and term children has been shown to be significantly different early post-natal with an overall enhanced diverse repertoire enriched in less mature, shorter CDR3 amino acid (CDR3aa) clones in premature children. Later on, they converge towards a similar TCR repertoire with regards to their clonality and richness profile [311]. Whether this convergence reflects the thymus intense capacity of newborns to compensate or whether environmental factors drive maturation is unknown.

Only a few studies have explored the TCR reconstitution after HSCT at a high-resolution level with a substantial level of variance in terms of the sampling scheme and cohort size. Some few common patterns can nevertheless be brought forward [308, 309, 312-314]:

- (i) An altered TCR restoration pattern in recipients was confirmed in all studies defined by an increased clonality and a decreased TCR richness.
- (ii) CMV serostatus majorly drives the TCR clonality and clone composition reconstitution, with CMV seropositive recipients displaying significantly higher TCR clonalities as compared to their negative counterpart.

Several intrinsic observations have been concluded out of these studies, to mention the most important findings:

- (i) Younger recipient and donor ages have been associated with an increased TCR diversity [308].
- (ii) The use of ATG and PTCy was associated with an increased clonality and a reduced TCR richness [312].
- (iii) Donor and recipient pre-HSCT inverse Simpson index are predictive of TCR diversity [313].
- (iv) Vaccination episodes post-HSCT greatly restrict the diversity of TCR repertoire as seen in HSCT recipients following Sars-Cov-2 vaccination [315].

2. Manuscript III

NK and T cell repertoire is established early after allogeneic HSCT and is imprinted by CMV reactivation

Antonia Schäfer¹, Zuleika Calderin Sollet¹, Marie-Priscille Hervé¹, Stéphane Buhler¹, Sylvie Ferrari-Lacraz¹, Paul J. Norman³, Katherine M. Kichula³, Ticiana D. J. Farias⁴, Stavroula Masouridi-Levrat², Anne-Claire Mamez², Amandine Pradier², Federico Simonetta², Yves Chalandon², Jean Villard^{1*}

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 *Department of Biomedical Informatics, University of Colorado School of Medicine, Aurora, CO, USA*

⁴ *Department of Biological Sciences, The University of North Carolina at Charlotte, Charlotte, NC, USA*

Corresponding author:

*Jean Villard, Transplantation Immunology Unit and National Reference Laboratory for Histocompatibility, Geneva University Hospitals, Gabrielle-Perret-Gentil 4, 1211 Geneva 4, +41 22 372 93 94/ +41 79 553 34 09, jean.villard@hcuge.ch

Keypoints:

- The TCR and NK cell repertoire establish within the first 90 days post-transplant
- CMV reactivation is associated with an enhanced TCR clonality and a concomitant shift towards enhanced maturation and functionality of NK cells

ABSTRACT

Besides genetic influences, non-genetic factors such as graft-versus-host disease (GvHD) and viral infections have been shown as important shapers of the immune reconstitution and diversification processes after hematopoietic stem cell transplantation (HSCT). However, the differential susceptibility to immune modulation by non-genetic factors is not fully understood. We determined to follow the reconstitution of the T cell receptor (TCR) repertoire through immune-sequencing, of natural killer (NK) cells using a 35-marker spectral flow cytometry panel and in relation to clinical events. Longitudinal investigation was performed on samples derived from 54 HSCT recipients during the first-year post-HSCT. We confirmed a significant contraction in TCR repertoire diversity with a remarkable stability over time. CMV reactivation had the ability to significantly change TCR repertoire clonality and composition, with a long-lasting imprint. Our data further revealed skewing of NK cell reconstitution in CMV reactivated recipients, with an increased frequency of KIR2DL2L3S2⁺ adaptive, cytolytic and functional CD107a⁺ NK cells concomitant with a reduced pool of NKG2A⁺ NK cells. We provided support that CMV might act as one of the more important drivers of peripheral homeostatic proliferation of circulating specific T and NK cells, which can be viewed as a compensatory mechanism to establish a new peripheral repertoire.

Keywords

T-cell receptor (TCR); Killer-cell immunoglobulin like receptor (KIR); Cytomegalovirus (CMV); Hematopoietic stem cell transplantation (HSCT); Systems immunology

1. INTRODUCTION

Allogeneic hematopoietic stem cell transplantation (alloHSCT) is a well-established immune therapy for patients having hematological malignancies such as leukemia or primary immunodeficiencies [161]. Although long term survival has strongly improved over the last decade, it remains a high risk treatment and its success is hampered by the occurrence of infections, immunological complications and disease relapse [316]. Its potential curative effect relies in the recreation of new competent hematopoietic and immune cells, combined with the eradication of residual tumoral cells through the graft-versus-leukemia effect. Restoration of the immune system's state thereafter is thought to be the contribution of a first wave of mature donor-derived immune cells followed by a second wave of *de novo* stem cell derived cells from the innate followed by the adaptive immune system [139, 166]. Alongside this process, the developing immune system is exposed to a wide variety of extrinsic factors, forcing this system to respond and adapt. Among these factors, we and others have demonstrated that cytomegalovirus (CMV) is an important immune state modulator imprinting the T cell receptor (TCR) repertoire at one year post-transplantation, and lasting over more than 5 years [308, 309]. The natural killer (NK) cell repertoire reconstitution has been less studied but similar observations regarding the altering role of CMV have been published [287, 289, 290]. These shifts are specifically driven by the massive clonal expansion of antigen-specific terminally differentiated CD8⁺ T cells and NKG2C⁺ CD57⁺ adaptive NK cells, combined with a skewed expression of self-specific inhibitory Killer-cell immunoglobulin like receptors (KIRs) [277, 317]. Recent findings hint towards a similar inflationary memory response with a lack of adaptive memory-like NK cell contraction as previously shown for CD8⁺ T cells during viral latency [317, 318].

In light of such findings and with regards to the limited data available on their combined reconstitution pattern, this report aims to dissect the inter-dependent temporal NK cell and TCR repertoire restoration in the unique human alloHSCT case model and further trace relations to clinical events. Comprehensive immune profiling was performed by leveraging high-dimensional single-cell proteomic technique with next-generation sequencing in 54 HSCT recipients monitored longitudinally over the time course of one year post-transplant.

2. MATERIALS AND METHODS

2.1 Study design

All patients having undergone an allogeneic HSCT at the Geneva University Hospital between December 2020 and April 2022 were enrolled in this study, encompassing a total of 54 donor/recipient (D/R) pairs. The cohort's detailed demographic and clinical characteristics are summarized in Table 1. Medical complications such as infectious events (viral, bacterial), acute (aGvHD) and chronic (cGvHD) graft-versus-host disease episodes, relapse of the initial disease, transplant-related or non-related deaths and CMV serologic status/reactivation (defined as CMV DNA in plasma above the limit of detection, i.e., 2.1×10^1 UI/ml, in patients with or without clinical symptoms) were recorded (Table 2, Supplementary Figs. S1A,B). All recipients with a positive CMV serology received a Letermovir CMV prophylaxis (480 mg per day with tacrolimus-based immunosuppression or 240 mg per day with ciclosporin-based immunosuppression) from the day of transplantation up until 100 days post-transplant. For all recipients with a negative CMV serology, no prophylaxis was given and CMV viremia was monitored on a regular basis after transplantation. All patients were treated for a CMV reactivation with ganciclovir, which was stopped once the viremia reached an undetectable level (< 21 copies).

This study was approved by the ethical committee of the institution (CER 06–208 and 08–208R) and performed according to the Declaration of Helsinki principles.

2.2 Sample collection and processing

Whole blood samples were serially collected from HSCT recipients at 3-, 6-, 9- and 12-months post-HSCT and from the donor prior to the transplant. Peripheral blood mononuclear cell (PBMC) isolation was performed by Ficoll density gradient centrifugation as well as DNA extraction (Supplementary material).

2.3 High resolution HLA typing

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).

2.4 Bulk TCR CDR3 β sequencing

High-throughput TCR complementary determining region 3 beta (CDR3 β) sequencing at survey resolution was performed using the immunoSEQ[®] Assay (Adaptive Biotechnologies, Seattle, WA, USA) according to the manufacturer's instructions (Supplementary material).

2.5 High resolution KIR genotyping

All donors were genotyped at high-resolution for all KIR loci. To this end, a DNA probe-based capture method was used as described here [83] (Supplementary material).

2.6 Spectral flow cytometry immunophenotyping

The antibody panel for the immunophenotyping comprised a total of 35 markers. Antibody clones, concentration, corresponding fluorochromes and supplier are summarized in supplementary Table S1. Surface and intracellular staining and sample acquisition were performed according to the standard protocol (Supplementary material).

2.7 CD107a mobilization assay

NK cell cytotoxicity potential assessment was based on the CD107a mobilization assay as previously described [319]. Shortly, after overnight resting, 100 μ l of the previously thawed PBMCs ($2-4 \times 10^6$) were transferred from each sample to a 5 ml PP tube and cocultured with 100 μ l of K562 cells at an effector: tumor cell ratio of 5:1 or without target for unstimulated controls. CD107a-BUV396 (BD Biosciences, San Diego, CA) antibody was added to each tube prior to the stimulation, including all unstimulated samples to assess any spontaneous degranulation. After 1 hour of stimulation, brefeldin A and monensin were added to the co-culture and cells were stimulated for another 4 hours at 37°C, 5% CO₂. After a total of 5h of stimulation, samples were washed for the subsequent flow cytometry staining as described in the section below. An internal control of stimulated cells without CD107a staining was added to assess positivity signals.

2.8 Computational and statistical analysis

2.8.1 PING bioinformatic pipeline

Pushing Immunogenetics to the Next Generation pipeline was applied for sequencing filtering, alignment, gene content and allelic genotype determination derived from the next generation

sequencing fastq files [82, 83]. High-resolution KIR genotyping, copy number and alleles were determined for all KIR genes (KIR2DS1, 2DS2, 2DS3, 2DS4, 2DS3/S5, 3DL1/3DS1, 2DL1, 2DL2/L3, 2DL4, 2DL5A/B, 3DL2, 3DL3) and the two pseudogenes (KIR2DP1, 3DP1).

2.8.2 TCR data processing

Raw data processing and sequencing analyses were performed using the immunoSEQ Analyzer 4.0 online platform (<http://www.adaptivebiotech.com/immunoseq>) and in R environment (version 4.1.3, R Core team, R foundation for statistical computing). TCR clonality, overlap analysis and identification of public and private TCR clones are detailed in the supplementary material.

2.8.3 Prediction and inference of antigenic specificity

For the assessment of pathology specificity, several public databases were downloaded (*VDJd* [320], *McPAS-TCR* [152], *ImmuneCODE* [321]), merged together and individual CDR3 amino acid (aa) sequences were filtered when belonging to the following categories: viral and bacterial disease, autoimmunity, allergenic, cancerous diseases and GvHD. Subsequently, individual CDR3aa sequences from the cohort were queried against the cleaned databases. HLA restriction was available for all CMV specific clonotypes and is listed in supplementary table S4.

2.8.4 Spectral flow cytometry data preprocessing and gating

FCS3.0 files were extracted from SpectroFlo® Software v3.1.0 after spectral unmixing and compensation adjustments and analysed by manual gating (Supplementary material).

2.8.5 Statistical analysis

Continuous data were reported as medians with interquartile ranges (IQR). Normality distribution of the data was tested using Shapiro-wilk test. Two-group comparisons were analysed by the two-tailed unpaired Wilcoxon rank-sum test and the matched-paired Wilcoxon signed-rank test. FDR correction was applied for multiple testing. Linear regression using the glm function in R was used to assess the contribution of the CMV-specific TCR clone frequency on the TCR clonality, assuming non-parametric data. All statistical analyses were performed in R environment (version 4.1.3, R Core team, R foundation for statistical computing) with a *p*-value of 0.05 set as threshold for significance.

3. RESULTS

In this study, we set out to characterize coordinated changes involved in the immune reconstitution of the NK and T cell compartment following alloHSCT and to interrogate the effect of non-genetic factors on the immune state restoration. To this end, 54 recipients and their respective donors were enrolled in a prospective manner at the Geneva University Hospital between December 2020 and April 2022. We monitored TCR repertoire and NK/T cell compartment restoration in longitudinal blood samples drawn at 3 (t_3), 6 (t_6), 9 (t_9) and 12 (t_{12}) months post-HSCT from the recipients and from their respective donors before HSCT (t_0). We performed targeted TCR CDR3 β chain high-throughput sequencing on a total of 254 samples and analyzed an overall of 2 710 450 productive TCR rearrangements after correcting for non-productive CDR3aa sequences. To interrogate temporal changes in the NK and T cell compartment post-HSCT, we further designed and applied a high-dimensional antibody panel targeting 35 surface and intracellular markers on a total of 228 samples derived from 47 recipients (Supplementary Fig.S1A). In a subset of recipients ($n = 24$), we additionally performed CD107a *in vitro* assay to interrogate the NK cell functional competence.

Recipients' TCR repertoire is subject to a long-lasting restricted diversity post-transplant

In the first place, we aimed to establish a comprehensive view of the TCR repertoire reconstitution dynamic during the first-year post-transplant with respect to its diversity, clone organization and composition. The recipient's TCR repertoires showed a significant increase in clonality (i.e., reduced diversity) over the first 3 months, which remained stable at subsequent timepoints ($p = 1.157e-08$ between t_0 and t_3 ; $p = 0.076$ between t_3 and t_{12} ; Fig. 1A). Despite the clonality being stable, the different clone size categories stratified according to their productive frequency were subject to changes post-transplant: While the repertoire during the first 3 months was almost exclusively dominated by expanded clones, a progressive significant increase in non-expanded clones (1-2 template) was apparent upon 6 months and lasted until 12 months ($p = 1.5e-06$) (Fig. 1B). To better understand the evolution of TCR composition alongside time post-HSCT, we computed the TCR overlap between any two time-points pre- and post-transplant on an intra-individual level, considering the abundance and the clone identity. Donor-recipient overlap indices were low and did not increase during the entire follow-up time until one-year post-transplant ($p = 0.566$) suggesting no further major donor-derived clone expansion (Fig. 1C).

The overlap between 3 and 12 months post-transplant was low (median = 0.228, IQR: 0.0556 – 0.537) and significantly increased from 6 months on, with the highest overlap being between 9 and 12 months (median = 0.885, IQR = 0.579 – 0.943; $p = 2.986\text{e-}06$ between t_3 - t_{12} and t_9 - t_{12}) indicating that the major clone size hierarchy is set up at an early time after transplantation while the TCR composition turnover and qualitative reconstitution stabilize over time (Fig. 1C).

CMV reactivation drives a strong TCR immune response

While genetics partly predicts the TCR repertoire state post-transplant, extrinsic factors play an important role in shaping immune states [306]. To interrogate their effect, we stratified our cohort into three distinct groups with respect to the CMV serostatus of the recipient [322]; CMV seronegative (D-/R- and D+/R- , $n = 20$), CMV seropositive but no documented reactivation (D-/R+ and D+/R+ , $n = 21$) and whether they were subject to a CMV reactivation (D-/R+ and D+/R+ , $n = 13$) (Supplementary Table S2, S3). Only recipients who were positive for CMV reactivated post-transplant and displayed CMV reactivation up until 9 months post-transplant (Supplementary Fig. S1B).

CMV seronegative and seropositive recipients displayed a post-transplant repertoire stable in terms of its diversity ($p = 0.87$; $p = 0.65$, respectively), as compared to CMV seropositive recipients subject to a CMV reactivation episode who experienced a significant increase in clonality ($p = 4.07\text{e-}02$) (Fig. 2A, Supplementary Fig. S7). In CMV reactivated recipients at 12 months, this shift in TCR repertoire clonality was dominated by a large-scale clone expansion of more than 1% (Fig. 2B). Representative reconstitution trajectories of selected recipients are provided in Supplementary Figure S8. The overlap between the third and twelve months was the lowest in recipients having a CMV reactivation (median = 0.035, IQR = 0.005 – 0.247; $p = 0.013$ between CMV seronegative and reactivated), likely reflecting a rapid turnover of TCR repertoire with major changes in the clone composition and expansion of competitive T cell clones (Fig. 2C). In light of such differential reconstitution trajectories, we thought to interrogate the origin and the presence of distinct molecular features associated with the presence of CMV. CMV reactivated recipients showed an incremental and significant increase in the cumulative frequency of hyperexpanded (>1%) public clones ($p = 0.046$) between the third and twelve months as compared to CMV seronegative and seropositive recipients ($p = 0.85$, $p = 0.945$, respectively) (Fig. 2D). To assess the immune responses specificity, we queried our dataset against public databases encompassing a total of 20'018 experimentally annotated CMV-specific TCR clones. All recipients had a good capacity to present CMV peptides according to their HLA class I restriction. We did not observe

any significant enrichment in the cumulative frequency of CMV-specific clones following CMV reactivation compared to CMV seronegative or seropositive recipients ($p = 0.687$ at t_{12}) (Fig. 2E). We further extended the query to all pathogen-associated clones, however, no significant increase in the cumulative frequency of potential cross-reactive clones was noted (Supplementary Fig. 3). However, under the assumption of transferred immunity, we observed a significant enrichment in the cumulative frequency of CMV-specific TCR clones shared between the donor and the recipient post-transplant in D+/R+ ($n = 4$) compared to D-/R+ CMV reactivated recipients ($n = 9$) ($p = 2.039\text{e-}06$) (Fig. 2F, Supplementary Fig. 4A). The cumulative frequency of shared anti-CMV clones in CMV mismatched non-reactivated recipients was comparable to the level in CMV matched non-reactivated recipients ($p > 0.05$) (Supplementary Fig. 4B). In line, we found that there was a significant contribution of CMV-specific TCR clone's productive frequency on the clonality in D+/R+ reactivated recipients ($p = 6.89\text{e-}07$, linear regression) while the effect was non-significant on the clonality in D-/R+ CMV reactivated recipients ($p = 0.443$, linear regression).

We have analyzed the impact on the TCR repertoire reconstitution according to the presence of post-transplant cyclophosphamide (PTCy), GvHD and Sars-Cov-2 with minor changes noticeable (Supplementary Table S5).

Dynamic immune cell regeneration during the first year post-transplant

Alongside TCR repertoire assessment, we systematically assessed the NK cell and T cell compartment and its recreation at the phenotypic level, based on a 35-color antibody panel. This global view showed differential recovery patterns for regeneration of the major canonical cell subsets: At three months, the immune compartment was dominated by NK cells and monocytes, with a descendent kinetic up until 12 months. NK cells were significantly enriched compared to their donor counterpart ($p = 0.018$, Supplementary Table S6). CD4⁺ and CD8⁺ T cells were significantly reduced early post-transplant ($p = 7.5\text{e-}06$, $p = 0.024$) and slowly recovered up until 12 months ($p = 4.4\text{e-}05$; $p = 8.1\text{e-}05$) (Fig. 3A, Supplementary Table S6). The frequency of circulating B cells was subject to a gradual increase from 3 months up until 12 months, without reaching statistical significance ($p = 0.11$) (Fig. 3A). CD56^{bright} NK cells expanded transiently after transplantation, significantly outnumbering donor values ($p = 7.1\text{e-}07$) and were subject to a sharp significant decrease from 3 months ($p = 8.4\text{e-}06$) on, with a progressive conversion towards NK cell maturation reflected by a significant increase in CD56^{dim} CD16⁺ NK cells ($p = 0.0094$) (Fig. 3B).

Then we focused on the NK cell repertoire using manual serial gating to profile the expression level and frequency of each functional marker on total CD56⁺, CD56^{dim} and CD56^{bright} NK cells. There was a significant enrichment of CD56^{dim} CD16⁺ CD57⁺ ($p = 2.0\text{e-}07$) NK cells at the expense of CD56^{dim} CD16⁺ NKG2A⁺ NK cells ($p = 1.1\text{e-}05$) (Fig. 3C). This equilibrium became stable at 9 months with no further significant changes of both subsets ($p = 0.705$, $p = 0.755$) (Fig. 3C). No significant changes were noted in terms of the frequency of NKp30, NKp46, NKp44, TIGIT, PD-1, NKG2D expressing CD56^{dim} NK cells between 3 and 12 months (Supplementary Fig. S5A). The maturation and functional competence of NK cells are tightly regulated by the interaction between inhibitory KIRs and their HLA-I ligand, dictating their cytotoxic strength [323]. Given the inter-individual differences in KIR gene content [324], we performed high-resolution KIR genotyping to accurately assess which KIRs could be expressed. Whereas HLA-B specific KIRs showed a stable behavior in the post-transplant course ($p = 0.97$), HLA-C specific inhibitory KIRs displayed differential reconstitution kinetics: KIR2DL1⁺ subsets were significantly downregulated compared to the donors ($p = 0.018$) and displayed a progressive reconstitution up until 12 months ($p = 0.017$) (Fig. 3D). KIR2DL2L3S2⁺ subsets showed similar frequency to the donors early post-transplant and showed a significant increase during time-course of one year post-transplant ($p = 0.022$), specifically KIR2DL2⁺S2⁺ subsets ($p = 0.0018$) (Fig. 3E). The proportion of KIR2DL5⁺, KIR2DS4⁺, KIR2DS1⁺, KIR3DS1⁺ cells did not change across all observed time-points ($p > 0.05$) (Supplementary Fig. S5B).

CMV reactivation induces an accelerated maturation of NK cell subsets

The presence of CMV has been shown to be also a decisive regulator of the NK cell immune state [277]. To assess whether these groups are subject to a divergent reconstitution trajectory, we first calculated an Aitchison's distance, measuring the difference in the cell composition based on 62 manually gated cell subsets. Intra-individual distances between 3 and 12 months were significantly higher in CMV reactivation recipients ($p = 0.006$) reflecting major changes in the NK and T cell immune state composition of CMV reactivated recipients (Fig. 4A). This observation was firstly driven by a significant increase in adaptive NK cells co-expressing NKG2C and CD57 being associated with the presence of CMV ($p = 1.97\text{e-}05$) (Fig. 4B). CMV seropositive recipients had a higher frequency of total CD8⁺ T cells in contrast to CMV reactivated recipients at 3 months ($p = 0.021$). At 12 months, there was a significant enrichment in CMV reactivated recipients in circulating CD56^{dim} NK cells expressing mature and cytolytic markers such as CD16⁺, CD57⁺, granzyme B, perforin and FasL, CD56^{neg} CD16^{pos} unconventional NK

cells and granzyme B expressing CD8⁺ and CD4⁺ T cells. The data also revealed a higher rate of CD56^{dim} expressing the checkpoint inhibitory marker PD-1 in CMV reactivated recipients ($p = 4.60e-02$) (Fig. 4C). Finally, the frequency of KIR2DL2L3S2⁺ CD56^{dim} NKG2C⁺ NK cells was significantly tuned upwards ($p = 0.01$) in CMV reactivated recipients in contrast to KIR2DL1⁺ and KIR3DL1⁺ CD56^{dim} NKG2C⁺ NK cells and showed an incremental increase in granzyme B content with time ($p = 0.0033$) (Fig. 4D). In comparison, the NK cell compartment in CMV seronegative recipients was dominated by immature CD56^{bright} NK cells and NKG2A⁺ expressing CD56^{dim} and CD56^{bright} NK cells (Fig. 4C). There was no difference in the frequency of NKp30, NKp44, NKp46, TIGIT, NKG2D, TRAIL expressing CD56^{dim} NK cells ($p > 0.05$) (data not shown). Representative reconstitution trajectories of selected recipients are provided in Supplementary Figure S8.

We finally interrogated the functional competence of NK cells using a stimulation assay, co-culturing NK cells against HLA-I deficient cancer cell lines (K562). Among a total of 62 stimulated samples, we assessed CD107a degranulation response in manually gated CD56^{dim} NK cells (Supplementary Figs. S6A, B). CMV seropositive and reactivated recipients produced a higher frequency of CD107a⁺ CD56^{dim} NK cells post stimulation, compared to seronegative recipients post-transplant. These differences were statistically significant ($p = 0.045$, $p = 6e-04$, respectively) (Fig. 4E).

We have analyzed the impact on the NK cell repertoire reconstitution according to the presence of PTCy, GvHD and Sars-Cov-2 with minor changes noticeable (Supplementary Table S5).

4. DISCUSSION

AlloH SCT represents a unique case study to understand immune ontogeny and its behavior under external non-genetic influences. In line with previous findings by us and others, we confirmed a significant contraction in the TCR repertoire diversity following transplantation [308, 309, 312-314]. Our data further revealed a remarkable stability in the clonality over one year follow-up, suggesting that the clone size hierarchy establishes at a very early stage after transplantation and doesn't fluctuate significantly over distinct immunomodulatory stages. Supportive data from mathematical modeling in healthy individuals profiled that adaptation of these systems early in life involves a rapid and long-lasting expansion of TCR clones as a direct response to the lymphopenic environment with a reduced clonal competition [325]. The reappearance of non-expanded clones at latter stages after transplantation in our data likely suggests the progressive increase in peripheral clonal competition and stabilization in the quantitative homeostatic turnover.

Although studies have demonstrated an influential effect of many factors, such the use of PTCy [190, 312], we believe based on our data that CMV reactivation has a dominant immunomodulatory effect. This was reflected by the ability to enhance TCR clonality and induce compositional changes with a long lasting effect [308, 326]. Unexpectedly, we weren't able to assess an enhanced frequency of CMV-specific or other antigen-experienced TCR clones in CMV reactivated recipients based on *in silico* matching. While undeniably these public databases lack completeness, thus inducing a bioinformatic bias in detecting antigen-specific TCR clones, several biological explanations might account for this observation: There is supportive data demonstrating a reduced thymic activity following CMV infection as a consequence of a damaging CD8⁺ T cell infiltration [327, 328]. Observation that is further supported by a reduced presence of circulating TREC following CMV reactivation [329]. Thus, we might hypothesize that a reduced thymic output associated with or as a consequence of CMV reactivation might lead to an even greater peripheral homeostatic proliferation of circulating clones as a compensatory mechanism. The absence of CMV-specific T cells could be explained by the co-presence of GvHD which has been showcased to impair MHC-I and II processes by dendritic cells, thus corrupting efficient CMV-specific T cell priming [330]. Alternatively, proliferation of non-specific T cells could be induced as a bystander effect after cytokine release or based on the fitness of non-specific unexperienced T cell clones [328]. Less likely, we might attribute this effect to a broader feature of the

adaptive immune system as clonal unspecific expansion has been reported in other settings i.e., BCR repertoire following a Sars-Cov-2 infection [331].

Similarly to the TCR repertoire, the NK cell repertoire showed a substantially high level of reconstitution at 3 months post-HSCT. Our observation on the differential reconstitution trajectory of KIRs alongside the post-transplant time course mirrors previous findings [174, 332, 333]. Expansion of adaptive NK cells has been observed in recipients with CMV reactivation [289, 290, 334]. Our data further point towards a long lasting immune reconstitution skewing in terms of highly mature, cytolytic and functional NK and T cell subsets in CMV reactivated recipients, reflecting the very highly activated immune environment following CMV. Specifically, we first showed a bias towards a KIR-mediated reactivity with the increase in KIR2DL2L3S2⁺ adaptive NK cells concomitant with a reduced pool of NKG2A expressing NK cells. Horowitz *et al.* showed a concomitant increase in HLA-C expression on monocytes and myeloid lineages alongside CMV reactivation, likely related to an upregulation in immunomodulatory cytokines such as IFN γ and TNF α [335]. The developing nature of NK cell compartment early post-HSCT could be particularly sensitive to changes in the HLA environment and substantiate questions about the contribution of HLA-C ligands to an accelerated and enhanced NK cell licensing, especially given the fact that dendritic cells have been shown important mediators of NK cell education [335]. Supportive data showed that donor and recipient immunogenetic backgrounds play a role: The magnitude of single KIR2DL2L3⁺ NKG2C⁺ NK cell expansion has been shown to be rendered in donors and recipients bearing C1/C1 and C1/C1/Bw4 HLA genotype [294]. A recent genetic study further demonstrated a protective effect of KIR centromeric B02 motifs encompassing KIR2DL2 and KIR2DS2 on the relapse risk [208]. In contrast, data on the impact of CMV infection in healthy individuals have shown an upward tuning of KIR2DL1 on adaptive NK cells, which we couldn't confirm in our cohort and might question the implication of KIR/HLA mismatch with regard to the KIR2DL1 locus [317].

Several groups have reported that skewing NK cell repertoire into a maturation status promotes graft-versus-leukemia effect with CD57⁺ NKG2C⁺ NK cells being acknowledged as the main cell contributor to this observation independently of the cell source origin [336-339]. Based on our findings showing the CMV-induced immune changes were still present at least 3 months after CMV reactivation, it remains to be examined whether these cells undergo changes in their functional competence and phenotype, as we observed an increase in NK cells expressing exhaustion markers such as PD-1. Indeed, single cell proteomics have similarly highlighted an enrichment in the expression of PD-1 and LAG-3 on adaptive

NK cells, suggesting a progressive exhaustion and dysfunction with time [293]. As to this regard, we might interrogate the presence of a certain immunological time window in which the graft-versus-leukemia effect might reach its maximum and vanish subsequently.

Finally, although our study questions the impact of clinical events on immune system reconfiguration, it remains a difficult task to fully disentangle cause and consequence, which we cannot fully distinguish in our study. Our data revealed an enhanced expansion of shared CMV-specific T cells in CMV matched reactivated recipients in contrast to CMV mismatched reactivated recipients. Thus, from a clinical point of view, consideration of the CMV serostatus would be key in predicting immune response as intrinsic donor's graft immunity could be "transferred" to the recipient with CMV-specific immune cells undergoing peripheral homeostatic proliferation and contributing to a protective CMV-specific immunity. In addition, most of the CMV seropositive recipients without reactivation showed no enhanced frequency of adaptive memory-like NK cells suggesting that memory CMV-specific T cell clones are sufficient for a first-line protection against further viral replication, while adaptive NK cells are required to unleash NK cell antiviral activity in an acute manifested CMV replication [340]. These observations could also indicate that patients at high risk for primary infection might benefit from the infusion of third-party CMV specific CD8⁺ T cells or in case of refractory severe CMV infection as it is performed in several centers [341, 342].

This study has several limitations. The limited number of patients doesn't give us sufficient power to rule out the effect of other co-dependent non-genetic influences, such as immunosuppressive therapy and specific infections on a small subgroup of patients. The restricted number of identified specific antigenic TCR epitopes limited our *in silico* matching and prevented us from ruling out that proliferation of some T cell clones that are considered bystanders could have specificity. Finally, despite a high-resolution KIR genotyping, the KIR phenotyping cannot represent the real repertoire because of the limited number of antibodies that cover the KIR proteins at the NK cell surface.

In summary, our data indicated that the T and NK cell repertoire reconstitution after alloHSCT is majorly established during the first three months, and seems to be driven to a large part by the CMV serostatus and CMV reactivation and is persistent for at least one year post-transplant.

Data availability

The dataset for the TCR sequencing generated during and/or analyzed during the current study is available in the Yareta repository, <https://doi.org/10.26037/yareta:uluquhaq7zf4tb2lxbxbkvzse>. The other datasets used and/or analyzed during the study are available upon reasonable request to the corresponding author.

Author contributions

J.V designed the study; A.S, MPH and ZCS performed the experiments; A.S acquired, analyzed the data and performed the statistical analysis; A.S and J.V wrote and drafted the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This study was supported by 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 Philanthropy Settlement. AS was supported by a grant of the Fondation de Reuter. PJN was supported by NIH U01AI090905.

Acknowledgements

The authors are grateful to the technicians of the LNRH for their most efficient support for HLA genotyping.

Competing Interests

YC: consulting fees from MSD, Novartis, Incyte, BMS, Pfizer, Abbvie, Roche, Jazz, Gilead, Amgen, Astra-Zeneca, Servier; Travel support from MSD, Roche, Gilead, Amgen, Incyte, Abbvie, Janssen, Astra-Zeneca, Jazz, Sanofi, all via the institution. The other authors declare no competing interests.

MAIN PART FIGURE LEGENDS

Fig. 1: TCR diversity restricted during the first-year post-HSCT. (A) Evolution of TCR productive Simpson clonality before alloHSCT (t_0 , $n = 53$) in donors and at serial timepoints t_3 ($n = 54$), t_6 ($n = 52$), t_9 ($n = 47$), t_{12} ($n = 48$) after alloHSCT in recipients. **(B)** Fractal clonal size organization defined by the productive frequency of clones at indicated timepoints (t_0 , $n = 53$; t_3 , $n = 54$; t_6 , $n = 52$; t_9 , $n = 47$; t_{12} , $n = 48$). Each bar represents a single individual. The color-coded legend bar represents the stratification according to the individual clone's productive frequency. **(C)** TCR repertoire overlap between donor and recipient and between post-HSCT timepoints in recipients calculated by the Morisita-horn index. Morisita-horn-indices vary between 0 (no overlap) and 1 (complete overlap) and are represented along the y axis. Box plots display medians and interquartile ranges (IQR), with whiskers representing $1.5 \times$ IQR. Wilcoxon rank sum test with FDR correction in (A). Kruskal-wallis test and post-hoc Dunn's test in (C). All p values were 2 sided. Statistical thresholds: (****) $p < 0.0001$ and (NS) not significant.

Fig. 2: TCR repertoire restoration in relation to CMV serostatus and infection/reactivation. Stratification according to following groups: seronegative ($n = 20$), seropositive ($n = 21$) and reactivated ($n = 13$) CMV recipients applies to plots (A) to (E). **(A)** Evolution of TCR productive Simpson clonality at serial timepoints after HSCT fitted using LOESS regression with 95% confidence interval. **(B)** Productive frequency distribution at 12 months post-HSCT with respect to the groups. **(C)** TCR repertoire overlap computed by the Morisita-horn index between 3- and 12-months post-HSCT. The Morisita-horn-indices vary between 0 (no overlap) and 1 (complete overlap) and are represented along the y axis. **(D)** Private (upper panel) (i.e., observed in only one donor/recipient pair) and public (lower panel) (i.e., observed in two or more donor/recipient pairs or matching clonotypes present in public databases with antigen-specific validated TCRs) fractal clonal size organization at serial timepoints post-HSCT. Each bar represents a single individual. The color-coded legend bar represents the stratification according to the individual clone's productive frequency. **(E)** Cumulative productive frequency of CMV-specific TCR clones identified by *in silico* matching with public databases at serial timepoints post-HSCT. **(F)** Cumulative productive frequency of donor-recipient non-shared and shared (i.e., between each donor/recipient pair) CMV-specific TCR clones in D-/R+ ($n = 9$) and D+/R+ ($n = 4$) CMV reactivated recipients at all timepoints combined. The cumulative frequencies at each timepoint are detailed in Fig. S4A. Box plots display medians and interquartile ranges (IQR), with whiskers representing $1.5 \times$ IQR.

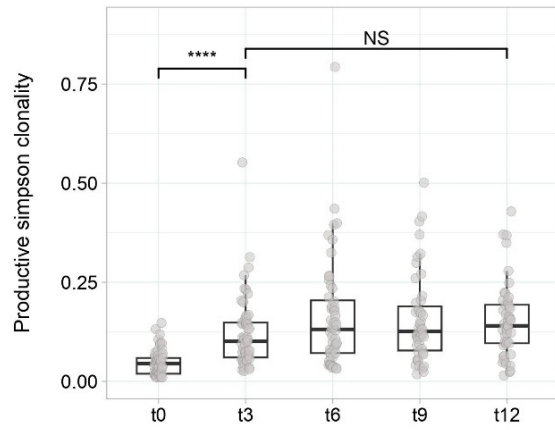
Wilcoxon rank sum test for pairwise comparisons in (B), (F). All p values were 2 sided. Statistical thresholds: (*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$ and (NS) not significant.

Fig. 3: Immune cell composition during the first-year post-HSCT (A) Relative proportions of six major cell populations: CD8⁺ and CD4⁺ T cells, natural killer T cell (NKT), natural killer cell (NK), monocytes and B cells assessed at indicated timepoints post-HSCT: t_0 ($n = 18$), t_3 ($n = 44$), t_6 ($n = 35$), t_9 ($n = 32$), t_{12} ($n = 37$). Each stacked bar represents an individual. **(B)** Evolution of the frequency of CD56^{bright} and CD56^{dim}, CD16⁺ NK cells at indicated timepoints post-HSCT: t_0 ($n = 18$), t_3 ($n = 44$), t_6 ($n = 35$), t_9 ($n = 32$), t_{12} ($n = 37$) fitted using LOESS regression with the 95% confidence interval. Day 0 indicates pre-HSCT state of the donor. Color-coded lines represent linear regression according to the cell subset. **(C)** Evolution of the frequency of CD56^{dim}, CD16⁺, NKG2A⁺, CD57^{neg} and CD56^{dim}, CD16⁺, NKG2A^{neg}, CD57⁺ NK cells at indicated timepoints post-HSCT: t_0 ($n = 18$), t_3 ($n = 44$), t_6 ($n = 35$), t_9 ($n = 32$), t_{12} ($n = 37$) fitted using LOESS regression with the 95% confidence interval. Day 0 indicate pre-HSCT state of the donor. Color-coded lines represent linear regression according to the cell subset. **(D) and (E)** Proportion of KIR⁺ CD56^{dim} NK cell at indicated timepoints post-HSCT: t_0 ($n = 18$), t_3 ($n = 44$), t_6 ($n = 35$), t_9 ($n = 32$), t_{12} ($n = 37$). Lines connect paired samples. Box plots display medians and interquartile ranges (IQR), with whiskers representing $1.5 \times$ IQR.

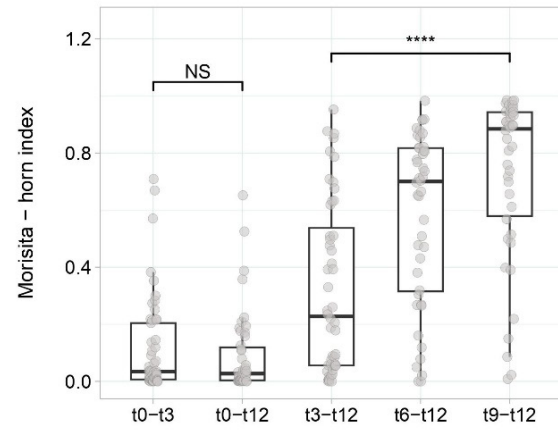
Fig. 4: NK cell compartment restoration in relation to CMV serostatus and infection/reactivation. Stratification according to following groups: seronegative ($n = 17$), seropositive ($n = 18$) and reactivated ($n = 12$) CMV recipients applies to plots (A) to (E). **(A)** Intra-individual distance (Aitchinson distance) between 3- and 12-months post-HSCT among the stratified groups based on 62 manually gated cell frequencies. **(B)** Evolution of the CD57⁺, NKG2C⁺ CD56^{dim} adaptive NK cells alongside one year-follow-up post-HSCT stratified among the groups fitted using LOESS regression with the 95% confidence interval. Color-coded lines represent linear regression according to the cell subset. **(C)** Heatmap of the frequency of cell subsets at 12 months that were statistically different between CMV seronegative (green) and CMV reactivated (blue) recipients. Heatmap rows display cell subsets and are colored by the Z-score normalized per row. Each column represents a recipient. **(D)** Granzyme B expression (MFI) in KIR2DL2L3S2⁺ NKG2C⁺ CD56^{dim} NK cells among the stratified groups at 3 and 12 months post-HSCT. **(E)** Frequency of CD107a⁺ CD56^{dim} NK cells in the stratified groups: seronegative ($n = 17$), seropositive ($n = 18$) and reactivated ($n = 12$). Box plots display medians and interquartile ranges (IQR),

with whiskers representing 1.5× IQR. Wilcoxon rank sum test in (A), (C), (D) and (E). All p values were 2 sided. Statistical thresholds: (*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$ and (NS) not significant.

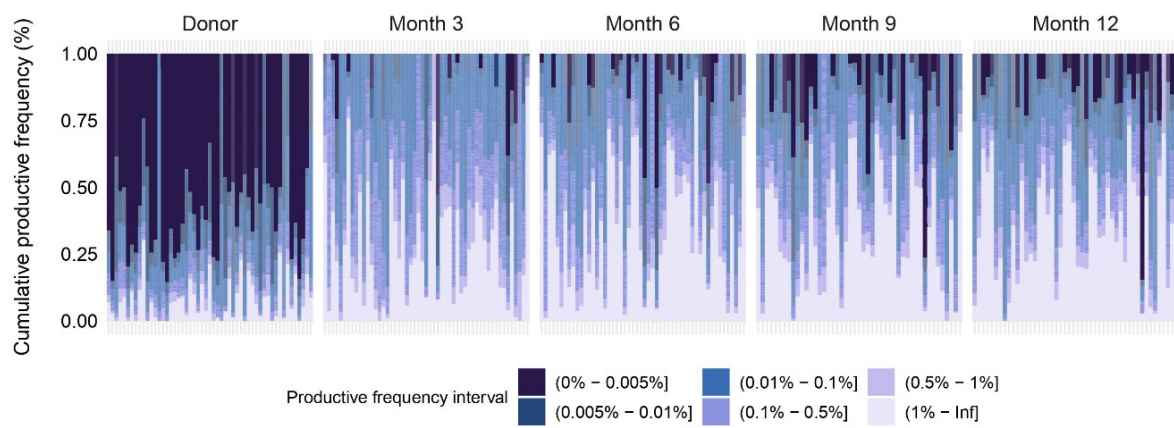
A

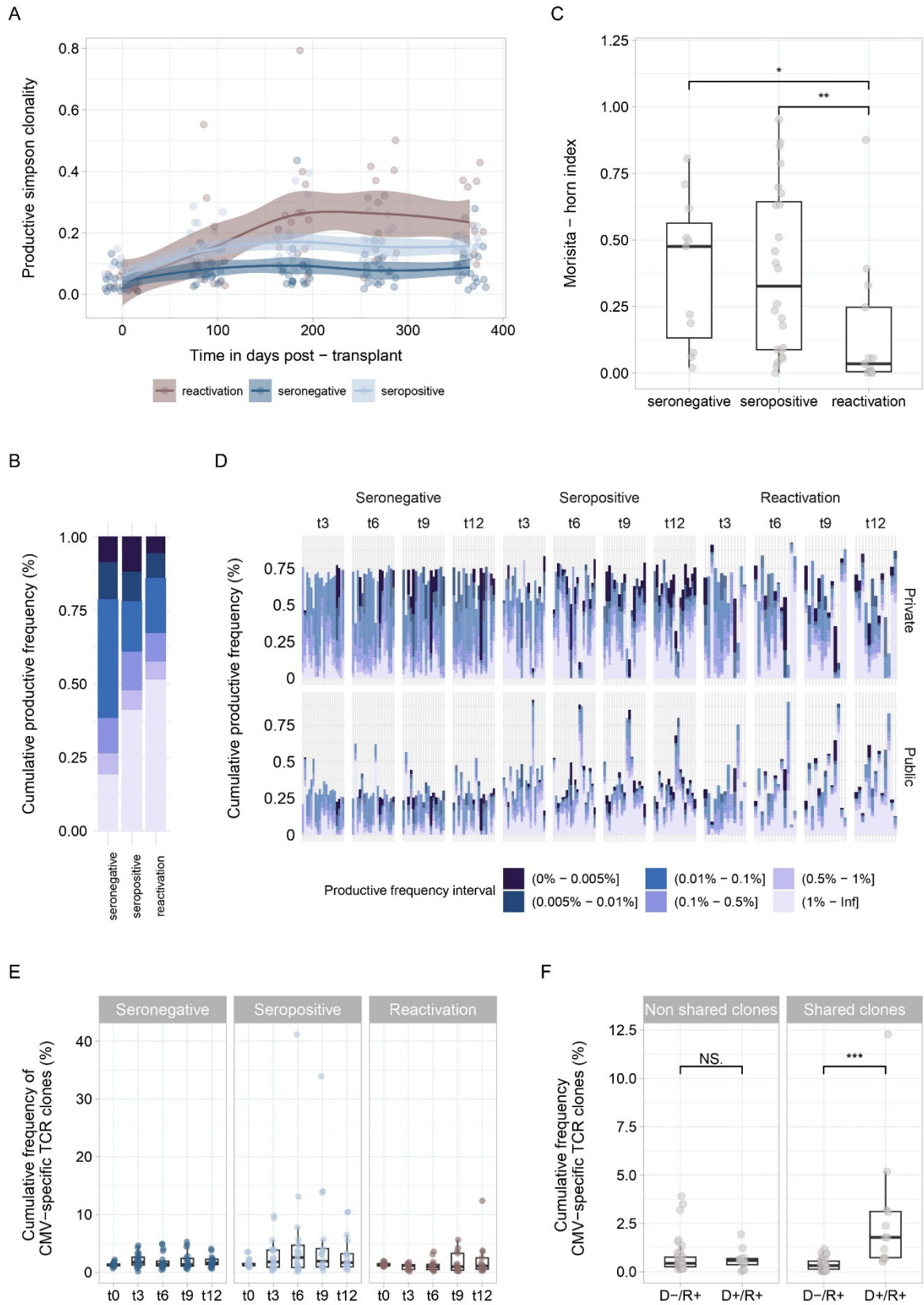


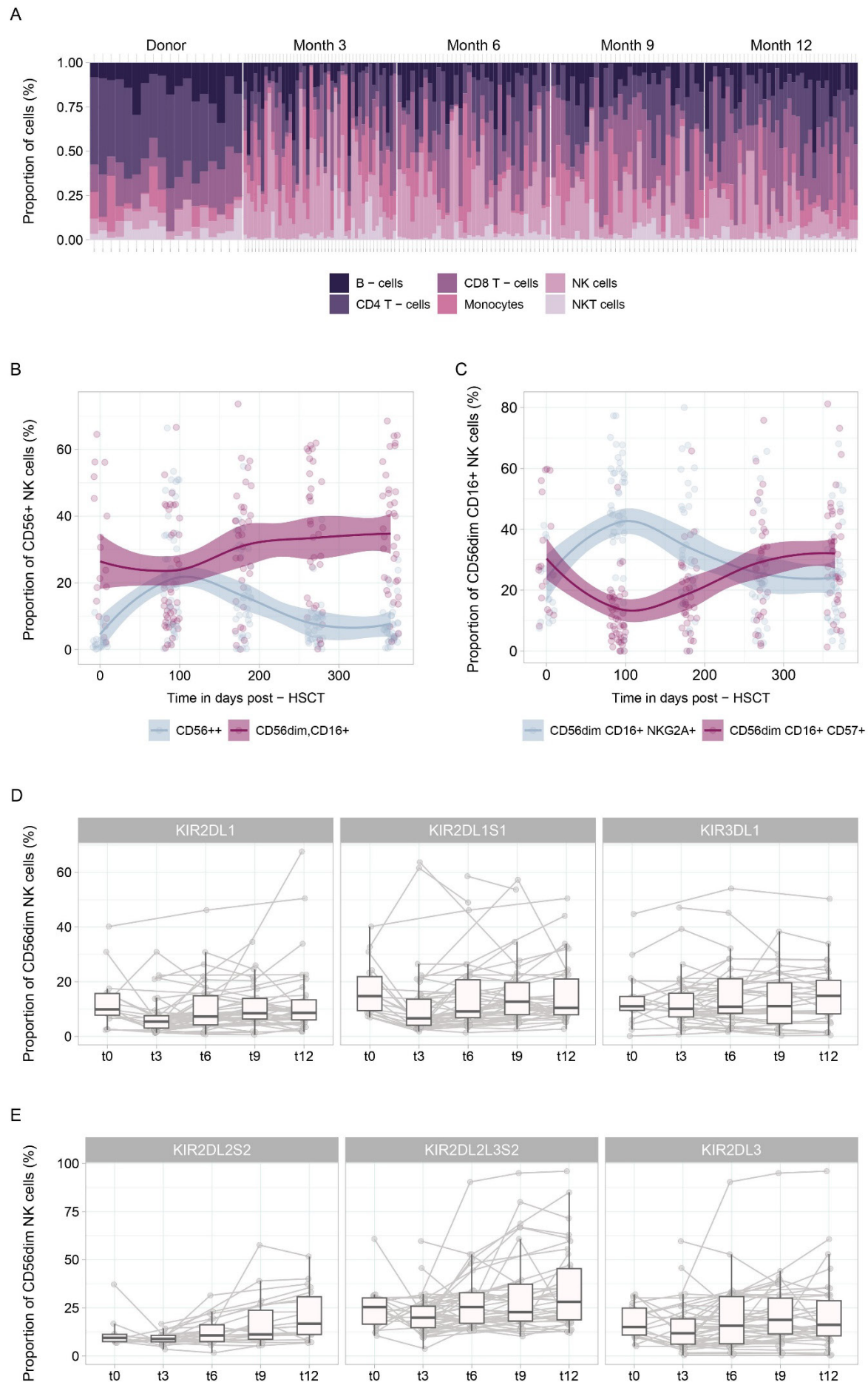
C



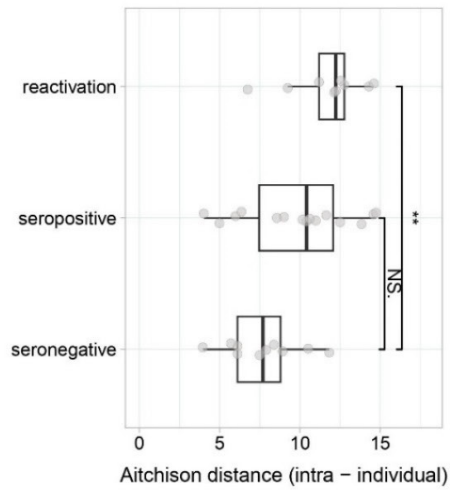
B



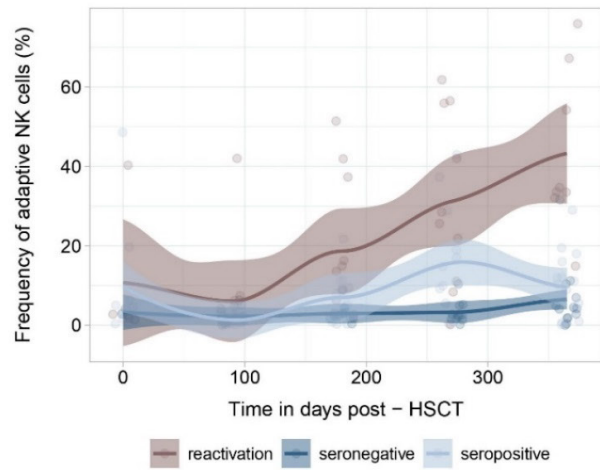




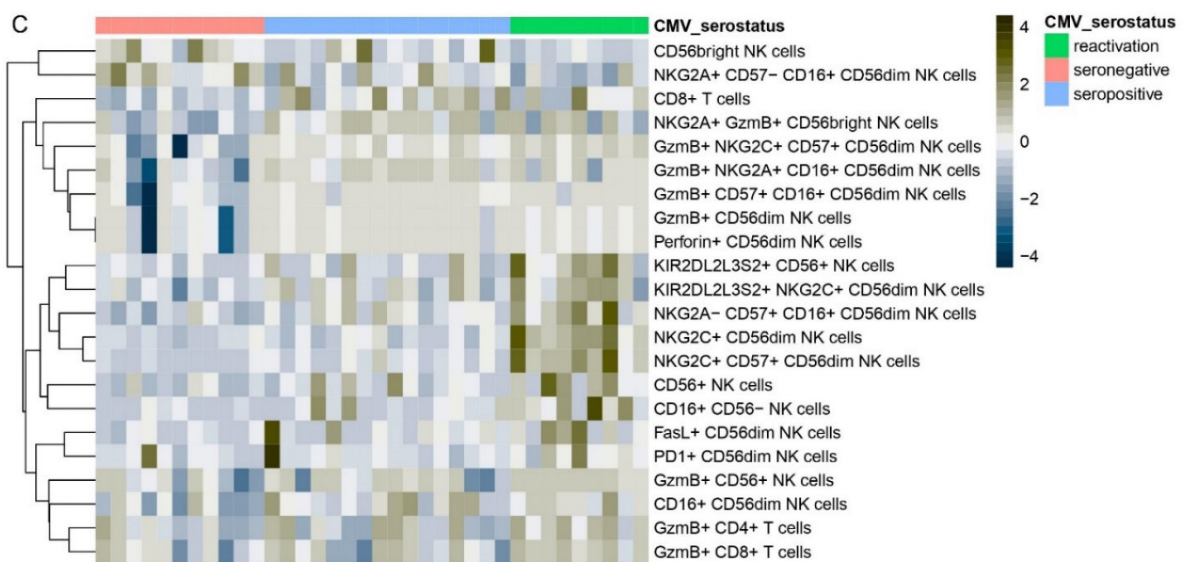
A



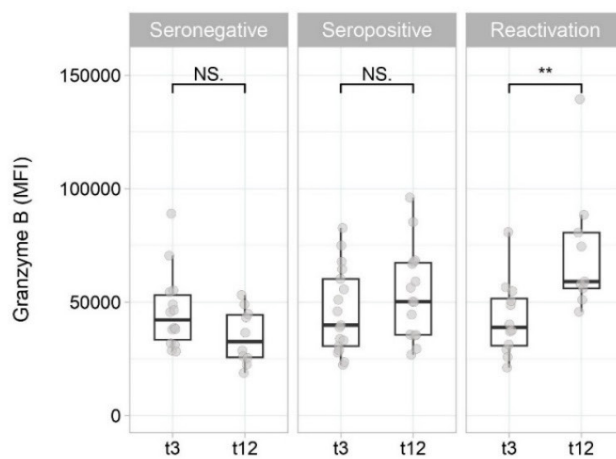
B



C



D



E

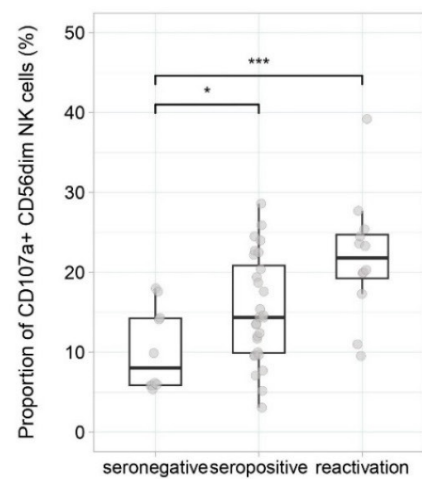


Table 1. Demographic and transplant related characteristics of the study cohort.

Parameter	All (<i>n</i> = 54)
Recipient's age at HSCT in yr (median, IQR)	57 (62)
Recipient's genre (M:F)	37 : 17
Donor's age in yr (median, IQR)	41 (21.5)
Donor's genre (M:F)	39 : 15
Underlying diagnosis, <i>n</i> (%)	
AML	27 (50)
ALL	4 (7.4)
AL non specific	2 (3.7)
CML /CLL	2 (3.7)
Lymphoma	4 (7.5)
Myeloma	1 (1.9)
MDPS /MDS /MPS	12 (22.2)
Hemoglobinopathy	2 (3.7)
Donor type, <i>n</i> (%)	
MUD	22 (40.7)
MMUD	2 (3.7)
MRD	9 (16.7)
Haplo-identical	21 (38.9)
First transplantation, <i>n</i> (%)	49 (90.7)
Conditioning with ATG, <i>n</i> (%)	39 (72.2)
PTCy, <i>n</i> (%)	22 (40.7)
No T-cell depletion, <i>n</i> (%)	53 (98.1)
Cryopreservation, <i>n</i> (%)	43 (79.6)
Stem cell source, <i>n</i> (%)	
BM	3 (5.5)
PBSC	51 (94.4)
CMV serostatus, <i>n</i> (%)	
D + /R +	22 (40.7)
D - /R +	12 (22.2)
D + /R -	7 (13)
D - /R -	13 (24.1)

Disease risk index, <i>n</i> (%)	
Low	4 (7.4)
Intermediate	41 (75.9)
High / very high	7 (13)
NA	2 (3.7)

ALL; acute lymphoblastic leukemic, AML; acute myeloid leukemia, ATG; anti-thymocyte globulin, BM; bone marrow, CLL; chronic lymphoblastic leukemia, CML; chronic myeloid leukemia, CMV; cytomegalovirus, IQR; interquartile range, MDS; myelodysplastic syndrome, MPS; myeloproliferative syndrome, MDPS; myelodysplastic/myeloproliferative syndrome, MMUD; mismatched unrelated donor, MRD; matched related donor, MUD; matched unrelated donor, PBSC; peripheral blood stem cells, PTCy; post-transplant cyclophosphamide

Table 2. Post-transplant clinical events and transplant outcomes of the study cohort.

Parameter	All (n = 54)
Immunogenic post-transplant complications	
aGvHD, n (%) *	40 (74.1)
> 1 episode within 1-year post-HSCT, n (%)	8 (14.8)
Severe aGvHD, grade $\geq$ III, n (%)	6 (11.1)
cGvHD, n (%)	13 (24.1)
> 1 episode within 1-year post-HSCT, n (%)	2 (3.7)
Moderate to severe cGvHD, n (%)	9 (16.7)
Infectious post-transplant complications	
Chronic viral infection, n (%)	
CMV	13 (24)
EBV	11 (20.4)
HHV-6	6 (11.1)
HHV-8	1 (1.9)
HSV	3 (5.6)
BKV	5 (9.3)
Respiratory viral infections, n (%)	
Influenza	5 (9.3)
Sars-Cov-2	16 (29.6)
RSV	4 (7.4)
Rhinovirus	5 (9.3)
Bacterial infections, n (%)	26 (48.2)
Transplant outcomes, n (%)	
Disease relapse	11 (20.4)
Overall survival at 1 year post-HSCT	50 (92.6)

aGvHD; acute graft-versus-host disease, BKV; polyomavirus BK, cGvHD; chronic graft-versus-host disease, CMV; cytomegalovirus, EBV; epstein-barr virus, HHV-6, human herpesvirus 6, HHV-8, human herpesvirus 6, HSV; herpes simplex virus, IQR; interquartile range, RSV; respiratory syncytial virus, Sars-Cov-2; severe acute respiratory syndrome coronavirus 2

*Classification of aGvHD grading according to the NIH criteria in which aGvHD can occur beyond 3 months based on clinical manifestations

Supplementary material

Antonia Schäfer¹, Zuleika Calderin Sollet¹, Marie-Priscille Hervé¹, Stéphane Buhler¹, Sylvie Ferrari-Lacraz¹, Paul J. Norman³, Katherine M. Kichula³, Ticiana D. J. Farias⁴, Stavroula Masouridi-Levrat², Anne-Claire Mamez², Amandine Pradier², Federico Simonetta², Yves Chalandon², Jean Villard^{1*}

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 *Department of Biomedical Informatics, University of Colorado School of Medicine, Aurora, CO, USA*

⁴ *Department of Biological Sciences, The University of North Carolina at Charlotte, Charlotte, NC, USA*

Corresponding author:

*Jean Villard, Transplantation Immunology Unit and National Reference Laboratory for
Histocompatibility, Geneva University Hospitals, Gabrielle-Perret-Gentil 4, 1211 Geneva 4, +41 22 372
93 94/ +41 79 553 34 09, jean.villard@hcuge.ch

Description of supplementary material

Supplementary materials and methods

1. Sample collection and processing
2. Bulk TCR CDR3 β sequencing
3. High resolution KIR genotyping
4. Spectral flow cytometry immunophenotyping
5. TCR data processing
6. Identification of public and private TCR clonotypes
7. Spectral flow cytometry data preprocessing and gating

Supplementary tables and figures

Table S1

Description: List of antibodies used for the full spectrum flow cytometry panel

Table S2

Description: Demographic and transplant related characteristics of the study cohort stratified according to the CMV groups (seronegative, seropositive and reactivated)

Table S3

Description: Post-transplant clinical events and transplant outcomes of the study cohort stratified according to the CMV groups (seronegative, seropositive and reactivated)

Table S4

Description: HLA restriction and peptide information related to the CMV specific TCR clonotypes

Table S5

Description: Association between TCR/NK cell repertoire reconstitution and different potential modulatory events

Table S6

Description: Absolute counts of NK and CD4⁺ / CD8⁺ T cells in the study cohort ($n = 47$)

Figure S1

- (A) Description: Study design overview.
- (B) Description: Swimmer plots detailing the patients' clinical course post-HSCT.

Figure S2

- (A) Description: Gating strategy for cleaning and identification of viable CD45⁺ immune cells.
- (B) Description: Gating hierarchy for identification of major immune cell subsets.

Figure S3

Description: Cumulative productive frequency of pathogen-specific TCR clones identified by *in silico* matching with public databases at serial timepoints post-HSCT.

Figure S4

- (A) Description: Cumulative productive frequency of donor-recipient non-shared and shared anti-CMV specific TCR clones in matched and mismatched CMV reactivated recipients at indicated timepoints post-HSCT.
- (B) Description: Cumulative productive frequency of donor-recipient non-shared and shared anti-CMV specific TCR clones according to the CMV seropositivity/reactivation.

Figure S5

- (A) Description: Proportion of NKp44⁺, NKp30⁺, NKp46⁺, TRAIL⁺, FasL⁺, NKG2D⁺, PD-1⁺, TIGIT⁺ CD56^{dim} cell subsets at indicated timepoints post-HSCT.
- (B) Description: Proportion of KIR2DL5⁺, KIR2DS1⁺, KIR2DS4⁺, KIR3DS1⁺ CD56^{dim} cell subsets at indicated timepoints post-HSCT.

Figure S6

- (A) Description: Representative single-cell plots displaying CD56 and CD107a expression frequency according to stimulation.
- (B) Description: Frequency of CD107a⁺ CD56⁺ NK cells in unstimulated and stimulated cells.

Figure S7

- (A) Description: CMV peak viral loads in CMV reactivated recipients in relation to the productive Simpson clonality of the TCR repertoire.
- (B) Description: CMV peak viral loads in CMV reactivated recipients in relation to the frequency of CD57⁺ NKG2C⁺ CD56^{dim} NK cells at 12 months post-alloHSCT.

Figure S8

Description: Immune reconstitution of representative CMV seronegative (a.) and CMV reactivated (b.), (c.) recipients.

Supplementary materials and methods

1. Sample collection and processing

Peripheral blood mononuclear cell (PBMC) isolation was performed by Ficoll density gradient centrifugation. Briefly, whole blood was first diluted with the equal amount of Phosphate-buffered saline (PBS), layered onto the Ficoll-Paque (Gibco Life Technologies, Oslo, Norway) and centrifuged at 800 x *g* for 20-25 min. Subsequently, PBMCs were isolated from the interface, washed twice with PBS and cryopreserved in fetal bovine serum supplemented with 10% dimethyl sulfoxide (Merck, Darmstadt, Germany). Samples were stored in liquid nitrogen until use. Genomic DNA was extracted from PBMCs using NucleoSpin kit (Machery-Nagel, Düren, Germany) or directly extracted from the whole blood and purified using the EZ1 Advanced XL machine (QIAGEN, GmbH, Hilden, Germany). All samples were stored at 4°C until usage.

2. Bulk TCR CDR3 β sequencing

DNA quality and concentration were assessed using *dsDNA* Qubit kit. The starting DNA input concentration for each sample was between 30 and 44 ng/ μ l. High-throughput TCR complementary determining region 3 beta (CDR3 β) sequencing at survey resolution was performed using the immunoSEQ® Assay (Adaptive Biotechnologies, Seattle, WA, USA) according to the manufacturer's instructions. Briefly, genomic DNA was amplified by a multiplex PCR reaction using forward and reverse amplification primers targeting all V and J gene segments. TCR pooled libraries were sequenced on an Illumina MiSeq instrument using v3 150 cycle kit chemistry (Illumina, San Diego, CA, USA).

3. High resolution KIR genotyping

DNA purity and concentration were assessed by Nanodrop 200 spectrophotometer. For library preparation, 500 ng of genomic DNA was first fragmented by digestive enzymes (New England Biolabs, Boston, MA, USA) followed by barcode ligation with unique adaptors. After postligation cleanup, dual size selection was performed with AMPure magnetic beads (Beckmann Coulter, Brea, CA, USA) to acquire fragment sizes of 800 to 1200 bp length. In a second step, a pool of oligonucleotide probes specific for KIR and HLA genomic region was used for the targeted capture. Final enriched libraries were normalized to a concentration of 12 pmol/l. Paired-end sequencing was performed using NovaSeq instrument with a sequencing length of 2 x 250 bp (Illumina, San Diego, CA, USA).

4. Spectral flow cytometry immunophenotyping

The antibody panel for the immunophenotyping comprised a total of 35 markers. Antibody clones, concentration, corresponding fluorochromes and supplier are summarized in supplementary Table S1. Antibody master mix was prepared fresh before each staining experiment. Cryopreserved samples were quickly thawed in a water-bath at 37°C and washed twice with warm RPMI-1640 (Gibco, Life Technologies, Oslo Norway) supplemented with 10% fetal bovine serum and – if required – with Pierce Universal Cell Nuclease (Thermo Scientific™, Waltham, MA, USA). Cells were subsequently resuspended in RPMI-1640 with 10% fetal bovine serum and rested overnight in a 37°C, 5 % CO₂ incubator to be revitalized. After overnight resting, samples were washed with PBS by centrifugation at 400 x g for 5 min. Viability staining was performed by adding 5 µl of a 1:500 diluted ViaDye Red viability staining solution (Cytek® Biosciences, Fremont, CA, USA) to the cells and incubated for 20 min at room temperature. Cells were subsequently washed with cell staining buffer (CSB) containing 1% BSA, 0.5 M EDTA and 0.02% NaNH₃ in PBS. For extracellular staining, Fc receptors were first blocked by adding 5 µl of Fc-receptor blocking (BioLegend, Fell, Germany) solution to each sample followed by 10 min incubation. Afterwards, 80 µl of antibody master mix were added to the cells and incubated for an additional 20 min at 4°C protected from the light. Following two washes with 2 ml of CSB, cells were fixed and permeabilized for the intracellular staining: to this end, cells were resuspended in 200 µl of Fix/Perm FoxP3 Solution (Foxp3/Transcription Factor Staining Buffer Set, eBiosciences™, San Diego, CA, USA) and left for 20 min at room temperature. Cells were then washed twice by pelleting with FoxP3 permeabilization buffer at 800 x g. After fixation and permeabilization, cells were resuspended in 50 µL of cytoplasmic /intracellular antibody cocktail and incubated for another 20 min at 4°C protected from the light. Cells were then washed twice by pelleting with CSB and resuspended in 400 µl CSB. On the day of acquisition, cells were finally filtered through a 35 µm nylon mesh filter. Samples were acquired on a 5-Laser Aurora system (Cytek® Biosciences, Fremont, CA, USA) using the SpectroFlo® Software v3.1.0. The instrument QC was done on a daily basis using SpectroFlo® QC Beads (Lot 2004) following manufacturer recommendations. Samples were acquired at a high flow rate of ~ 50 µl / sec. For compensation, cells were unmixed using the stored reference controls with autofluorescence extraction option selected.

5. TCR data processing

Raw data processing and sequencing analyses were performed using the immunoSEQ Analyzer 4.0 online platform (<http://www.adaptivebiotech.com/immunoseq>) and in R environment (version 4.1.3, R Core team, R foundation for statistical computing). At first, raw sequencing datasets were cleaned as follows: (i) According to the criteria of the International ImMunoGeneTics Collaboration, the TCR β CDR3 region starts at the second conserved cysteine encoded by the 3' position of the V β gene segment and ends with the conserved phenylalanine encoded by the 5' position of the J β gene segment. Thus, all sequences matching this requirement were kept while all other sequences generating non-productive events were discarded. (ii) We collapsed sequences with the same CDR3 composition (including sequences with different V-D-J rearrangements) into the same clonotype and computed for each clonotype, the sum of the productive frequencies of all the different DNA sequences coding for that same CDR3.

The Simpson clonality was computed as a TCR repertoire diversity marker as follows: *Simpson index* = $1 - \lambda = 1 - \sum_{i=1}^N p_i^2$. Output values range from 0 to 1, with values tending to 1 indicating an uneven with dominant clones distributed repertoire and inversely values tending to 0 a more even polyclonal repertoire. The degree of overlap in terms of TCR clone identity and abundance was computed using the Morisita-Horn index, mathematically defined as follows: *Morisita – Horn index* = $MHI(i) = \frac{2\sum(x_i y_i)}{(D_x + D_y)XY}$.

6. Identification of public and private TCR clonotypes

CDR3aa sequences were used to assess TCR clones publicness. Public TCR CDR3 β clonotypes were either defined as matching clonotypes present in public databases with antigen-specific validated TCR (section 2.8.3) or being present in ≥ 2 individuals in our cohort. Identification of identical clonotypes within a D/R pair was considered private unless the requirement of full chimerism was not set.

7. Spectral flow cytometry data preprocessing and gating

FCS3.0 files were extracted from SpectroFlo® Software v3.1.0 after spectral unmixing and compensation adjustments. For further downstream analysis, non-randomized data were uploaded in FlowJo v10.7.2 software (Tree Star, Ashland, OR, USA): (i) the datasets were checked whether they follow appropriate quality standards using the FlowAI plugin. (ii) the datasets were then manually pre-

gated on single, alive, CD45⁺ immune cells (Supplementary Fig. 2A). (iii) Cell populations of interest were further identified via manual gating (Supplementary Fig. 2B).

Supplementary table S1. List of antibodies used for the full spectrum flow cytometry panel.

Target	Fluorochrome	Antibody clone	Provider	ug/test	Staining step
CD45	cFluor V547	HI30	Cytek custom	0.2	Surface
CD3	APC/Fire 810	SK7	Biolegend	0.1	Surface
CD4	PE/Fire 640	SK3	Biolegend	0.1	Surface
CD8a	BV570	RPA-T8	Biolegend	0.4	Surface
CD19	BV750	HIB19	Biolegend	0.4	Surface
CD14	BUV805	M5E2	BD biosciences	0.2	Surface
CD56	BV785	5.1H11	Biolegend	0.05	Surface
CD16	BUV496	3G8	BD biosciences	0.05	Surface
CD57	BV605	QA17A04	Biolegend	0.4	Surface
NKG2A	PE-Cy5	S19004C	Biolegend	0.4	Surface
NKG2C	BUV737	134591	BD biosciences	0.2	Surface
NKG2D	cFluor BYG710	1D11	Cytek custom	2.5	Surface
NKp30	BUV563	P30-15	BD biosciences	1	Surface
NKp44	BUV615	P44-8	BD biosciences	1	Surface
NKp46	BV711	9E2	BD biosciences	1	Surface
KIR2DL1	PE-Vio 770	REA284	Miltenyi	1	Surface
KIR2DL1/ S1	APC-Vio 770	REA1010	Miltenyi	1	Surface
KIR2DL2/ L3 /S2	BB700	CH-L	BD biosciences	0.4	Surface
KIR2DL3	APC	REA147	Miltenyi	2.5	Surface
KIR2DL4	PE	181703	R&D Systems	2.5	Surface
KIR2DL5	BV421	UPR1	BD biosciences	1	Surface
KIR2DS4	PE – Vio 615	REA860	Miltenyi	2.5	Surface
KIR3DL1	VioGreen	REA1005	Miltenyi	1	Surface
KIR3DL2	Alexa Fluor 647	539304	R&D Systems	2.5	Surface
KIR3DL3	Alexa Fluor 700	1136B	R&D Systems	2.5	Surface
KIR3DL1 / S1	FITC	REA168	Miltenyi	1	Surface
CD107a	BUV395	H4A3	BD biosciences	0.05	Intra – cellular
TNFa	BV650	Mab11	Biolegend	1	Intra – cellular
PD-1	BUV661	EH12.1	BD biosciences	0.4	Surface
TIGIT	cFluor BYG750	A15153G	Cytek custom	1	Surface
TRAIL	VioBright B515	REA1113	Miltenyi	2.5	Surface
Fas-L	BV480	NOK1	BD biosciences	1	Surface
Granzyme B	Pacific Blue	GB11	Biolegend	0.2	Intra – cellular
Perforin	cFluor R685	DG9	Cytek custom	1	Intra – cellular
Viability	ViaDyeRed		Cytek custom		Surface

Table S2. Demographic and transplant related characteristics of the study cohort stratified according to CMV the groups (seronegative, seropositive and reactivated).

Parameter	All (n = 54)	CMV seronegative n = 20	CMV seropositive n = 20	CMV reactivation (D-/R+) n = 9	CMV reactivation (D+/R+) n = 4
Recipient age at HSCT in yr (median, IQR)	57 (62)	57 (10)	54 (18)	63 (12)	54.5 (33.5)
Recipient genre (M:F)	37 : 17	15:5	13:8	2:2	7:2
Donor age in yr (median, IQR)	41 (21.5)	38.5 (26.3)	51 (19)	38 (19)	40.5 (5.5)
Donor genre (M:F)	39 : 15	15:5	14:7	7:2	3:1
Underlying diagnosis, n (%)					
AML	27 (50)	8 (40)	10 (47.6)	7 (77.8)	2 (50)
ALL	4 (7.4)	3 (15)	1 (4.8)	-	-
AL non specific	2 (3.7)	1 (5)	1 (4.8)	-	-
CML /CLL	2 (3.7)	-	2 (9.5)	-	-
Lymphoma	4 (7.5)	1 (5)	1 (4.8)	2 (22.2)	-
Myeloma	1 (1.9)	1 (5)	-	-	-
MDPS /MDS /MPS	12 (22.2)	6 (30)	4 (19)	-	2 (50)
Hemoglobinopathy	2 (3.7)	-	2 (9.5)	-	-
Donor type, n (%)					
MUD	22 (40.7)	11 (55)	9 (42.9)	1 (11.1)	1 (25)
MMUD	2 (3.7)	1 (5)	1 (4.8)	-	-
MRD	9 (16.7)	5 (25)	4 (19)	-	-
Haplo-identical	21 (38.9)	3 (15)	7 (33.3)	8 (88.9)	3 (75)
First transplantation, n (%)	49 (90.7)	18 (90)	19 (90.5)	8 (88.9)	4 (100)
Conditioning with ATG, n (%)	39 (72.2)	15 (75)	15 (71.4)	7 (77.8)	2 (50)
PTCy, n (%)	22 (40.7)	4 (20)	8 (38.1)	8 (88.9)	2 (50)
No T-cell depletion, n (%)	53 (98.1)	20 (100)	20 (95.2)	9 (100)	4 (100)
Cryopreservation, n (%)	43 (79.6)	19 (95)	14 (66.7)	6 (66.7)	4 (100)
Stem cell source, n (%)					
BM	3 (5.5)	-	3 (14.3)	-	-
PBSC	51 (94.4)	20 (100)	18 (85.7)	9 (100)	4 (100=

CMV serostatus, <i>n</i> (%)					
D + /R +	22 (40.7)	-	18 (85.7)	-	4 (100)
D - /R +	12 (22.2)	-	3 (14.3)	9 (100)	-
D + /R -	7 (13)	7 (35)	-	-	-
D - /R -	13 (24.1)	13 (65)	-	-	-
Disease risk index, <i>n</i> (%)					
Low	4 (7.4)	-	4 (19)	-	-
Intermediate	41 (75.9)	17 (85)	13 (61.9)	7 (77.8)	4 (100)
High / very high	7 (13)	3 (15)	2 (9.5)	2 (22.2)	-
NA	2 (3.7)	-	2 (9.5)	-	-

ALL; acute lymphoblastic leukemic, AML; acute myeloid leukemia, ATG; anti-thymocyte globulin, BM; bone marrow, CLL; chronic lymphoblastic leukemia, CML; chronic myeloid leukemia, CMV; cytomegalovirus, IQR; interquartile range, MDS; myelodysplastic syndrome, MPS; myeloproliferative syndrome, MDPS; myelodysplastic/myeloproliferative syndrome, MMUD; mismatched unrelated donor, MRD; matched related donor, MUD; matched unrelated donor, PBSC; peripheral blood stem cells, PTCy; post-transplant cyclophosphamide

Table S3. Post-transplant clinical events and transplant outcomes of the study cohort stratified according to the CMV groups (seronegative, seropositive and reactivated).

Parameter	All (n = 54)	CMV seronegative n = 20	CMV seropositive n = 21	CMV reactivation (D-/R+) n = 9	CMV reactivation (D+/R+) n = 4
Immunogenic post-transplant complications					
aGvHD, n (%)	40 (74.1)	15 (75)	14 (66.7)	8 (88.9)	3 (75)
> 1 episode within 1-year post-HSCT, n (%)	8 (14.8)	4 (20)	2 (9.5)	2 (22.2)	-
Severe aGvHD, grade $\geq$ III, n (%)	6 (11.1)	3 (15)	2 (9.5)	-	1 (25)
cGvHD, n (%)	13 (24.1)	6 (30)	4 (19)	2 (22.2)	1 (25)
> 1 episode within 1-year post-HSCT, n (%)	2 (3.7)	1 (5)	-	-	1 (25)
Moderate to severe cGvHD, n (%)	9 (16.7)	4 (20)	3 (14.3)	1 (11.1)	1 (25)
Infectious post-transplant complications					
Chronic viral infection, n (%)					
CMV	13 (24)	-	-	9 (100)	4 (100)
EBV	11 (20.4)	2 (10)	6 (28.6)	1 (11.1)	2 (50)
HHV-6	6 (11.1)	2 (10)	3 (14.3)	1 (11.1)	-
HHV-8	1 (1.9)	-	1 (4.7)	-	-
HSV	3 (5.6)	1 (5)	1 (4.7)	-	1 (25)
BKV	5 (9.3)	-	3 (14.3)	2 (22.2)	-
Respiratory viral infections, n (%)					
Influenza	5 (9.3)	1 (5)	1 (4.7)	3 (33.3)	-
Sars-Cov-2	16 (29.6)	5 (25)	8 (38.1)	1 (11.1)	2 (50)
RSV	4 (7.4)	2 (10)	2 (9.5)	-	-
Rhinovirus	5 (9.3)	2 (10)	2 (9.5)	1 (11.1)	-
Bacterial infections, n (%)	26 (48.2)	8 (40)	10 (47.6)	6 (66.7)	2 (50)
Transplant outcomes, n (%)					
Disease relapse	11 (20.4)	4 (20)	4 (19)	3 (33.4)	-
Overall survival at 1 year post-HSCT	50 (92.6)	19 (95)	18 (85.7)	9 (100)	4 (100)

Table S4. HLA restriction and peptide information related to the CMV specific TCR clonotypes.

HLA	Peptide	<i>n</i>	frequency
HLA-A*03:01	KLGGALQAK	12657	0,632280947
HLA-A*02	NLVPMVATV	4048	0,202218004
HLA-A*02:01	NLVPMVATV	748	0,03736637
HLA-C*07:02	CRVLCCYVL	430	0,021480667
HLA-B*07:02	TPRVTGGGAM	421	0,021031072
HLA-DRA1*01	LLQTGIHVRVSQPSL	302	0,015086422
HLA-C*07:02	FRCPRRFCF	263	0,013138176
HLA-B*07:02	RPHERNGFTVL	224	0,011189929
HLA-A*01:01	VTEHDTLLY	194	0,009691278
HLA-B*35:01	IPSINVHHY	93	0,004645819
HLA-A*01:01	YSEHPTFTSQY	80	0,003996403
HLA-B*44	NEGVKAAW	76	0,003796583
HLA-A*02	MLNIPSINV	72	0,003596763
HLA-A1	VTEHDTLLY	69	0,003446898
HLA-A*24:02	QYDPVAALF	43	0,002148067
HLA-B*44:03:08	NEGVKAAW	40	0,001998202
HLA-B*08:01	QIKVRVKMV	35	0,001748426
HLA-A*24:02	AYAQKIFKI	32	0,001598561
HLA-DRA1*01	EHPTFTSQYRIQGKL	29	0,001448696
HLA-A*02:01	YSEHPTFTSQY	22	0,001099011
HLA-B*07	TPRVTGGGAM	19	0,000949146
HLA-B*08	QIKVRVDMV	14	0,000699371
HLA-B*08:01	ELRRKMMYM	13	0,000649416
HLA-A*02:01	VLEETSVML	10	0,00049955
HLA-B*35	IPSINVHHY	10	0,00049955
HLA-B*35:08	FPTKDVAL	10	0,00049955
HLA-B*08:01	QIKVRVDMV	9	0,000449595
HLA-B7	TPRVTGGGAM	9	0,000449595
HLA-A*02	VLEETSVML	6	0,00029973
HLA-A2	NLVPMVATV	5	0,000249775
HLA-B*18	ELKRKMIYM	5	0,000249775
HLA-A*02:01	ARNLVPMVATVQGQN	3	0,000149865
HLA-A*02:01	YILEETSVM	3	0,000149865
HLA-B*08	ELRRKMMYM	3	0,000149865
HLA-A*01	VTEHDTLLY	2	9,99101E-05
HLA-A*02:01	LSEFCRVLCCYVLEE	2	9,99101E-05
HLA-B*07	RPHERNGFTVL	2	9,99101E-05
HLA-B*18	CVETMCNEY	2	9,99101E-05
HLA-B*18	DEEDAIAAY	2	9,99101E-05
HLA-B*35:08	CPSQEPMSIYVY	2	9,99101E-05
HLA-E*01:01:01:03	VMAPRTLIL	2	9,99101E-05
HLA-A*01	VLEETSVML	1	4,9955E-05

HLA-A*01	YSEHPTFTSQY	1	4,9955E-05
HLA-A*02:01	EDVPSGKLFMHVTLG	1	4,9955E-05
HLA-A*02:01:98	NLVPMVATV	1	4,9955E-05
HLA-B*07	RIPHERNGFTV	1	4,9955E-05
HLA-B*12	EFFWDANDIY	1	4,9955E-05
HLA-B*35:42:01	IPSINVHHY	1	4,9955E-05

Table S5. Association between TCR/NK cell repertoire reconstitution and different potential modulatory events within the study cohort ($n = 54$).

Parameter	PTCy	aGvHD/cGvHD	Sars-Cov-2
	$n = 22$ with PTCy $n = 32$ without PTCy	$n = 43$ with GvHD $n = 11$ without GvHD	$n = 16$ with PTCy $n = 38$ without PTCy
TCR clonality	No significant difference between the groups except at T6 with a higher clonality in PTCy treated recipients ($p = 0.043$). Between 3 and 12 months, stable TCR clonality ($p = 0.29$).	No significant difference between these different groups ($p > 0.05$). Between 3 and 12 months, significant increase in recipients with GvHD ($p = 0.0275$).	No significant difference between the groups and between 3 and 12 months ($p > 0.05$).
TCR overlap	Significant reduced TCR overlap in PTCy treated recipients early post-HSCT (t_0 - t_3 , $p = 0.049$; t_3 - t_6 , $p = 0.049$). At latter timepoints, no significant difference was noted between the groups (t_3 - t_{12} , $p = 0.39$, t_9 - t_{12} , $p = 0.75$).	No significant difference between the groups ($p > 0.05$).	No significant difference between the groups ($p > 0.05$).
NK/T cell subsets	Significant higher frequency of CD8 ⁺ T cells at 12 months in PTCy treated recipients ($p = 0.015$). Significant reduced proportion of CD56 ^{bright} NK cells in PTCy treated recipients ($p = 0.02$).	Significant reduced proportion of CD16 ⁺ CD56 ^{dim} NK cells at 12 months in GvHD recipients ($p = 0.0067$). Significant increase in monocytes at 6 ($p = 0.045$) and 12 months in GvHD recipients ($p = 0.0093$).	Significant reduced frequency of NK cells at 12 months in Sars-Cov-2 infected recipients ($p = 0.0485$).

aGvHD; acute graft-versus-host disease, cGvHD; chronic graft-versus-host disease, PTCy; post-transplant cyclophosphamide, Sars-Cov-2; severe acute respiratory syndrome coronavirus 2

Table S6. Absolute counts of NK and CD4⁺ / CD8⁺ T cells in the study cohort (*n* = 47).

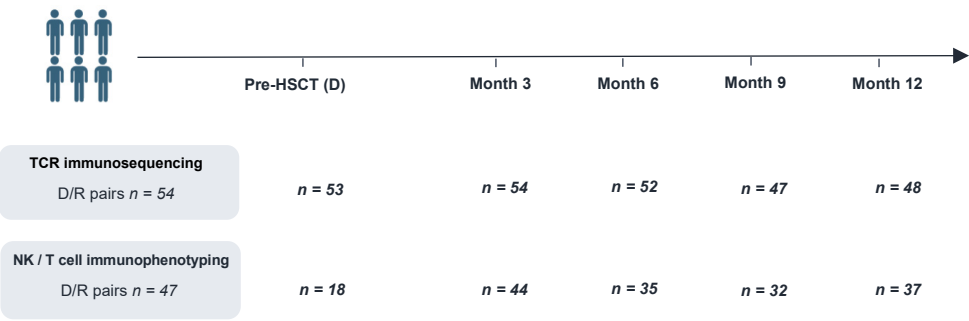
Sample	Timepoint	CD3+ cells	CD4 T - cells	CD8 T - cells	NK cells
p01	T03	263094	62020	109808	227643
p01	T06	324462	86547	210836	215048
p01	T09	95543	17822	74315	48434
p01	T0	422729	309570	95769	54294
p02	T03	91848	65370	20393	103544
p02	T06	89754	68221	19287	88844
p02	T12	113093	59593	32259	28934
p02	T0	807453	542799	233852	48206
p03	T03	122230	21728	43086	25747
p03	T06	90670	16723	64314	62911
p03	T09	107064	21281	77442	70971
p04	T03	75953	39164	24037	33431
p04	T06	287193	60500	193684	62014
p04	T09	383971	107038	191291	40555
p05	T03	5489	3927	1458	4539
p05	T06	5751	2942	2575	5776
p05	T09	7162	3320	3479	3689
p07	T03	165749	159809	4628	52397
p07	T06	157909	103023	13404	179783
p07	T09	261847	180679	40901	101134
p07	T12	144689	94466	22791	58976
p07	T0	400697	302317	67540	23531
p09	T03	76082	18150	47572	117489
p09	T06	89972	24354	47722	132241
p09	T12	75650	19631	48971	83250
p10	T03	4051	811	1486	50527
p10	T06	81425	59793	16424	57699
p10	T12	5857	3871	1031	1703
p11	T03	21597	2383	11106	13695
p11	T06	106398	41918	43788	56447
p11	T09	93065	42748	29285	59268
p11	T12	225686	90311	86755	76215
p12	T03	39022	8129	16105	27926
p12	T06	31096	11441	11726	2727
p12	T12	21575	2784	14589	4186
p13	T03	432997	82727	321251	42994
p13	T06	473462	114716	328086	73480
p13	T09	45564	26968	9828	22047
p13	T12	418132	137397	253077	67192
p13	T0	294424	179721	89258	34084
p14	T03	23477	18902	3333	25880
p14	T06	340769	88474	222208	82304

p14	T12	19383	3893	14547	2219
p17	T03	98094	60693	23446	183950
p17	T06	117668	23547	18185	530672
p17	T09	63646	48651	10420	112683
p17	T12	142149	76990	33095	219829
p18	T06	3138	2605	112	1670
p18	T09	8415	6080	360	9173
p19	T03	71375	29980	35686	34045
p19	T09	252263	37774	183117	61218
p19	T12	182217	39686	128158	20714
p19	T0	234019	178033	44889	24114
p20	T03	210766	69139	87892	40218
p20	T06	396341	72192	189134	45924
p20	T09	248377	44514	110125	11546
p20	T12	427469	75637	235740	62850
p20	T0	352118	194778	101113	63124
p21	T03	62973	61146	908	28168
p21	T09	172986	54559	101058	14875
p21	T12	49047	27824	13881	48400
p21	T0	361461	216965	94604	101416
p22	T03	12862	3114	5008	45964
p22	T12	44505	11274	21994	7564
p23	T03	73563	38268	28304	49289
p23	T06	116616	84230	27922	15440
p23	T09	119487	75777	39339	7251
p23	T12	87892	59013	21815	9472
p23	T0	160169	93437	33874	20612
p24	T03	77860	43099	23167	23918
p24	T09	188904	44935	101407	136110
p24	T12	202856	69889	105677	142659
p25	T03	2679	720	1312	118280
p25	T06	266787	15082	187193	51196
p25	T09	460331	50207	333912	78266
p25	T12	378281	58136	272489	17123
p25	T0	545673	360047	133792	64103
p27	T03	23351	8286	11656	88239
p27	T06	137555	51599	40647	98126
p27	T09	160276	58344	79405	113573
p27	T12	40213	13910	20983	18077
p28	T03	127161	30231	78210	65526
p28	T12	287788	93382	157991	43985
p30	T03	55721	12597	30820	46796
p30	T12	203548	86528	92954	20768
p32	T03	10022	5967	403	4089
p32	T06	1393	680	130	655
p32	T0	63098	39750	18412	3318
p34	T03	19853	12854	6085	214031

p34	T06	18180	8807	6764	28694
p34	T09	101029	38327	47488	64041
p34	T12	225987	81963	118721	175827
p34	T0	202960	134161	43271	37639
p35	T03	327128	84631	212950	20589
p35	T09	514120	93212	382631	26047
p35	T12	462108	97810	339818	36455
p35	T0	332195	194985	61686	75072
p36	T03	4400	3134	618	59987
p36	T09	88182	24086	55111	118589
p36	T12	117559	47844	55994	80544
p36	T0	246898	188586	50885	25662
p37	T03	246845	64689	105435	60467
p37	T06	177777	50295	92768	19616
p37	T09	348585	95503	200902	61437
p37	T12	376810	118873	214527	74913
p37	T0	215739	137749	57752	31961
p38	T03	20162	10160	5793	96279
p38	T06	237302	42008	100008	38563
p38	T09	132948	31441	68944	12770
p39	T03	22108	18701	2024	8310
p39	T09	3216	1960	662	503
p39	T12	159401	114539	38809	13337
p40	T03	4580	2271	962	528
p40	T06	71010	34601	25318	2166
p40	T12	387881	185000	149620	25708
p41	T03	9453	1527	2952	7248
p41	T06	258847	43006	175467	83339
p41	T12	249514	90242	134902	27181
p42	T06	13150	4801	6109	16498
p42	T12	49698	23454	21400	26034
p42	T0	55862	28961	22785	8254
p43	T03	10203	3508	1439	35726
p43	T06	257316	14021	68119	70095
p43	T09	285954	36990	73786	91004
p43	T12	295962	74231	108200	50126
p44	T06	12943	5975	3894	2056
p44	T12	331116	141680	155674	53608
p44	T0	50099	22107	21503	8727
p45	T03	154404	49067	85459	76956
p45	T06	84767	26532	46076	11116
p45	T12	202913	55066	123126	24928
p47	T03	28456	22186	4375	14039
p47	T06	41728	28115	11915	6781
p47	T09	52995	33775	17995	2452
p48	T03	206113	17105	113743	51914
p48	T06	242572	38843	153596	61852

p48	T09	269283	36875	188478	96328
p48	T0	376273	184628	141439	31090
p49	T03	18315	9220	6394	3156
p49	T06	17663	3058	7765	6713
p49	T09	128245	36320	76847	37633
p49	T12	48047	19380	23779	7543
p49	T0	131243	63733	42438	8613
p50	T03	38236	7575	21568	29622
p50	T06	27492	6637	12084	12878
p50	T09	66818	18499	38151	66406
p50	T12	61996	24299	30427	30126
p52	T03	45715	27278	14862	91698
p52	T06	29930	18198	7984	76376
p52	T09	29423	22575	4027	53120
p53	T03	47030	28214	12890	7365
p53	T06	46827	24490	13934	4034
p53	T09	431502	238648	138050	24892
p53	T12	142678	86561	45180	2798
p55	T03	107422	39501	60046	26500
p55	T09	159123	61336	83996	8559
p55	T12	273310	65192	178918	17227
p59	T03	18576	11922	3499	19598
p59	T09	101015	85777	11187	22302
p59	T12	234907	115046	64881	26412
p61	T03	125145	58908	34386	15040
p61	T12	404667	111497	220583	64000
p66	T03	85143	4027	43664	5936
p66	T06	586291	101094	306430	98920
p66	T09	502321	174829	260046	81323
p66	T12	199037	80890	91169	27526

A



B

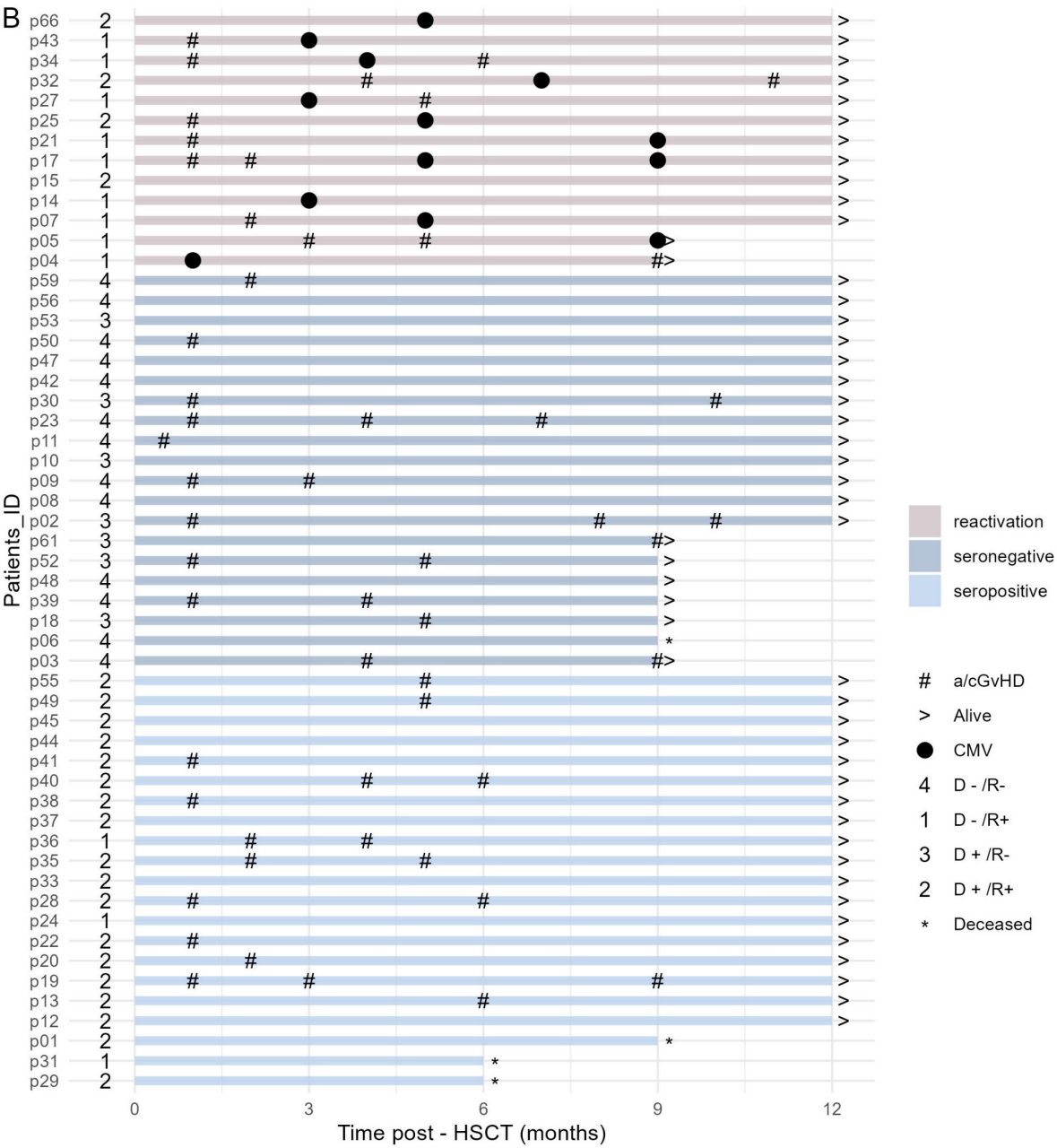
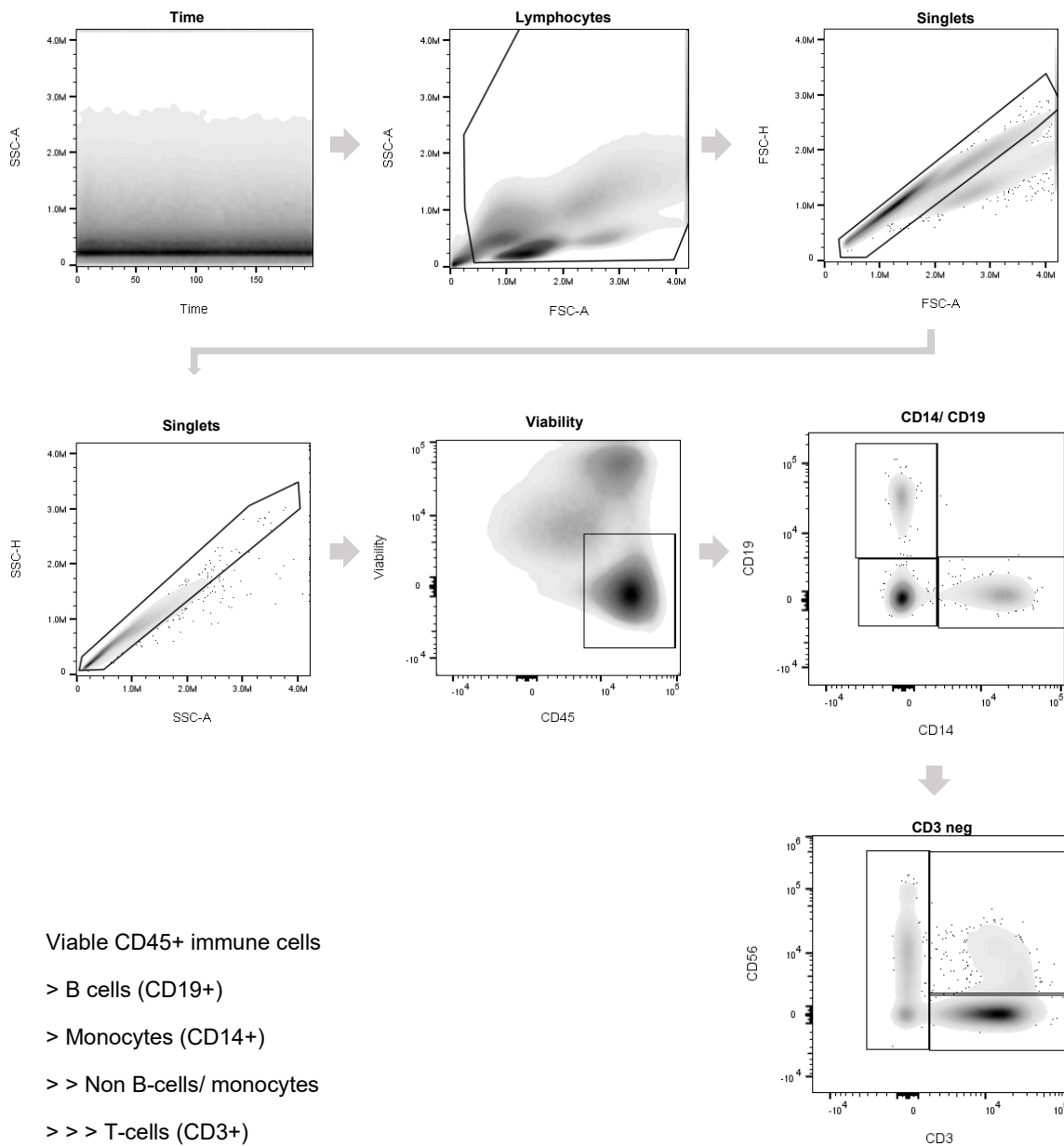


Fig. S1 (A) Peripheral blood samples derived from five different timepoints (pre-HSCT from the donor and 3-, 6-, 9-, 12-months from the recipient) were analyzed by spectral flow cytometry and TCR immunosequencing **(B)** Swimmer plots detailing the patients' clinical course with selective clinical events and stratified according CMV serostatus and infection/reactivation episodes. Bottom horizontal axis display months from HSCT and each line represents an HSCT recipient.

A

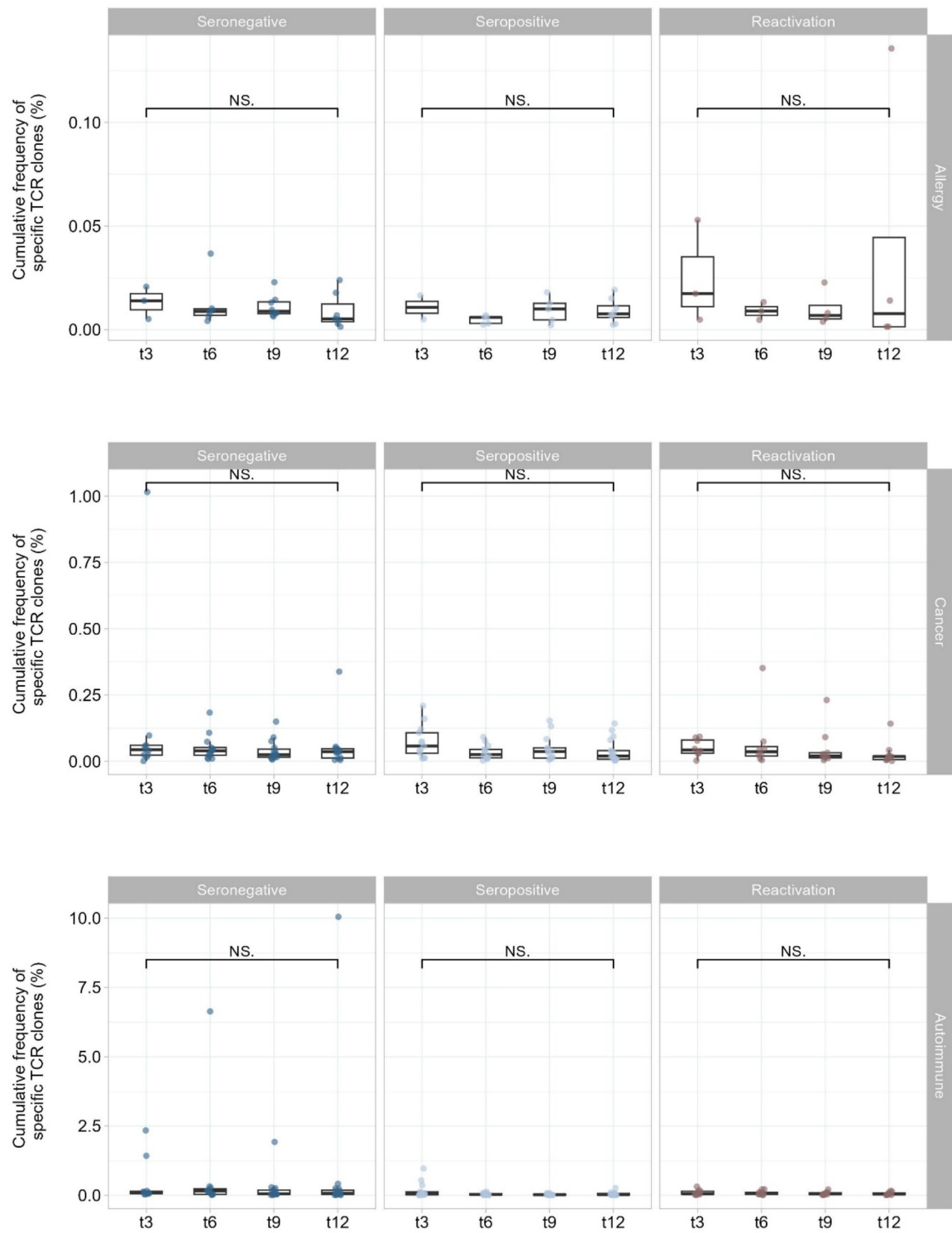


B

Viable CD45+ immune cells

- > B cells (CD19+)
- > Monocytes (CD14+)
- >> Non B-cells/ monocytes
- >>> T-cells (CD3+)
- >>>> CD4+ T-cells (CD4+ CD8a-)
- >>>> CD8+ T-cells (CD4- CD8a+)
- >>>> CD4+ CD8+ DP T-cells
- >>>> CD8- CD4- DN T-cells
- >>> NKT cells (CD3+ CD56+)
- >>> Non T cells (CD3-)
- >>>> NK-cells (CD56+ CD16+)
- >>>>> CD56bright NK-cells (CD56++ CD16+/-)
- >>>>> CD56dim NK-cells (CD56+ CD16+/-)
- >>>>> Unconventional NK-cells (CD56- CD16+)

Fig. S2 (A) Gating strategy for cleaning and identification of viable CD45⁺ immune cells **(B)** Gating hierarchy for identification of major immune cell subsets based on following markers: CD3, CD8a, CD4, CD19, CD14, CD16 and CD56.



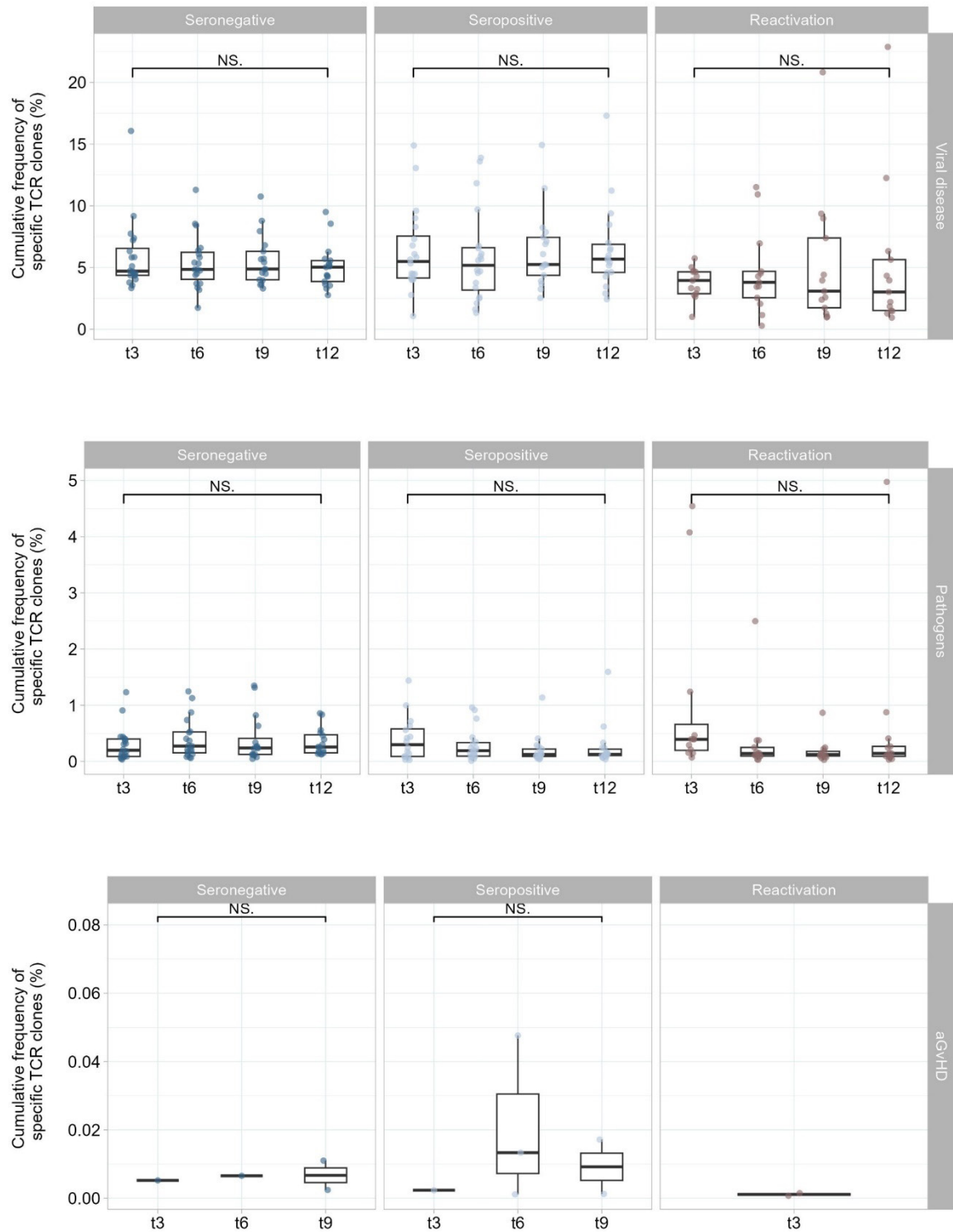


Fig. S3 Cumulative productive frequency of pathogen-specific TCR clones identified by *in silico* matching with public databases at serial timepoints post-HSCT (t0, $n = 53$; t3, $n = 54$; t6, $n = 52$; t9, $n = 47$; t12, $n = 48$). Box plots display medians and interquartile ranges (IQR), with whiskers representing $1.5 \times$ IQR. Wilcoxon sum rank test in (A). All p values were 2 sided. Statistical thresholds: (NS) not significant.

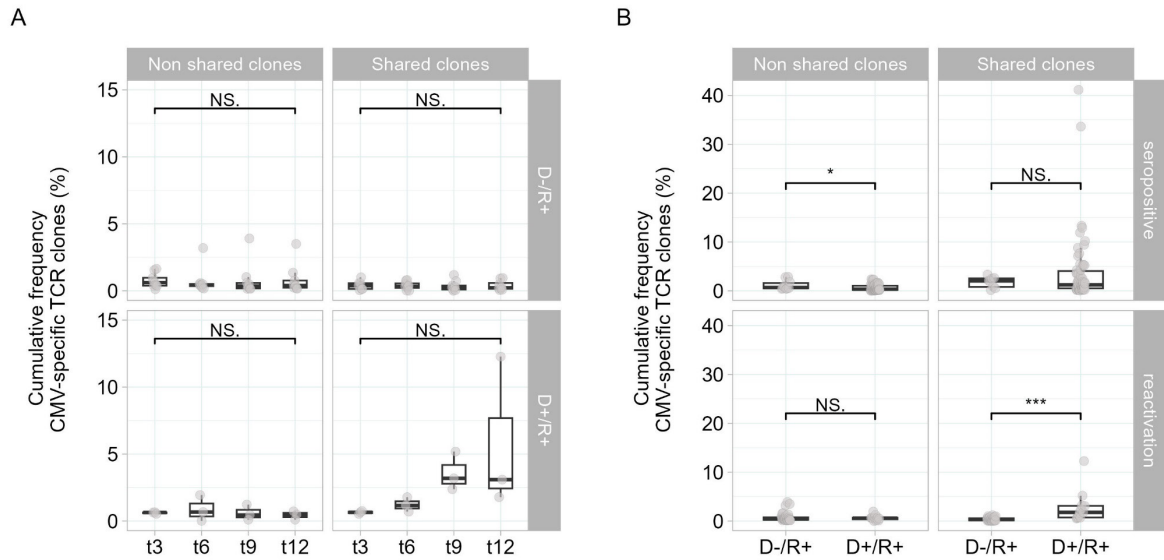
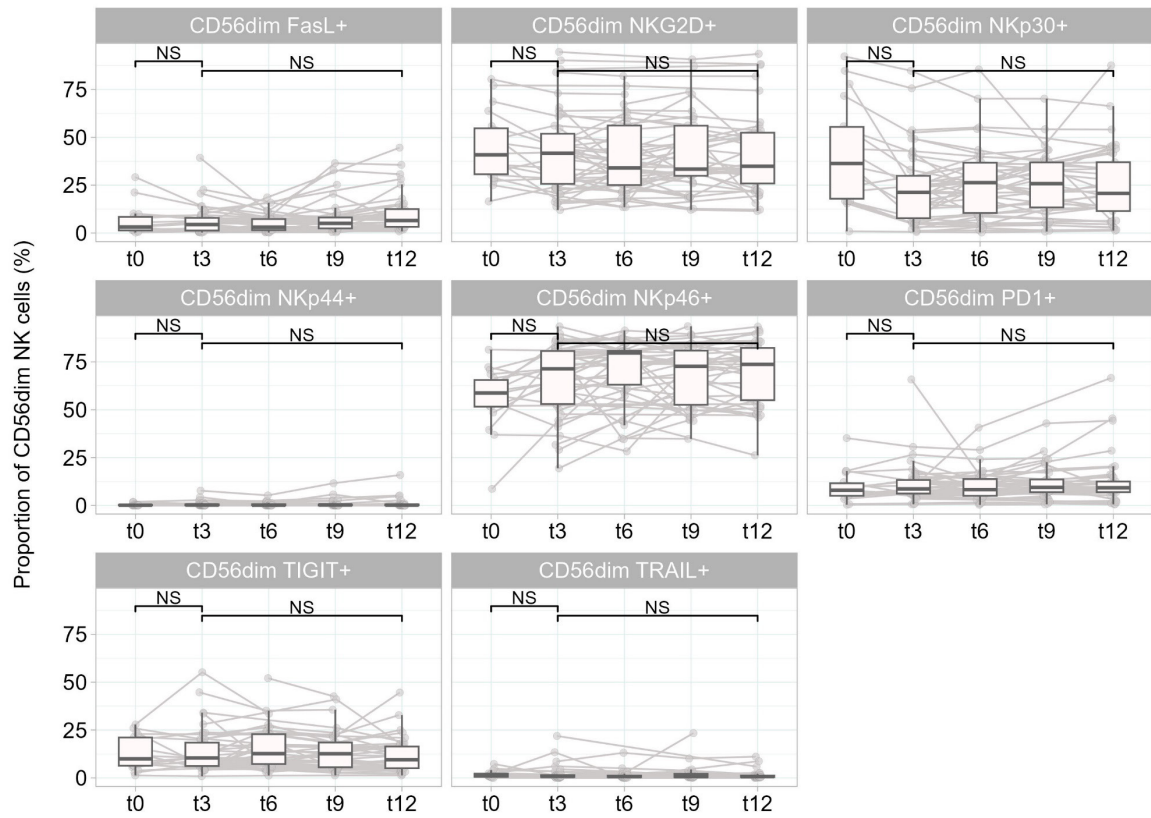


Fig. S4 Stratification according to following groups: seronegative ($n = 20$), seropositive ($n = 21$) and reactivated ($n = 13$) CMV recipients applies to plot (A) and (B). **(A)** Cumulative productive frequency of donor-recipient non-shared and shared CMV- specific TCR clones in D-/R+ ($n = 9$) and D+/R+ ($n = 4$) CMV reactivated recipients at indicated timepoints post-HSCT. **(B)** Cumulative productive frequency of donor-recipient non-shared and shared CMV-specific TCR clones according to CMV serostatus and reactivation. Box plots display medians and interquartile ranges (IQR), with whiskers representing $1.5 \times$ IQR. Wilcoxon rank sum test in (A), (B). All p values were 2 sided. Statistical thresholds: (*) $p < 0.05$, (***) $p < 0.001$ and (NS) not significant.

A



B

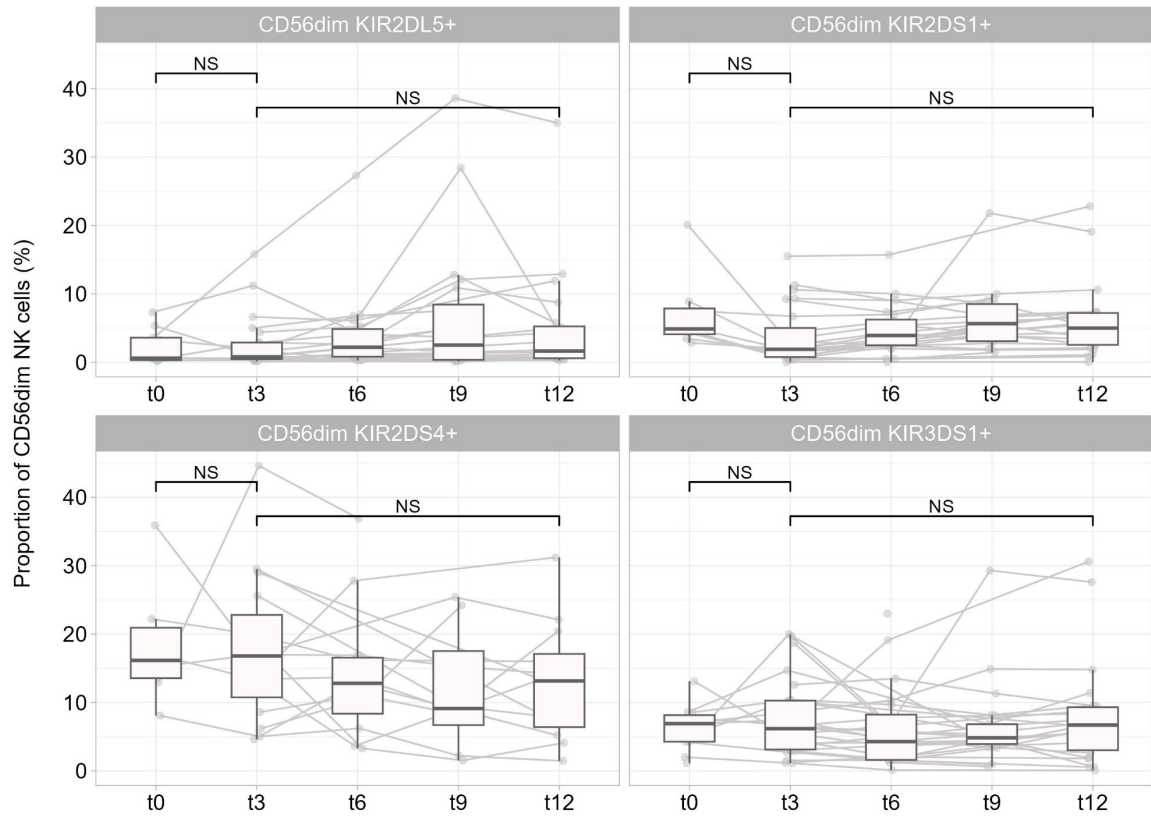


Fig. S5 (A) Proportion of NKp44⁺, NKp30⁺, NKp46⁺, TRAIL⁺, FasL⁺, NKG2D⁺, PD-1⁺, TIGIT⁺ CD56^{dim} cell subsets at indicated timepoints post-HSCT: t0 (n = 18), t3 (n = 44), t6 (n = 35), t9 (n = 32), t12 (n = 37). **(B)** Proportion of KIR2DL5⁺, KIR2DS1⁺, KIR2DS4⁺, KIR3DS1⁺ CD56^{dim} cell subsets at indicated timepoints post-HSCT: t0 (n = 18), t3 (n = 44), t6 (n = 35), t9 (n = 32), t12 (n = 37). Lines connect paired samples. Box plots display medians and interquartile ranges (IQR), with whiskers representing 1.5× IQR. Wilcoxon rank sum test with FDR correction in (A), (B). All *p* values were 2 sided. Statistical thresholds: (NS) not significant.

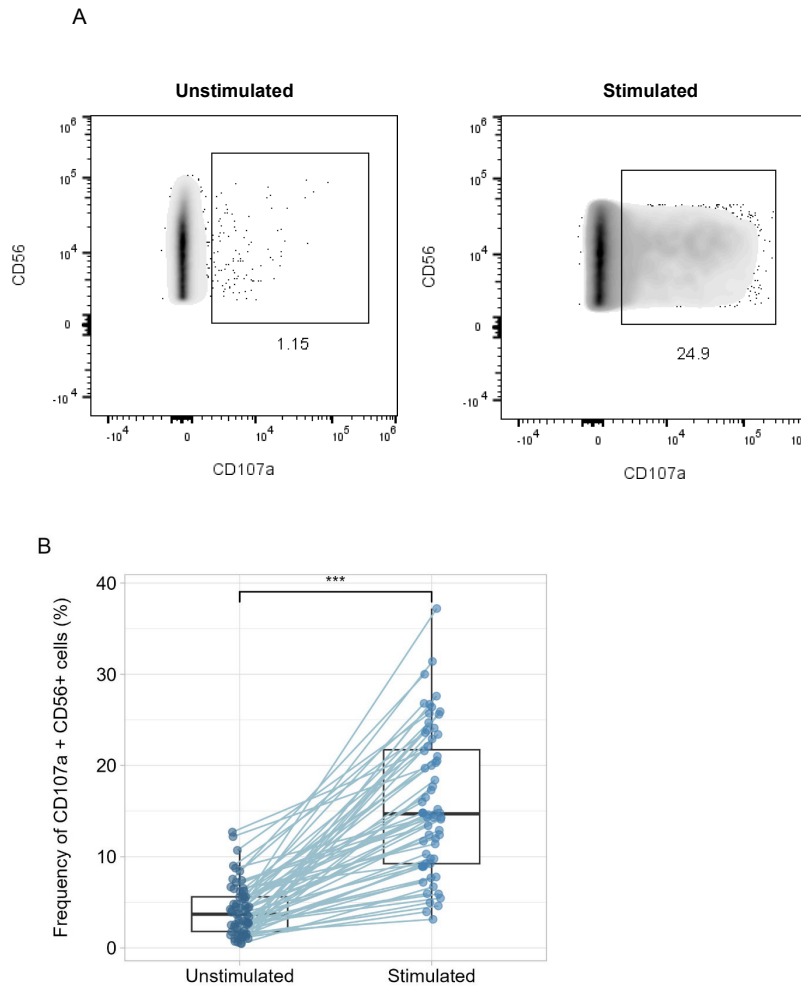


Fig. S6 (A) Representative single-cell plots displaying CD56 and CD107a expression frequency without K562 and post-K562 stimulation. **(B)** Frequency of CD107a⁺ CD56⁺ NK cells in unstimulated and stimulated cells. Lines connect paired samples. Box plots display medians and interquartile ranges (IQR), with whiskers representing 1.5× IQR. Wilcoxon rank sum test in (B). All *p* values were 2 sided. Statistical thresholds: (***) *p* < 0.001.

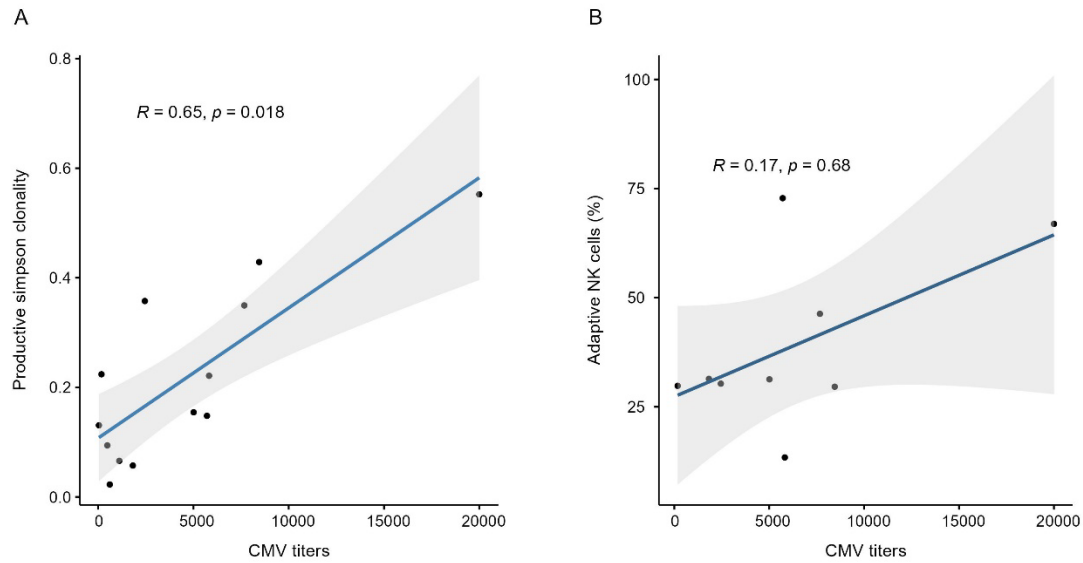
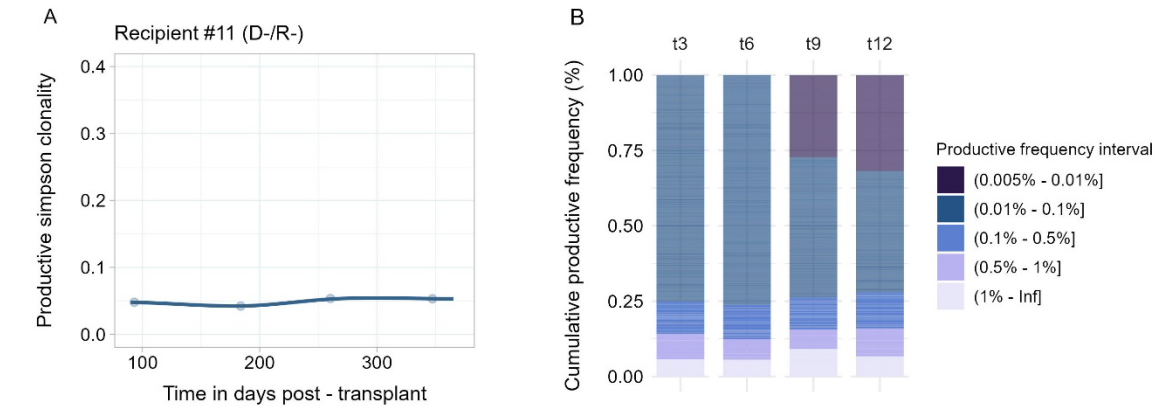
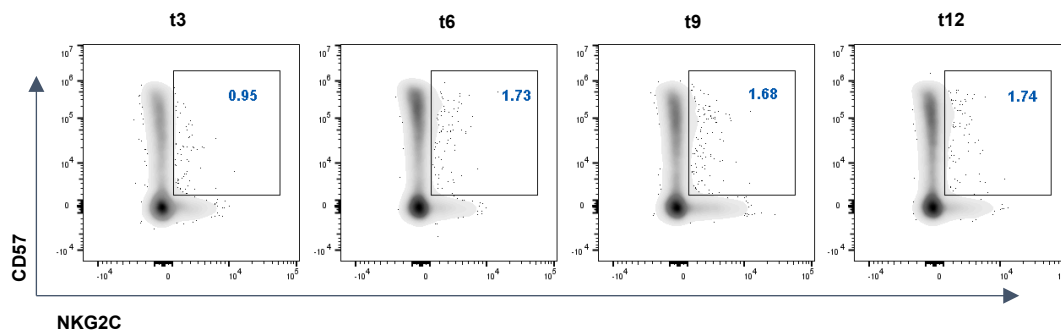


Fig. S7: **(A)** Scatter plot showing CMV peak viral loads in CMV reactivated recipients (x axis) in relation to the productive Simpson clonality of the TCR repertoire (y axis) **(B)** Scatter plot showing CMV peak viral loads in CMV reactivated recipients (x axis) in relation to the frequency of CD57⁺ NKG2C⁺ CD56^{dim} NK cells at 12 months post-alloHSCT (y axis). Spearman rank correlation coefficient is indicated on each plot with the corresponding p -value.

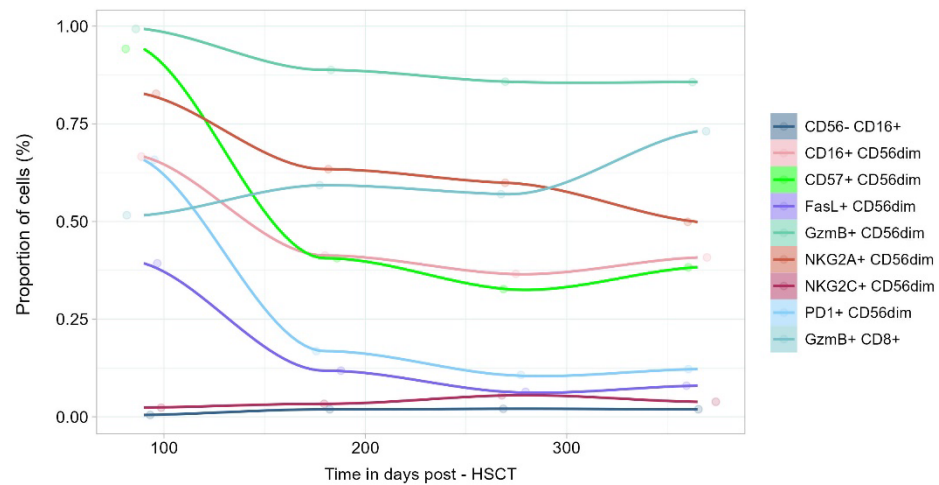
a.



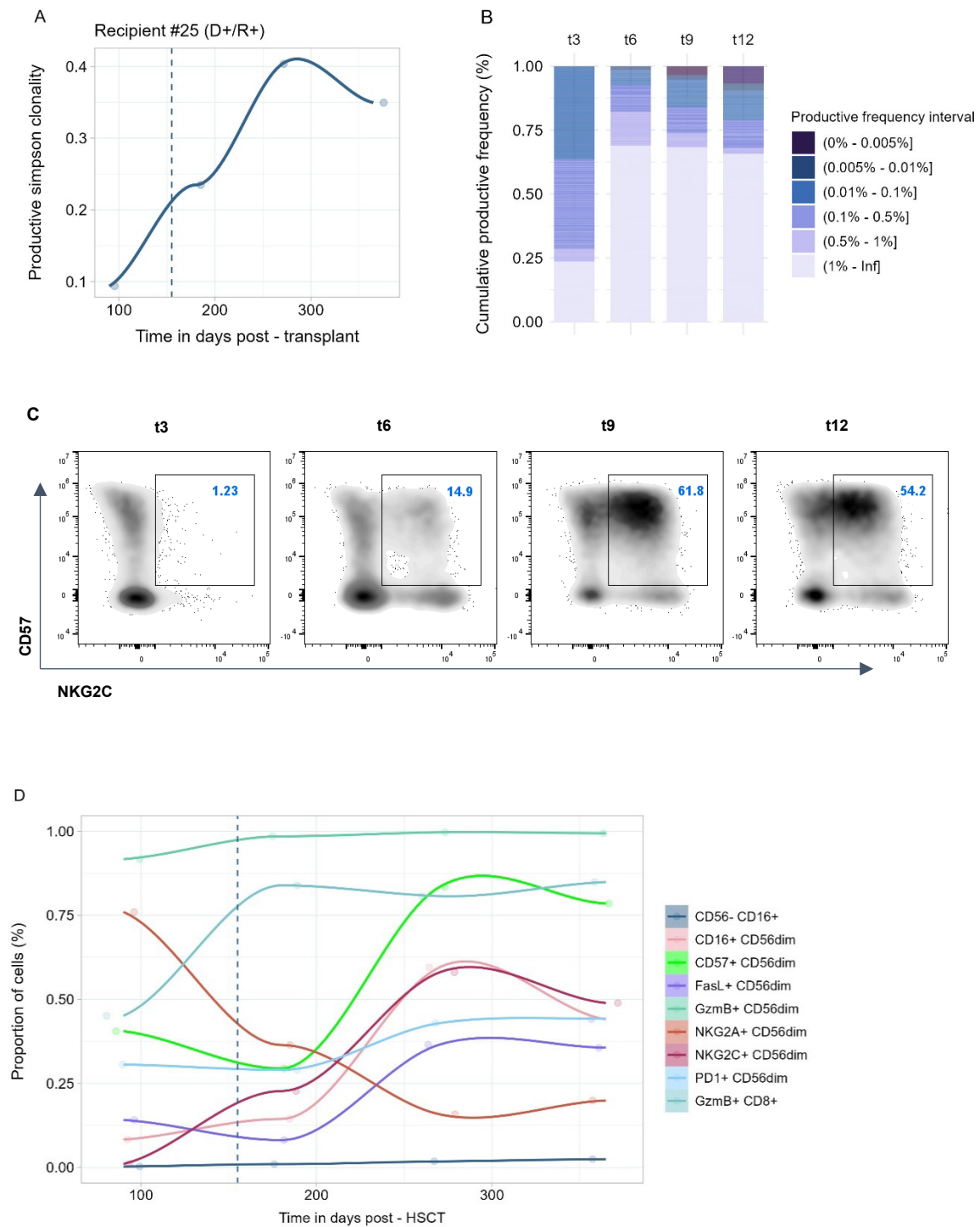
c



D



b.



C.

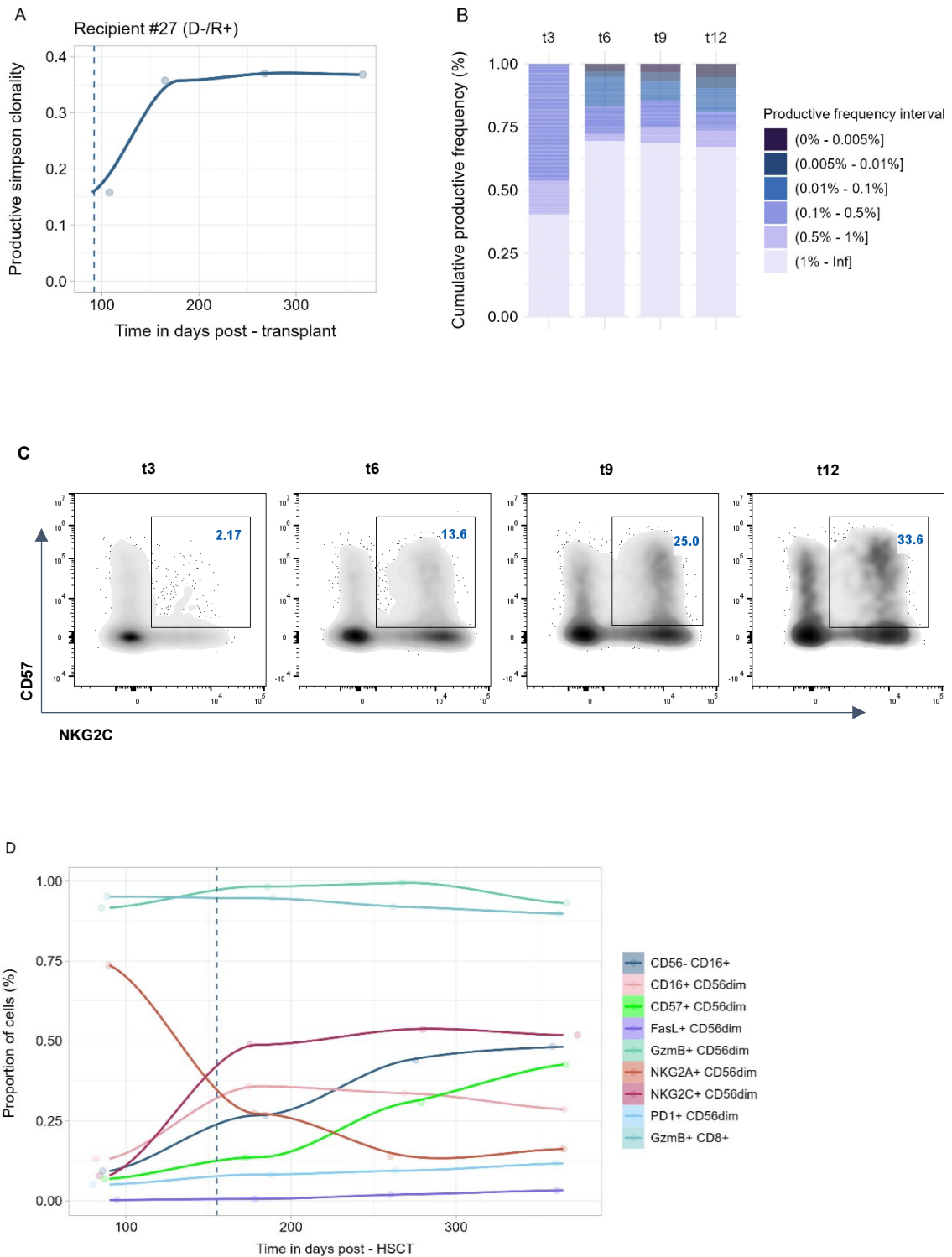


Fig. S8: Immune reconstitution of representative CMV seronegative (a.) and CMV reactivated (b.), (c.) recipients. **(A)** Evolution of the TCR repertoire post-alloHSCT based on the productive Simpson clonality (y axis) in a representative recipient. The vertical dotted line represents time of CMV reactivation. **(B)** TCR fractal clonal size organization defined by the productive frequency of clones at indicated timepoints. The color-coded legend bar represents the cumulative frequency of clones stratified according to the individual clone's productive frequency. **(C)** Representative flow cytometry plot of NKG2C⁺ CD57⁺ NK cells gated on CD56^{dim} NK cells at different timepoints post-alloHSCT. **(D)** Evolution of the frequency of distinct cell subsets at different timepoints post-alloHSCT. The color-coded legend represents the distinct cell subsets.

PERSPECTIVES

This thesis is an attempt to contribute to the pressing need of improving HSCT success rates by exploring the potential of the KIR immunogenetic system to refine the donor selection algorithm and by increasing learnings on the complex re-establishment of immune compartments and its interplay between immune regulators. This last part of the thesis will be dedicated to outline some future research avenues that could be investigated building up on this work.

KIR-mediated NK cell alloreactivity in allogeneic HSCT: how can we move forward?

In the first chapter of this thesis, we sort out to revisit the predictive power of high-resolution KIR genotyping using a state-of-the-art KIR sequencing workflow in a large retrospective cohort of allogeneic HSCT transplanted recipients. We could demonstrate that KIR high-resolution genotyping informs post-transplant NK cell alloreactivity mainly driven by KIR2DS4, KIR2DL2/L3 and KIR3DL1 alleles. Our results are an important step towards dissecting the KIR system's contribution to transplant outcomes.

Although disease association studies are a powerful research study design, lacking biological knowledge may lead to potential biological misinterpretation, thus hampering accurate clinical application. While the advent of high-resolution KIR sequencing has greatly increased the granularity of KIR immunogenetic studies, there are still several outstanding questions and unresolved issues.

Addressing mechanisms by which NK cells render into a hyporesponsiveness or activated status by gaining deeper molecular knowledge of the NK cell immunological synapse might be a first important step. A recent study showcased how the spatial ligand distribution on target cells influences the functional outcome of receptor stimulation [343]. They demonstrated that the spatial ligand distribution on target cells shaped as a donut was subject to a delayed NK cell degranulation as compared to a dot-shaped ligand distribution and represents an additional mechanism of immune evasion against NK cell recognition and killing [344]. We are further lacking data regarding the spatial engagement or clustering of KIRs on the NK cell surface. In a similar fashion to the presence of KIR allotype binding affinities, to what extent are activating and inhibitory KIRs able to cluster, thus increasing the binding avidity and enhancing or curtailing the NK cell activation threshold? Current technological advents such as spatial proteomics through molecular pixelation holds the ability to assess the protein abundance and the degree of spatial clustering or polarization in stimulated settings, which might be of great interest to fill these gaps [345].

Complex tumor immune escape mechanisms greatly challenge a comprehensive understanding of NK cell alloreactivity post-transplant. A well-known tumor feature to evade HLA-I mediated CD8⁺ T cell recognition is the epigenetic downregulation and loss of heterozygosity of histocompatibility ligands [346-348], thus on a theoretical scale potentiating the NK cell anti-tumor activity. It would be interesting to characterize the HLA ligand distribution in different hematological cancer settings. We could further

think of applying allele-specific HLA loss sequencing and bioinformatic identification tools in stem cell transplanted recipients to differentiate the HLA ligand expression landscape on different tumor types and investigate the effect of a predicted missing-self immunogenicity of NK cells following classical licensing rules [349]. As a way of example, we might expect that specific loss of HLA-C2*08:02 would unleash KIR2DL1⁺ NK cell cytotoxicity. In case this requirement is not set, it could hint us towards an abrogated KIR-mediated cytotoxicity due to, e.g., the upregulation of inhibitory receptor ligands such as PD-L1 or the shedding of activating ligands such as MICA/B in a tumor microenvironment.

There is mounting evidence that HLA-I downregulation is not a universal immune escape mechanism and might be maintained on cancerous cells [350]. Based on our work, we might speculate a superiority of classical missing-self immunogenicity of NK cells over an environment susceptible to induce tolerance disruption. It might be thus of interest to acknowledge whether in HLA-matched settings without HLA-I downregulation or HLA-loss, i.e., without proofed missing-self immunogenicity, NK cell alloreactivity could be unleashed. One mechanism by which tumors hinder immunological recognition is genetic alterations leading to the presentation of neoantigens on MHC complexes [351]. While peptide recognition represents one of the core pathways utilized by T cells, peptide dependency by NK cells is less well studied. However, there is mounting evidence of an immunopeptidome contribution to the NK cell activity, with peptide selectivity modulating, for example the binding affinity or the ligand specificity [352]. Pathological settings bear the ability to escape NK cell recognition by inducing the presentation of peptides sensitive to recognition by inhibitory KIRs, as demonstrated in hepatitis C and HIV settings [353, 354]. Could this NK cell-based peptide selectivity be potentially implicated in cancer immunity? We might postulate that peptide alterations induced by tumors might either disrupt activating KIRs interactions, thus reducing their activation threshold or in case of inhibitory KIRs, abrogate a missing-self situation.

It has become evident in recent years that tumor microenvironments have the ability to upregulate the expression of various inhibitory checkpoint receptor ligands such as PD-L1 and CD155 with the potential to curtail KIR-mediated alloreactivity and render NK cells into an exhaustion state [355]. Beside extensive immunophenotypic studies to assess its expression in specific hematological cancer settings, it would be of interest to test the effect of checkpoint blockade antibodies in combination with favorable KIR/HLA allelic configurations to restore the NK cell activity. A novel discovery has been the identification of the KIR3DL3 ligand, HHLA-2, receptor belonging to the B7 family with a propensity to be expressed on various solid tumors [356]. Recent data unveiled a high expression level of KIR3DL3 on CD8⁺T cells and to a lesser extent on NK cells with KIR3DL3 allele polymorphism modulating the expression level [357, 358]. Given that KIR3DL3 is a ubiquitous gene expressed in each individual at two copies, it represents a potential interesting target to investigate in non-solid tumors and whether its expression abrogates NK cell alloreactivity post-transplant.

T and NK cell alloreactivity have for long been seen as competing mechanisms, with a propensity of the T cell alloreactivity to overweight KIR-mediated NK cell alloreactivity. Recent work unveiled a potential direct crosstalk between NK and T cells through KIR and HLA interactions, with the number of interactions predicting the T cell longevity [182, 234]. We might thus hypothesize that the level of

inhibitory KIR and HLA interactions might magnify the effect of T cell alloreactivity with the potential to modulate the durability and the extent of a transplant outcome (e.g., GvL and GvHD), which might be an interesting research axis to investigate in the future.

Although many studies, including our research analysis, point towards a promising implication of the KIR immunogenetic system in allogeneic HSCT, the extent to which we can harness NK cells anti-tumoral properties still needs further mechanistical and associative studies on confirmatory cohorts. Gaining deeper knowledge on the molecular level and increasing data regarding the crosstalk between NK cells and other immune cells might be paramount to understand the differential results gathered until now and why some diseases are presumably resistant to NK cell alloreactivity in comparison to others.

Immune reconstitution modulation by non-genetic factors: implications and potential mechanisms

Allogeneic hematopoietic stem cell transplantation represents a powerful human *in vivo* model to decipher the immune system establishment, otherwise only feasible in newborns with a major logistical burden or in mice models. In the last chapter of this thesis, we have provided insights into the reconstitution of the T and NK cell repertoire after allogeneic HSCT leveraging high-dimensional immunophenotyping and next-generation sequencing. We showed that these repertoires establish at an early timepoint post-transplant and undergo coordinated inflationary changes upon CMV reactivation. Our data paves the way for further exploration, some of which will be briefly discussed in the following section.

To what extent does timing after transplantation affect our results? According to our findings, the majority of the TCR repertoire establishes on a quantitative level during the first 90 days, whereas a progressive increase in the overlap thereafter suggests a more refined remodeling and renewing of TCR clonotypes at the periphery, reflecting (i) the central output of new TCRs (ii) the competitive survival at the periphery, rendering clones with a higher fitness. In future work, it would be of interest to get a deeper understanding of reconstitution pathways during this key window of 90 to 100 days post-transplant and as to what extent the immune reconstitution differs in consequence to early and late occurring clinical events. Could we hypothesize that enhanced reconstituted immune repertoires display a higher resistance towards immune perturbations, i.e., a reduced susceptibility to immune changes, or in contrast, is the immune system during the first 3 months extremely adaptable with the capacity to re-equilibrate potential shifts? There is supporting evidence in newborns that 3 months might represent a setting point upon which immune repertoires are shaped by environmental events [278].

We have already partly assessed in the current state of research that multiple genetic and non-genetic factors have been recognized to imprint our immune state and functionality (cf. current state of research, chapter IV). As a result of the research work in our lab and findings from this current thesis, CMV seems to be one of the leading immunomodulatory events, although there are still several uncertainties with regard to its imprinting behavior. The resolution of the *in silico* TCR specificity prediction was surprisingly low in our study, especially regarding the dominant clones in CMV reactivated recipients. Specificity and peptide binding TCR prediction could greatly benefit from complementary bioinformatic approaches to

enrich the TCR analysis, as experimentally derived epitope identification is time, cost consuming and thus lacks completeness [359]. As a way of example, recent literature showcased that the TCR repertoire is subject to a shift from high affinity dominant clones towards low affinity clones in persistent CMV antigen exposure, termed as reverse evolution [360]. This pathway likely reflects the adaptation of the T cell compartment to chronic peripheral low viral exposure. It would be of interest to apply existing bioinformatic tools to capture TCR-pMHC complex affinity spectra and dissect whether CMV occurrence in reconstituting immune systems follows this classical pathway to re-equilibrate the TCR repertoire or whether it renders a highly specific and sustained response. To what extent do these differential reconstitution trajectories impact the recipient's post-transplant immunity and susceptibility to infections? There is recent evidence of a correlation between levels of exhausted tumor-infiltrating lymphocytes and disease persistence in cancer immunology, with the avidity of exhausted T cells being inversely proportional to the antigen burden in melanoma [361]. In-depth immunophenotypic approaches would be helpful in deciphering whether persistent antigen stimulation might render the immune system into a cell senescence and exhaustion state, which in turn would lead to a reduced T-cell mediated immunity.

Adaptive NK cell biology has gained substantial attention in recent years, attributable to its unique persistent memory state as an innate lymphoid cell following a presumably similar inflationary expansion as T cells. Recent data have provided insights into the adaptive NK cell repertoire being subject to more heterogeneity than initially assumed based on epigenetic traits [362]. Albeit our data suggests a skewing towards KIR⁺ selection and increased functionality, there is compelling evidence that the NK cell adaptive compartment renders into states of anergy, exhaustion and senescence, similar to what has been described for T cells [363]. It would be useful to study adaptive NK cell expansion in a more dynamic fashion during the transition from an acute infection into a memory and long-term phase to (i) understand NK cell functional states and (ii) receptor expression behavior in cases of CMV rechallenge under the assumption of an inflationary expansion.

Individuals are thought to display a spectrum of immune state variation as a result of additive or reversing internal and external factors [364]. Immune reconstitution trajectories are likely to result from the effect of multifactorial events acting synergistically, which represents a remarkable challenge to disentangle. Extending the number of recipients in future studies would be instrumental to identify reoccurring patterns using state-of-the-art systems' immunology technical and bioinformatic tools. Potential intriguing factors besides the ones stated in the current state of research would be vaccination episodes which have partially been addressed in the following publication [315], but enhanced data is lacking. The expansion of memory-specific T lymphocytes through vaccination might shift back the state of exhaustion and enhance T cell cytotoxicity in CMV-reactivated recipients. Further factors that might be taken into account is the gut microbiome and the use of antibiotics after transplantation.

With an increasing amount of generated data, standardization across experimental protocols and application of high-resolution techniques, we might enhance our comprehension about the susceptibility of recipients' immune systems to various regulators in the reconstitution phase and how interindividual variation differentially impacts transplant outcomes.

REFERENCES

1. Angelini, D.F., et al., *FcgammaRIII discriminates between 2 subsets of Vgamma9Vdelta2 effector cells with different responses and activation pathways*. Blood, 2004. **104**(6): p. 1801-7.
2. Clemenceau, B., et al., *Effector memory alphabeta T lymphocytes can express FcgammaRIIIa and mediate antibody-dependent cellular cytotoxicity*. J Immunol, 2008. **180**(8): p. 5327-34.
3. Vivier, E., et al., *Tyrosine phosphorylation of the Fc gamma RIII(CD16): zeta complex in human natural killer cells. Induction by antibody-dependent cytotoxicity but not by natural killing*. J Immunol, 1991. **146**(1): p. 206-10.
4. Prager, I. and C. Watzl, *Mechanisms of natural killer cell-mediated cellular cytotoxicity*. J Leukoc Biol, 2019. **105**(6): p. 1319-1329.
5. Ortaldo, J.R., A.T. Mason, and J.J. O'Shea, *Receptor-induced death in human natural killer cells: involvement of CD16*. J Exp Med, 1995. **181**(1): p. 339-44.
6. Dustin, M.L. and E.O. Long, *Cytotoxic immunological synapses*. Immunol Rev, 2010. **235**(1): p. 24-34.
7. Warren, H.S. and B.F. Kinnear, *Quantitative analysis of the effect of CD16 ligation on human NK cell proliferation*. J Immunol, 1999. **162**(2): p. 735-42.
8. Anegón, I., et al., *Interaction of Fc receptor (CD16) ligands induces transcription of interleukin 2 receptor (CD25) and lymphokine genes and expression of their products in human natural killer cells*. J Exp Med, 1988. **167**(2): p. 452-72.
9. Fauriat, C., et al., *Regulation of human NK-cell cytokine and chemokine production by target cell recognition*. Blood, 2010. **115**(11): p. 2167-76.
10. Eissens, D.N., et al., *Defining early human NK cell developmental stages in primary and secondary lymphoid tissues*. PLoS One, 2012. **7**(2): p. e30930.
11. Dogra, P., et al., *Tissue Determinants of Human NK Cell Development, Function, and Residence*. Cell, 2020. **180**(4): p. 749-763 e13.
12. Freud, A.G., et al., *Evidence for discrete stages of human natural killer cell differentiation in vivo*. J Exp Med, 2006. **203**(4): p. 1033-43.
13. Abel, A.M., et al., *Natural Killer Cells: Development, Maturation, and Clinical Utilization*. Front Immunol, 2018. **9**: p. 1869.
14. Moroso, V., et al., *NK cells can generate from precursors in the adult human liver*. Eur J Immunol, 2011. **41**(11): p. 3340-50.
15. Vacca, P., et al., *CD34+ hematopoietic precursors are present in human decidua and differentiate into natural killer cells upon interaction with stromal cells*. Proc Natl Acad Sci U S A, 2011. **108**(6): p. 2402-7.
16. Torcellan, T., et al., *Circulating NK cells establish tissue residency upon acute infection of skin and mediate accelerated effector responses to secondary infection*. Immunity, 2024. **57**(1): p. 124-140 e7.

17. Cichocki, F., B. Grzywacz, and J.S. Miller, *Human NK Cell Development: One Road or Many?* Front Immunol, 2019. **10**: p. 2078.
18. Yang, C., et al., *Heterogeneity of human bone marrow and blood natural killer cells defined by single-cell transcriptome*. Nat Commun, 2019. **10**(1): p. 3931.
19. Melsen, J.E., et al., *Single-cell transcriptomics in bone marrow delineates CD56(dim)GranzymeK(+) subset as intermediate stage in NK cell differentiation*. Front Immunol, 2022. **13**: p. 1044398.
20. Mace, E.M., *Human natural killer cells: Form, function, and development*. J Allergy Clin Immunol, 2023. **151**(2): p. 371-385.
21. Seo, S. and E.M. Mace, *Diversity of human NK cell developmental pathways defined by single-cell analyses*. Curr Opin Immunol, 2022. **74**: p. 106-111.
22. Schorr, C., M.S. Krishnan, and M. Capitano, *Deficits in our understanding of natural killer cell development in mouse and human*. Curr Opin Hematol, 2023. **30**(4): p. 106-116.
23. Ivarsson, M.A., et al., *Differentiation and functional regulation of human fetal NK cells*. J Clin Invest, 2013. **123**(9): p. 3889-901.
24. Bjorkstrom, N.K., H.G. Ljunggren, and J. Michaelsson, *Emerging insights into natural killer cells in human peripheral tissues*. Nat Rev Immunol, 2016. **16**(5): p. 310-20.
25. Wilson, M.J., M. Torkar, and J. Trowsdale, *Genomic organization of a human killer cell inhibitory receptor gene*. Tissue Antigens, 1997. **49**(6): p. 574-9.
26. Colonna, M., et al., *A high-resolution view of NK-cell receptors: structure and function*. Immunol Today, 2000. **21**(9): p. 428-31.
27. Uhrberg, M., et al., *Human diversity in killer cell inhibitory receptor genes*. Immunity, 1997. **7**(6): p. 753-63.
28. Martin, A.M., et al., *The genomic organization and evolution of the natural killer immunoglobulin-like receptor (KIR) gene cluster*. Immunogenetics, 2000. **51**(4-5): p. 268-80.
29. Beziat, V., et al., *Deciphering the killer-cell immunoglobulin-like receptor system at super-resolution for natural killer and T-cell biology*. Immunology, 2017. **150**(3): p. 248-264.
30. Moesta, A.K., et al., *Humans differ from other hominids in lacking an activating NK cell receptor that recognizes the C1 epitope of MHC class I*. J Immunol, 2010. **185**(7): p. 4233-7.
31. Venstrom, J.M., et al., *HLA-C-dependent prevention of leukemia relapse by donor activating KIR2DS1*. N Engl J Med, 2012. **367**(9): p. 805-16.
32. Graef, T., et al., *KIR2DS4 is a product of gene conversion with KIR3DL2 that introduced specificity for HLA-A*11 while diminishing avidity for HLA-C*. J Exp Med, 2009. **206**(11): p. 2557-72.
33. Blokhuis, J.H., et al., *KIR2DS5 allotypes that recognize the C2 epitope of HLA-C are common among Africans and absent from Europeans*. Immun Inflamm Dis, 2017. **5**(4): p. 461-468.
34. Hilton, H.G., et al., *Polymorphic HLA-C Receptors Balance the Functional Characteristics of KIR Haplotypes*. J Immunol, 2015. **195**(7): p. 3160-70.
35. Rajagopalan, S. and E.O. Long, *KIR2DL4 (CD158d): An activation receptor for HLA-G*. Front Immunol, 2012. **3**: p. 258.

36. Foley, B.A., et al., *The reactivity of Bw4+ HLA-B and HLA-A alleles with KIR3DL1: implications for patient and donor suitability for haploidentical stem cell transplantations*. Blood, 2008. **112**(2): p. 435-43.
37. Gumperz, J.E., et al., *The Bw4 public epitope of HLA-B molecules confers reactivity with natural killer cell clones that express NKB1, a putative HLA receptor*. J Exp Med, 1995. **181**(3): p. 1133-44.
38. Forlenza, C.J., et al., *KIR3DL1 Allelic Polymorphism and HLA-B Epitopes Modulate Response to Anti-GD2 Monoclonal Antibody in Patients With Neuroblastoma*. J Clin Oncol, 2016. **34**(21): p. 2443-51.
39. Hansasuta, P., et al., *Recognition of HLA-A3 and HLA-A11 by KIR3DL2 is peptide-specific*. Eur J Immunol, 2004. **34**(6): p. 1673-9.
40. Dohring, C., et al., *A human killer inhibitory receptor specific for HLA-A1,2*. J Immunol, 1996. **156**(9): p. 3098-101.
41. Bhatt, R.S., et al., *KIR3DL3 Is an Inhibitory Receptor for HHLA2 that Mediates an Alternative Immunoinhibitory Pathway to PD1*. Cancer Immunol Res, 2021. **9**(2): p. 156-169.
42. Kula, A., et al., *The Importance of HHLA2 in Solid Tumors-A Review of the Literature*. Cells, 2024. **13**(10).
43. Luu, K., H. Schwarz, and A. Lundqvist, *B7-H7 Is Inducible on T Cells to Regulate Their Immune Response and Serves as a Marker for Exhaustion*. Front Immunol, 2021. **12**: p. 682627.
44. Oliveira, M.L.G., et al., *Genetic diversity of the LILRB1 and LILRB2 coding regions in an admixed Brazilian population sample*. HLA, 2022. **100**(4): p. 325-348.
45. Zeller, T., et al., *Perspectives of targeting LILRB1 in innate and adaptive immune checkpoint therapy of cancer*. Front Immunol, 2023. **14**: p. 1240275.
46. Lazetic, S., et al., *Human natural killer cell receptors involved in MHC class I recognition are disulfide-linked heterodimers of CD94 and NKG2 subunits*. J Immunol, 1996. **157**(11): p. 4741-5.
47. Horowitz, A., et al., *Class I HLA haplotypes form two schools that educate NK cells in different ways*. Sci Immunol, 2016. **1**(3).
48. Wang, X., H. Xiong, and Z. Ning, *Implications of NKG2A in immunity and immune-mediated diseases*. Front Immunol, 2022. **13**: p. 960852.
49. Guma, M., et al., *Expansion of CD94/NKG2C+ NK cells in response to human cytomegalovirus-infected fibroblasts*. Blood, 2006. **107**(9): p. 3624-31.
50. Siemaszko, J., A. Marzec-Przyszlak, and K. Bogunia-Kubik, *Activating NKG2C Receptor: Functional Characteristics and Current Strategies in Clinical Applications*. Arch Immunol Ther Exp (Warsz), 2023. **71**(1): p. 9.
51. Basilio-Queiros, D. and E. Mischak-Weissinger, *Natural killer cells- from innate cells to the discovery of adaptability*. Front Immunol, 2023. **14**: p. 1172437.
52. Wu, J., et al., *An activating immunoreceptor complex formed by NKG2D and DAP10*. Science, 1999. **285**(5428): p. 730-2.
53. Siemaszko, J., A. Marzec-Przyszlak, and K. Bogunia-Kubik, *NKG2D Natural Killer Cell Receptor-A Short Description and Potential Clinical Applications*. Cells, 2021. **10**(6).

54. Barrow, A.D., C.J. Martin, and M. Colonna, *The Natural Cytotoxicity Receptors in Health and Disease*. Front Immunol, 2019. **10**: p. 909.
55. Pesce, S., et al., *PD/1-PD-Ls Checkpoint: Insight on the Potential Role of NK Cells*. Front Immunol, 2019. **10**: p. 1242.
56. Quatrini, L., et al., *The Immune Checkpoint PD-1 in Natural Killer Cells: Expression, Function and Targeting in Tumour Immunotherapy*. Cancers (Basel), 2020. **12**(11).
57. Zhang, Q., et al., *Blockade of the checkpoint receptor TIGIT prevents NK cell exhaustion and elicits potent anti-tumor immunity*. Nat Immunol, 2018. **19**(7): p. 723-732.
58. Cao, Y., et al., *Immune checkpoint molecules in natural killer cells as potential targets for cancer immunotherapy*. Signal Transduct Target Ther, 2020. **5**(1): p. 250.
59. Wang, J.M., et al., *KLRG1 negatively regulates natural killer cell functions through the Akt pathway in individuals with chronic hepatitis C virus infection*. J Virol, 2013. **87**(21): p. 11626-36.
60. Witt, C.S., A. Martin, and F.T. Christiansen, *Detection of KIR2DL4 alleles by sequencing and SSCP reveals a common allele with a shortened cytoplasmic tail*. Tissue Antigens, 2000. **56**(3): p. 248-57.
61. Hsu, K.C., et al., *Killer Ig-like receptor haplotype analysis by gene content: evidence for genomic diversity with a minimum of six basic framework haplotypes, each with multiple subsets*. J Immunol, 2002. **169**(9): p. 5118-29.
62. Gardiner, C.M., et al., *Different NK cell surface phenotypes defined by the DX9 antibody are due to KIR3DL1 gene polymorphism*. J Immunol, 2001. **166**(5): p. 2992-3001.
63. Pando, M.J., et al., *The protein made from a common allele of KIR3DL1 (3DL1*004) is poorly expressed at cell surfaces due to substitution at positions 86 in Ig domain 0 and 182 in Ig domain 1*. J Immunol, 2003. **171**(12): p. 6640-9.
64. Moesta, A.K., et al., *Synergistic polymorphism at two positions distal to the ligand-binding site makes KIR2DL2 a stronger receptor for HLA-C than KIR2DL3*. J Immunol, 2008. **180**(6): p. 3969-79.
65. Schonberg, K., et al., *Analyses of HLA-C-specific KIR repertoires in donors with group A and B haplotypes suggest a ligand-instructed model of NK cell receptor acquisition*. Blood, 2011. **117**(1): p. 98-107.
66. O'Connor, G.M., et al., *Mutational and structural analysis of KIR3DL1 reveals a lineage-defining allotypic dimorphism that impacts both HLA and peptide sensitivity*. J Immunol, 2014. **192**(6): p. 2875-84.
67. O'Connor, G.M., et al., *Analysis of binding of KIR3DS1*014 to HLA suggests distinct evolutionary history of KIR3DS1*. J Immunol, 2011. **187**(5): p. 2162-71.
68. Bari, R., et al., *KIR2DL2/2DL3-E(35) alleles are functionally stronger than -Q(35) alleles*. Sci Rep, 2016. **6**: p. 23689.
69. Frazier, W.R., et al., *Allelic variation in KIR2DL3 generates a KIR2DL2-like receptor with increased binding to its HLA-C ligand*. J Immunol, 2013. **190**(12): p. 6198-208.
70. Stewart, C.A., et al., *Recognition of peptide-MHC class I complexes by activating killer immunoglobulin-like receptors*. Proc Natl Acad Sci U S A, 2005. **102**(37): p. 13224-9.

71. Bari, R., et al., *Significant functional heterogeneity among KIR2DL1 alleles and a pivotal role of arginine 245*. Blood, 2009. **114**(25): p. 5182-90.
72. Trompeter, H.I., et al., *Three structurally and functionally divergent kinds of promoters regulate expression of clonally distributed killer cell Ig-like receptors (KIR), of KIR2DL4, and of KIR3DL3*. J Immunol, 2005. **174**(7): p. 4135-43.
73. Horowitz, A., et al., *Genetic and environmental determinants of human NK cell diversity revealed by mass cytometry*. Sci Transl Med, 2013. **5**(208): p. 208ra145.
74. Cichocki, F., et al., *Cutting edge: KIR antisense transcripts are processed into a 28-base PIWI-like RNA in human NK cells*. J Immunol, 2010. **185**(4): p. 2009-12.
75. Davies, G.E., et al., *Identification of bidirectional promoters in the human KIR genes*. Genes Immun, 2007. **8**(3): p. 245-53.
76. Santourlidis, S., et al., *Crucial role of DNA methylation in determination of clonally distributed killer cell Ig-like receptor expression patterns in NK cells*. J Immunol, 2002. **169**(8): p. 4253-61.
77. Wright, P.W., et al., *The KIR2DL1 intermediate upstream element participates in gene activation*. Immunogenetics, 2023. **75**(6): p. 495-506.
78. Xia, M., et al., *Epigenetic Regulation of NK Cell-Mediated Antitumor Immunity*. Front Immunol, 2021. **12**: p. 672328.
79. Andersson, S., et al., *KIR acquisition probabilities are independent of self-HLA class I ligands and increase with cellular KIR expression*. Blood, 2009. **114**(1): p. 95-104.
80. Sleiman, M., et al., *NK cell killer Ig-like receptor repertoire acquisition and maturation are strongly modulated by HLA class I molecules*. J Immunol, 2014. **192**(6): p. 2602-10.
81. Dhuyser, A., et al., *KIR in Allogeneic Hematopoietic Stem Cell Transplantation: Need for a Unified Paradigm for Donor Selection*. Front Immunol, 2022. **13**: p. 821533.
82. Marin, W.M., et al., *High-throughput Interpretation of Killer-cell Immunoglobulin-like Receptor Short-read Sequencing Data with PING*. PLoS Comput Biol, 2021. **17**(8): p. e1008904.
83. Norman, P.J., et al., *Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing*. Am J Hum Genet, 2016. **99**(2): p. 375-91.
84. Kim, S., et al., *Licensing of natural killer cells by host major histocompatibility complex class I molecules*. Nature, 2005. **436**(7051): p. 709-13.
85. Yokoyama, W.M. and S. Kim, *How do natural killer cells find self to achieve tolerance?* Immunity, 2006. **24**(3): p. 249-57.
86. Yokoyama, W.M. and S. Kim, *Licensing of natural killer cells by self-major histocompatibility complex class I*. Immunol Rev, 2006. **214**: p. 143-54.
87. Fernandez, N.C., et al., *A subset of natural killer cells achieves self-tolerance without expressing inhibitory receptors specific for self-MHC molecules*. Blood, 2005. **105**(11): p. 4416-23.
88. Raulet, D.H. and R.E. Vance, *Self-tolerance of natural killer cells*. Nat Rev Immunol, 2006. **6**(7): p. 520-31.
89. Brodin, P. and P. Hoglund, *Beyond licensing and disarming: a quantitative view on NK-cell education*. Eur J Immunol, 2008. **38**(11): p. 2934-7.

90. Joncker, N.T., et al., *NK cell responsiveness is tuned commensurate with the number of inhibitory receptors for self-MHC class I: the rheostat model*. J Immunol, 2009. **182**(8): p. 4572-80.
91. Goodson-Gregg, F.J., S.A. Krepel, and S.K. Anderson, *Tuning of human NK cells by endogenous HLA-C expression*. Immunogenetics, 2020. **72**(4): p. 205-215.
92. Goodson-Gregg, F.J., et al., *Tuning of NK-Specific HLA-C Expression by Alternative mRNA Splicing*. Front Immunol, 2019. **10**: p. 3034.
93. Bjorkstrom, N.K., et al., *Expression patterns of NKG2A, KIR, and CD57 define a process of CD56dim NK-cell differentiation uncoupled from NK-cell education*. Blood, 2010. **116**(19): p. 3853-64.
94. Fauriat, C., et al., *Education of human natural killer cells by activating killer cell immunoglobulin-like receptors*. Blood, 2010. **115**(6): p. 1166-74.
95. Hallner, A., et al., *The HLA-B -21 dimorphism impacts on NK cell education and clinical outcome of immunotherapy in acute myeloid leukemia*. Blood, 2019. **133**(13): p. 1479-1488.
96. Hammer, Q., et al., *SARS-CoV-2 Nsp13 encodes for an HLA-E-stabilizing peptide that abrogates inhibition of NKG2A-expressing NK cells*. Cell Rep, 2022. **38**(10): p. 110503.
97. Leijonhufvud, C., et al., *LIR-1 educates expanded human NK cells and defines a unique antitumor NK cell subset with potent antibody-dependent cellular cytotoxicity*. Clin Transl Immunology, 2021. **10**(10): p. e1346.
98. Chen, S., et al., *The Self-Specific Activation Receptor SLAM Family Is Critical for NK Cell Education*. Immunity, 2016. **45**(2): p. 292-304.
99. Orr, M.T., W.J. Murphy, and L.L. Lanier, *'Unlicensed' natural killer cells dominate the response to cytomegalovirus infection*. Nat Immunol, 2010. **11**(4): p. 321-7.
100. Potempa, M., et al., *Influence of Self-MHC Class I Recognition on the Dynamics of NK Cell Responses to Cytomegalovirus Infection*. J Immunol, 2022. **208**(7): p. 1742-1754.
101. Mahmoud, A.B., et al., *Influenza Virus Targets Class I MHC-Educated NK Cells for Immunevasion*. PLoS Pathog, 2016. **12**(2): p. e1005446.
102. Boudreau, J.E., et al., *KIR3DL1 and HLA-B Density and Binding Calibrate NK Education and Response to HIV*. J Immunol, 2016. **196**(8): p. 3398-410.
103. Sim, M.J., et al., *Canonical and Cross-reactive Binding of NK Cell Inhibitory Receptors to HLA-C Allotypes Is Dictated by Peptides Bound to HLA-C*. Front Immunol, 2017. **8**: p. 193.
104. Fadda, L., et al., *Peptide antagonism as a mechanism for NK cell activation*. Proc Natl Acad Sci U S A, 2010. **107**(22): p. 10160-5.
105. Strong, R.K., et al., *HLA-E allelic variants. Correlating differential expression, peptide affinities, crystal structures, and thermal stabilities*. J Biol Chem, 2003. **278**(7): p. 5082-90.
106. Nersesian, S., et al., *Killer instincts: natural killer cells as multifactorial cancer immunotherapy*. Front Immunol, 2023. **14**: p. 1269614.
107. Lin, Z., et al., *HLA class I signal peptide polymorphism determines the level of CD94/NKG2-HLA-E-mediated regulation of effector cell responses*. Nat Immunol, 2023. **24**(7): p. 1087-1097.
108. Chaplin, D.D., *Overview of the immune response*. J Allergy Clin Immunol, 2010. **125**(2 Suppl 2): p. S3-23.

109. Medina, K.L., *Overview of the immune system*. Handb Clin Neurol, 2016. **133**: p. 61-76.
110. Bonilla, F.A. and H.C. Oettgen, *Adaptive immunity*. J Allergy Clin Immunol, 2010. **125**(2 Suppl 2): p. S33-40.
111. Dong, D., et al., *Structural basis of assembly of the human T cell receptor-CD3 complex*. Nature, 2019. **573**(7775): p. 546-552.
112. Shah, K., et al., *T cell receptor (TCR) signaling in health and disease*. Signal Transduct Target Ther, 2021. **6**(1): p. 412.
113. Morath, A. and W.W. Schamel, *alphabeta and gammadelta T cell receptors: Similar but different*. J Leukoc Biol, 2020. **107**(6): p. 1045-1055.
114. Ribot, J.C., N. Lopes, and B. Silva-Santos, *gammadelta T cells in tissue physiology and surveillance*. Nat Rev Immunol, 2021. **21**(4): p. 221-232.
115. Wieczorek, M., et al., *Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation*. Front Immunol, 2017. **8**: p. 292.
116. Gascoigne, N.R., et al., *TCR Signal Strength and T Cell Development*. Annu Rev Cell Dev Biol, 2016. **32**: p. 327-348.
117. Jung, D. and F.W. Alt, *Unraveling V(D)J recombination; insights into gene regulation*. Cell, 2004. **116**(2): p. 299-311.
118. Bassing, C.H., W. Swat, and F.W. Alt, *The mechanism and regulation of chromosomal V(D)J recombination*. Cell, 2002. **109** Suppl: p. S45-55.
119. Krangel, M.S., *Mechanics of T cell receptor gene rearrangement*. Curr Opin Immunol, 2009. **21**(2): p. 133-9.
120. Sewell, A.K., *Why must T cells be cross-reactive?* Nat Rev Immunol, 2012. **12**(9): p. 669-77.
121. Hopkins, J.R., et al., *Unconventional modes of peptide-HLA-I presentation change the rules of TCR engagement*. Discov Immunol, 2022. **1**(1): p. kyac001.
122. Bradley, P. and P.G. Thomas, *Using T Cell Receptor Repertoires to Understand the Principles of Adaptive Immune Recognition*. Annu Rev Immunol, 2019. **37**: p. 547-570.
123. Kersh, G.J. and P.M. Allen, *Structural basis for T cell recognition of altered peptide ligands: a single T cell receptor can productively recognize a large continuum of related ligands*. J Exp Med, 1996. **184**(4): p. 1259-68.
124. Robins, H.S., et al., *Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells*. Blood, 2009. **114**(19): p. 4099-107.
125. Qi, Q., et al., *Diversity and clonal selection in the human T-cell repertoire*. Proc Natl Acad Sci U S A, 2014. **111**(36): p. 13139-44.
126. Sethna, Z., et al., *Insights into immune system development and function from mouse T-cell repertoires*. Proc Natl Acad Sci U S A, 2017. **114**(9): p. 2253-2258.
127. Lythe, G., et al., *How many TCR clonotypes does a body maintain?* J Theor Biol, 2016. **389**: p. 214-24.
128. Lythe, G. and C. Molina-Paris, *Some deterministic and stochastic mathematical models of naive T-cell homeostasis*. Immunol Rev, 2018. **285**(1): p. 206-217.
129. de la Higuera, L., et al., *Fate of a Naive T Cell: A Stochastic Journey*. Front Immunol, 2019. **10**: p. 194.

130. Kurd, N. and E.A. Robey, *T-cell selection in the thymus: a spatial and temporal perspective*. Immunol Rev, 2016. **271**(1): p. 114-26.
131. Irla, M., *Instructive Cues of Thymic T Cell Selection*. Annu Rev Immunol, 2022. **40**: p. 95-119.
132. Soto, C., et al., *High Frequency of Shared Clonotypes in Human T Cell Receptor Repertoires*. Cell Rep, 2020. **32**(2): p. 107882.
133. Venturi, V., et al., *Sharing of T cell receptors in antigen-specific responses is driven by convergent recombination*. Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18691-6.
134. Venturi, V., et al., *The molecular basis for public T-cell responses?* Nat Rev Immunol, 2008. **8**(3): p. 231-8.
135. Murugan, A., et al., *Statistical inference of the generation probability of T-cell receptors from sequence repertoires*. Proc Natl Acad Sci U S A, 2012. **109**(40): p. 16161-6.
136. Madi, A., et al., *T-cell receptor repertoires share a restricted set of public and abundant CDR3 sequences that are associated with self-related immunity*. Genome Res, 2014. **24**(10): p. 1603-12.
137. Rechavi, E., et al., *Timely and spatially regulated maturation of B and T cell repertoire during human fetal development*. Sci Transl Med, 2015. **7**(276): p. 276ra25.
138. Park, J.E., et al., *Prenatal development of human immunity*. Science, 2020. **368**(6491): p. 600-603.
139. Velardi, E., J.J. Tsai, and M.R.M. van den Brink, *T cell regeneration after immunological injury*. Nat Rev Immunol, 2021. **21**(5): p. 277-291.
140. Kawabe, T., et al., *Memory-phenotype CD4(+) T cells spontaneously generated under steady-state conditions exert innate T(H)1-like effector function*. Sci Immunol, 2017. **2**(12).
141. Min, B., et al., *Spontaneous proliferation, a response of naive CD4 T cells determined by the diversity of the memory cell repertoire*. Proc Natl Acad Sci U S A, 2004. **101**(11): p. 3874-9.
142. Troy, A.E. and H. Shen, *Cutting edge: homeostatic proliferation of peripheral T lymphocytes is regulated by clonal competition*. J Immunol, 2003. **170**(2): p. 672-6.
143. Faint, J.M., et al., *Quantitative flow cytometry for the analysis of T cell receptor Vbeta chain expression*. J Immunol Methods, 1999. **225**(1-2): p. 53-60.
144. Salameire, D., et al., *Efficient characterization of the TCR repertoire in lymph nodes by flow cytometry*. Cytometry A, 2009. **75**(9): p. 743-51.
145. Cochet, M., et al., *Molecular detection and in vivo analysis of the specific T cell response to a protein antigen*. Eur J Immunol, 1992. **22**(10): p. 2639-47.
146. Correia-Neves, M., et al., *The shaping of the T cell repertoire*. Immunity, 2001. **14**(1): p. 21-32.
147. Sant'Angelo, D.B., et al., *A molecular map of T cell development*. Immunity, 1998. **9**(2): p. 179-86.
148. De Simone, M., G. Rossetti, and M. Pagani, *Single Cell T Cell Receptor Sequencing: Techniques and Future Challenges*. Front Immunol, 2018. **9**: p. 1638.
149. Goodwin, S., J.D. McPherson, and W.R. McCombie, *Coming of age: ten years of next-generation sequencing technologies*. Nat Rev Genet, 2016. **17**(6): p. 333-51.
150. Six, A., et al., *The past, present, and future of immune repertoire biology - the rise of next-generation repertoire analysis*. Front Immunol, 2013. **4**: p. 413.

151. Hu, T., et al., *Next-generation sequencing technologies: An overview*. Hum Immunol, 2021. **82**(11): p. 801-811.
152. Tickotsky, N., et al., *McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences*. Bioinformatics, 2017. **33**(18): p. 2924-2929.
153. Shugay, M., et al., *VDJdb: a curated database of T-cell receptor sequences with known antigen specificity*. Nucleic Acids Res, 2018. **46**(D1): p. D419-D427.
154. Huang, H., et al., *Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening*. Nat Biotechnol, 2020. **38**(10): p. 1194-1202.
155. Glanville, J., et al., *Identifying specificity groups in the T cell receptor repertoire*. Nature, 2017. **547**(7661): p. 94-98.
156. Dash, P., et al., *Quantifiable predictive features define epitope-specific T cell receptor repertoires*. Nature, 2017. **547**(7661): p. 89-93.
157. Sidhom, J.W., et al., *DeepTCR is a deep learning framework for revealing sequence concepts within T-cell repertoires*. Nat Commun, 2021. **12**(1): p. 1605.
158. Tong, Y., et al., *SETE: Sequence-based Ensemble learning approach for TCR Epitope binding prediction*. Comput Biol Chem, 2020. **87**: p. 107281.
159. Fischer, D.S., et al., *Predicting antigen specificity of single T cells based on TCR CDR3 regions*. Mol Syst Biol, 2020. **16**(8): p. e9416.
160. Katayama, Y., et al., *Machine Learning Approaches to TCR Repertoire Analysis*. Front Immunol, 2022. **13**: p. 858057.
161. Saccardi, R., et al., *Benchmarking of survival outcomes following Haematopoietic Stem Cell Transplantation (HSCT): an update of the ongoing project of the European Society for Blood and Marrow Transplantation (EBMT) and Joint Accreditation Committee of ISCT and EBMT (JACIE)*. Bone Marrow Transplant, 2023. **58**(6): p. 659-666.
162. Giralt, S. and M.R. Bishop, *Principles and overview of allogeneic hematopoietic stem cell transplantation*. Cancer Treat Res, 2009. **144**: p. 1-21.
163. Bacigalupo, A., et al., *Defining the intensity of conditioning regimens: working definitions*. Biol Blood Marrow Transplant, 2009. **15**(12): p. 1628-33.
164. Nagler, A. and A. Shimoni, *Conditioning*, in *The EBMT Handbook: Hematopoietic Stem Cell Transplantation and Cellular Therapies*, E. Carreras, et al., Editors. 2019: Cham (CH). p. 99-107.
165. Storek, J., et al., *Immune reconstitution after allogeneic marrow transplantation compared with blood stem cell transplantation*. Blood, 2001. **97**(11): p. 3380-9.
166. Storek, J., et al., *Reconstitution of the immune system after hematopoietic stem cell transplantation in humans*. Semin Immunopathol, 2008. **30**(4): p. 425-37.
167. Bosch, M., F.M. Khan, and J. Storek, *Immune reconstitution after hematopoietic cell transplantation*. Curr Opin Hematol, 2012. **19**(4): p. 324-35.
168. Jacobs, R., et al., *CD16- CD56+ natural killer cells after bone marrow transplantation*. Blood, 1992. **79**(12): p. 3239-44.

169. Ullah, M.A., G.R. Hill, and S.K. Tey, *Functional Reconstitution of Natural Killer Cells in Allogeneic Hematopoietic Stem Cell Transplantation*. Front Immunol, 2016. **7**: p. 144.
170. Dulphy, N., et al., *An unusual CD56(bright) CD16(low) NK cell subset dominates the early posttransplant period following HLA-matched hematopoietic stem cell transplantation*. J Immunol, 2008. **181**(3): p. 2227-37.
171. Nguyen, S., et al., *NK-cell reconstitution after haploidentical hematopoietic stem-cell transplantations: immaturity of NK cells and inhibitory effect of NKG2A override GvL effect*. Blood, 2005. **105**(10): p. 4135-42.
172. Shilling, H.G., et al., *Reconstitution of NK cell receptor repertoire following HLA-matched hematopoietic cell transplantation*. Blood, 2003. **101**(9): p. 3730-40.
173. Savani, B.N., et al., *Rapid natural killer cell recovery determines outcome after T-cell-depleted HLA-identical stem cell transplantation in patients with myeloid leukemias but not with acute lymphoblastic leukemia*. Leukemia, 2007. **21**(10): p. 2145-52.
174. Vago, L., et al., *Temporal, quantitative, and functional characteristics of single-KIR-positive alloreactive natural killer cell recovery account for impaired graft-versus-leukemia activity after haploidentical hematopoietic stem cell transplantation*. Blood, 2008. **112**(8): p. 3488-99.
175. Triplett, B.M., et al., *Effects of activating NK cell receptor expression and NK cell reconstitution on the outcomes of unrelated donor hematopoietic cell transplantation for hematologic malignancies*. Leukemia, 2009. **23**(7): p. 1278-87.
176. Yu, J., et al., *Breaking tolerance to self, circulating natural killer cells expressing inhibitory KIR for non-self HLA exhibit effector function after T cell-depleted allogeneic hematopoietic cell transplantation*. Blood, 2009. **113**(16): p. 3875-84.
177. Foley, B., et al., *NK cell education after allogeneic transplantation: dissociation between recovery of cytokine-producing and cytotoxic functions*. Blood, 2011. **118**(10): p. 2784-92.
178. Buhmann, L., et al., *Lymphocyte subset recovery and outcome after T-cell replete allogeneic hematopoietic SCT*. Bone Marrow Transplant, 2011. **46**(10): p. 1357-62.
179. Minculescu, L., et al., *Early Natural Killer Cell Reconstitution Predicts Overall Survival in T Cell-Replete Allogeneic Hematopoietic Stem Cell Transplantation*. Biol Blood Marrow Transplant, 2016. **22**(12): p. 2187-2193.
180. Blazar, B.R., G.R. Hill, and W.J. Murphy, *Dissecting the biology of allogeneic HSCT to enhance the GvT effect whilst minimizing GvHD*. Nat Rev Clin Oncol, 2020. **17**(8): p. 475-492.
181. Sweeney, C. and P. Vyas, *The Graft-Versus-Leukemia Effect in AML*. Front Oncol, 2019. **9**: p. 1217.
182. Feldman, H.A., H. Cevik, and S.N. Waggoner, *Negativity begets longevity in T cells*. J Clin Invest, 2023. **133**(12).
183. Malard, F., et al., *Acute graft-versus-host disease*. Nat Rev Dis Primers, 2023. **9**(1): p. 27.
184. Wong, E., et al., *Strategies to enhance the graft versus tumour effect after allogeneic haematopoietic stem cell transplantation*. Bone Marrow Transplant, 2019. **54**(2): p. 175-189.
185. Chang, Y.J., X.Y. Zhao, and X.J. Huang, *Strategies for Enhancing and Preserving Anti-leukemia Effects Without Aggravating Graft-Versus-Host Disease*. Front Immunol, 2018. **9**: p. 3041.

186. Pryce, A., et al., *Genetic Findings of Potential Donor Origin following Hematopoietic Cell Transplantation: Recommendations on Donor Disclosure and Genetic Testing from the World Marrow Donor Association*. Transplant Cell Ther, 2024. **30**(2): p. 143-154.
187. Tiercy, J.M., *How to select the best available related or unrelated donor of hematopoietic stem cells?* Haematologica, 2016. **101**(6): p. 680-7.
188. Petersdorf, E.W., et al., *High HLA-DP Expression and Graft-versus-Host Disease*. N Engl J Med, 2015. **373**(7): p. 599-609.
189. Fleischhauer, K., 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): p. 366-74.
190. Russo, A., et al., *NK cell recovery after haploidentical HSCT with posttransplant cyclophosphamide: dynamics and clinical implications*. Blood, 2018. **131**(2): p. 247-262.
191. Haroun-Izquierdo, A., et al., *Effect of mTOR Inhibition with Sirolimus on Natural Killer Cell Reconstitution in Allogeneic Stem Cell Transplantation*. Transplant Cell Ther, 2023. **29**(6): p. 376 e1-376 e11.
192. Merli, P., et al., *TCRalpha/beta/CD19 depleted HSCT from an HLA-haploidentical relative to treat children with different nonmalignant disorders*. Blood Adv, 2022. **6**(1): p. 281-292.
193. Naik, S. and B.M. Triplett, *Selective depletion of naive T cells by targeting CD45RA*. Front Oncol, 2022. **12**: p. 1009143.
194. Bleakley, M., et al., *Outcomes of acute leukemia patients transplanted with naive T cell-depleted stem cell grafts*. J Clin Invest, 2015. **125**(7): p. 2677-89.
195. Choi, I., et al., *Donor-derived natural killer cells infused after human leukocyte antigen-haploidentical hematopoietic cell transplantation: a dose-escalation study*. Biol Blood Marrow Transplant, 2014. **20**(5): p. 696-704.
196. Choi, I., et al., *Donor-Derived Natural Killer Cell Infusion after Human Leukocyte Antigen-Haploidentical Hematopoietic Cell Transplantation in Patients with Refractory Acute Leukemia*. Biol Blood Marrow Transplant, 2016. **22**(11): p. 2065-2076.
197. Shapiro, R.M., et al., *Expansion, persistence, and efficacy of donor memory-like NK cells infused for posttransplant relapse*. J Clin Invest, 2022. **132**(11).
198. Romee, R., et al., *Cytokine-induced memory-like natural killer cells exhibit enhanced responses against myeloid leukemia*. Sci Transl Med, 2016. **8**(357): p. 357ra123.
199. Song, Y., et al., *IL-12/IL-18-preactivated donor NK cells enhance GVL effects and mitigate GvHD after allogeneic hematopoietic stem cell transplantation*. Eur J Immunol, 2018. **48**(4): p. 670-682.
200. Romee, R., et al., *First-in-human phase 1 clinical study of the IL-15 superagonist complex ALT-803 to treat relapse after transplantation*. Blood, 2018. **131**(23): p. 2515-2527.
201. Ruggeri, L., et al., *Effectiveness of donor natural killer cell alloreactivity in mismatched hematopoietic transplants*. Science, 2002. **295**(5562): p. 2097-100.
202. Zhao, X.Y., et al., *Donor and host coexpressing KIR ligands promote NK education after allogeneic hematopoietic stem cell transplantation*. Blood Adv, 2019. **3**(24): p. 4312-4325.

203. Nguyen, S., et al., *Clinical impact of NK-cell reconstitution after reduced intensity conditioned unrelated cord blood transplantation in patients with acute myeloid leukemia: analysis of a prospective phase II multicenter trial on behalf of the Societe Francaise de Greffe de Moelle Osseuse et Therapie Cellulaire and Eurocord*. Bone Marrow Transplant, 2017. **52**(10): p. 1428-1435.
204. Boudreau, J.E., et al., *KIR3DL1/HLA-B Subtypes Govern Acute Myelogenous Leukemia Relapse After Hematopoietic Cell Transplantation*. J Clin Oncol, 2017. **35**(20): p. 2268-2278.
205. Schetelig, J., et al., *External validation of models for KIR2DS1/KIR3DL1-informed selection of hematopoietic cell donors fails*. Blood, 2020. **135**(16): p. 1386-1395.
206. Schetelig, J., et al., *Haplotype Motif-Based Models for KIR-Genotype Informed Selection of Hematopoietic Cell Donors Fail to Predict Outcome of Patients With Myelodysplastic Syndromes or Secondary Acute Myeloid Leukemia*. Front Immunol, 2020. **11**: p. 584520.
207. Bari, R., et al., *Effect of donor KIR2DL1 allelic polymorphism on the outcome of pediatric allogeneic hematopoietic stem-cell transplantation*. J Clin Oncol, 2013. **31**(30): p. 3782-90.
208. Guethlein, L.A., et al., *Following Transplantation for Acute Myelogenous Leukemia, Donor KIR Cen B02 Better Protects against Relapse than KIR Cen B01*. J Immunol, 2021. **206**(12): p. 3064-3072.
209. Schetelig, J., et al., *Donor KIR genotype based outcome prediction after allogeneic stem cell transplantation: no land in sight*. Front Immunol, 2024. **15**: p. 1350470.
210. van der Ploeg, K., et al., *HLA-A alleles influencing NK cell function impact AML relapse following allogeneic hematopoietic cell transplantation*. Blood Adv, 2020. **4**(19): p. 4955-4964.
211. Wright, P.A., et al., *Donor KIR2DL1 Allelic Polymorphism Influences Posthematopoietic Progenitor Cell Transplantation Outcomes in the T Cell Depleted and Reduced Intensity Conditioning Setting*. Transplant Cell Ther, 2024. **30**(5): p. 488 e1-488 e15.
212. Legrand, N., et al., *Non-Expressed Donor KIR3DL1 Alleles May Represent a Risk Factor for Relapse after T-Replete Haploidentical Hematopoietic Stem Cell Transplantation*. Cancers (Basel), 2023. **15**(10).
213. Dubreuil, L., et al., *Centromeric KIR AA Individuals Harbor Particular KIR Alleles Conferring Beneficial NK Cell Features with Implications in Haplo-Identical Hematopoietic Stem Cell Transplantation*. Cancers (Basel), 2020. **12**(12).
214. Gowdavally, S., et al., *KIR2DS4 and Its Variant KIR1D in KIR-AA Genotype Donors Showed Differential Survival Impact in Patients with Lymphoid Disease after HLA-Matched Unrelated Hematopoietic Stem Cell Transplantation*. Transplant Cell Ther, 2023. **29**(7): p. 457 e1-457 e10.
215. Shaffer, B.C., et al., *Prospective KIR genotype evaluation of hematopoietic cell donors is feasible with potential to benefit patients with AML*. Blood Adv, 2021. **5**(7): p. 2003-2011.
216. Wolf, N.K., D.U. Kissiov, and D.H. Raulet, *Roles of natural killer cells in immunity to cancer, and applications to immunotherapy*. Nat Rev Immunol, 2023. **23**(2): p. 90-105.
217. Cooley, S., et al., *KIR reconstitution is altered by T cells in the graft and correlates with clinical outcomes after unrelated donor transplantation*. Blood, 2005. **106**(13): p. 4370-6.

218. Amorim, L.M., et al., *High-Resolution Characterization of KIR Genes in a Large North American Cohort Reveals Novel Details of Structural and Sequence Diversity*. Front Immunol, 2021. **12**: p. 674778.
219. Pyo, C.W., et al., *Different patterns of evolution in the centromeric and telomeric regions of group A and B haplotypes of the human killer cell Ig-like receptor locus*. PLoS One, 2010. **5**(12): p. e15115.
220. Vierra-Green, C., et al., *Allele-level haplotype frequencies and pairwise linkage disequilibrium for 14 KIR loci in 506 European-American individuals*. PLoS One, 2012. **7**(11): p. e47491.
221. Maxwell, L.D., et al., *A common KIR2DS4 deletion variant in the human that predicts a soluble KIR molecule analogous to the KIR1D molecule observed in the rhesus monkey*. Tissue Antigens, 2002. **60**(3): p. 254-8.
222. Gonzalez-Galarza, F.F., et al., *Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools*. Nucleic Acids Res, 2020. **48**(D1): p. D783-D788.
223. Naiyer, M.M., et al., *KIR2DS2 recognizes conserved peptides derived from viral helicases in the context of HLA-C*. Sci Immunol, 2017. **2**(15).
224. Liu, J., et al., *Activating killer cell immunoglobulin-like receptor 2DS2 binds to HLA-A*11*. Proc Natl Acad Sci U S A, 2014. **111**(7): p. 2662-7.
225. Martin, M.P., et al., *Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS*. Nat Genet, 2002. **31**(4): p. 429-34.
226. Cooley, S., et al., *Donor selection for natural killer cell receptor genes leads to superior survival after unrelated transplantation for acute myelogenous leukemia*. Blood, 2010. **116**(14): p. 2411-9.
227. Krieger, E., et al., *Increased donor inhibitory KIR with known HLA interactions provide protection from relapse following HLA matched unrelated donor HCT for AML*. Bone Marrow Transplant, 2021. **56**(11): p. 2714-2722.
228. Garrod, K.R., et al., *NK cell patrolling and elimination of donor-derived dendritic cells favor indirect alloreactivity*. J Immunol, 2010. **184**(5): p. 2329-36.
229. Saunders, P.M., et al., *The Role of the HLA Class I alpha2 Helix in Determining Ligand Hierarchy for the Killer Cell Ig-like Receptor 3DL1*. J Immunol, 2021. **206**(4): p. 849-860.
230. Burek Kamenaric, M., et al., *The impact of KIR2DS4 gene on clinical outcome after hematopoietic stem cell transplantation*. Hum Immunol, 2017. **78**(2): p. 95-102.
231. Farias, T.D.J., et al., *HLA-DPB1*13:01 associates with enhanced, and KIR2DS4*001 with diminished protection from developing severe COVID-19*. HLA, 2024. **103**(1): p. e15251.
232. Merino, A.M., et al., *KIR2DS4 promotes HIV-1 pathogenesis: new evidence from analyses of immunogenetic data and natural killer cell function*. PLoS One, 2014. **9**(6): p. e99353.
233. Morishima, S., 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): p. 491-8.
234. Boelen, L., et al., *Inhibitory killer cell immunoglobulin-like receptors strengthen CD8(+) T cell-mediated control of HIV-1, HCV, and HTLV-1*. Sci Immunol, 2018. **3**(29).

235. Martinez-Jimenez, F., et al., *Genetic immune escape landscape in primary and metastatic cancer*. Nat Genet, 2023. **55**(5): p. 820-831.
236. Apps, R., et al., *Relative expression levels of the HLA class-I proteins in normal and HIV-infected cells*. J Immunol, 2015. **194**(8): p. 3594-600.
237. Baysoy, A., et al., *The technological landscape and applications of single-cell multi-omics*. Nat Rev Mol Cell Biol, 2023. **24**(10): p. 695-713.
238. Bandura, D.R., et al., *Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry*. Anal Chem, 2009. **81**(16): p. 6813-22.
239. Futamura, K., et al., *Novel full-spectral flow cytometry with multiple spectrally-adjacent fluorescent proteins and fluorochromes and visualization of in vivo cellular movement*. Cytometry A, 2015. **87**(9): p. 830-42.
240. Nolan, J.P. and D. Condello, *Spectral flow cytometry*. Curr Protoc Cytom, 2013. **Chapter 1**: p. 1 27 1-1 27 13.
241. Robinson, J.P., *Spectral flow cytometry-Quo vadimus?* Cytometry A, 2019. **95**(8): p. 823-824.
242. Roederer, M., *Spectral compensation for flow cytometry: visualization artifacts, limitations, and caveats*. Cytometry, 2001. **45**(3): p. 194-205.
243. Bendall, S.C., et al., *Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum*. Science, 2011. **332**(6030): p. 687-96.
244. Robinson, J.P., N. Li, and P.K. Narayanan, *High Throughput-Based Mitochondrial Function Assays by Multi-Parametric Flow Cytometry*. Curr Protoc Cytom, 2015. **73**: p. 9 48 1-9 48 9.
245. Robinson, J.P., et al., *Computational analysis of high-throughput flow cytometry data*. Expert Opin Drug Discov, 2012. **7**(8): p. 679-93.
246. Robinson, J.P., *Flow cytometry: past and future*. Biotechniques, 2022. **72**(4): p. 159-169.
247. Novo, D., G. Gregori, and B. Rajwa, *Generalized unmixing model for multispectral flow cytometry utilizing nonsquare compensation matrices*. Cytometry A, 2013. **83**(5): p. 508-20.
248. R Core Team R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>, 2022.
249. Hahne, F., et al., *flowCore: a Bioconductor package for high throughput flow cytometry*. BMC Bioinformatics, 2009. **10**: p. 106.
250. Van Gassen, S., et al., *FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data*. Cytometry A, 2015. **87**(7): p. 636-45.
251. Hadley, W., *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.
252. Finak, G., W. Jiang, and R. Gottardo, *CytoML for cross-platform cytometry data sharing*. Cytometry A, 2018. **93**(12): p. 1189-1196.
253. Finak, G., *flowWorkspace: Infrastructure for representing and interacting with gated and ungated cytometry data sets*. R package version 4.10.0. <https://www.bioconductor.org/packages/release/bioc/html/flowWorkspace.html>, 2022.
254. Matthews, B.W., *Comparison of the predicted and observed secondary structure of T4 phage lysozyme*. Biochim Biophys Acta, 1975. **405**(2): p. 442-51.

255. Pegram, H.J., et al., *Activating and inhibitory receptors of natural killer cells*. Immunol Cell Biol, 2011. **89**(2): p. 216-24.
256. Mahnke, Y.D., M.H. Beddall, and M. Roederer, *OMIP-029: Human NK-cell phenotypization*. Cytometry A, 2015. **87**(11): p. 986-8.
257. Vanikova, S., A. Koladiya, and J. Musil, *OMIP-080: 29-Color flow cytometry panel for comprehensive evaluation of NK and T cells reconstitution after hematopoietic stem cells transplantation*. Cytometry A, 2022. **101**(1): p. 21-26.
258. Kay, A.W., D.M. Strauss-Albee, and C.A. Blish, *Application of Mass Cytometry (CyTOF) for Functional and Phenotypic Analysis of Natural Killer Cells*. Methods Mol Biol, 2016. **1441**: p. 13-26.
259. Rahim, M.M., *The downside of human natural killer cell diversity in viral infection revealed by mass cytometry*. Ann Transl Med, 2016. **4**(24): p. 546.
260. Chattopadhyay, P.K., et al., *High-Parameter Single-Cell Analysis*. Annu Rev Anal Chem (Palo Alto Calif), 2019. **12**(1): p. 411-430.
261. Barcenilla, H., et al., *Mass Cytometry Identifies Distinct Subsets of Regulatory T Cells and Natural Killer Cells Associated With High Risk for Type 1 Diabetes*. Front Immunol, 2019. **10**: p. 982.
262. Ferrer-Font, L., et al., *High-Dimensional Data Analysis Algorithms Yield Comparable Results for Mass Cytometry and Spectral Flow Cytometry Data*. Cytometry A, 2020. **97**(8): p. 824-831.
263. Jaimes, M.C., et al., *Full spectrum flow cytometry and mass cytometry: A 32-marker panel comparison*. Cytometry A, 2022.
264. Oetjen, K.A., et al., *Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow cytometry*. JCI Insight, 2018. **3**(23).
265. Gadalla, R., et al., *Validation of CyTOF Against Flow Cytometry for Immunological Studies and Monitoring of Human Cancer Clinical Trials*. Front Oncol, 2019. **9**: p. 415.
266. van der Pan, K., et al., *Performance of spectral flow cytometry and mass cytometry for the study of innate myeloid cell populations*. Front Immunol, 2023. **14**: p. 1191992.
267. Kashima, Y., et al., *Intensive single-cell analysis reveals immune-cell diversity among healthy individuals*. Life Sci Alliance, 2022. **5**(7).
268. Lo Tartaro, D., et al., *Molecular and cellular immune features of aged patients with severe COVID-19 pneumonia*. Commun Biol, 2022. **5**(1): p. 590.
269. Leipold, M.D. and H.T. Maecker, *Mass cytometry: protocol for daily tuning and running cell samples on a CyTOF mass cytometer*. J Vis Exp, 2012(69): p. e4398.
270. Spitzer, M.H. and G.P. Nolan, *Mass Cytometry: Single Cells, Many Features*. Cell, 2016. **165**(4): p. 780-91.
271. Bonilla, D.L., G. Reinin, and E. Chua, *Full Spectrum Flow Cytometry as a Powerful Technology for Cancer Immunotherapy Research*. Front Mol Biosci, 2020. **7**: p. 612801.
272. Maecker, H.T. and A. Harari, *Immune monitoring technology primer: flow and mass cytometry*. J Immunother Cancer, 2015. **3**: p. 44.
273. Ornatsky, O., et al., *Highly multiparametric analysis by mass cytometry*. J Immunol Methods, 2010. **361**(1-2): p. 1-20.

274. Brummelman, J., et al., *Development, application and computational analysis of high-dimensional fluorescent antibody panels for single-cell flow cytometry*. Nat Protoc, 2019. **14**(7): p. 1946-1969.
275. Del Zotto, G., et al., *Comprehensive Phenotyping of Human PB NK Cells by Flow Cytometry*. Cytometry A, 2020. **97**(9): p. 891-899.
276. Iyer, A., A.A.J. Hamers, and A.B. Pillai, *CyTOF((R)) for the Masses*. Front Immunol, 2022. **13**: p. 815828.
277. Brodin, P., et al., *Variation in the human immune system is largely driven by non-heritable influences*. Cell, 2015. **160**(1-2): p. 37-47.
278. Olin, A., et al., *Stereotypic Immune System Development in Newborn Children*. Cell, 2018. **174**(5): p. 1277-1292 e14.
279. Pradier, A., et al., *Small-Molecule Immunosuppressive Drugs and Therapeutic Immunoglobulins Differentially Inhibit NK Cell Effector Functions in vitro*. Front Immunol, 2019. **10**: p. 556.
280. Wang, H., et al., *The unexpected effect of cyclosporin A on CD56+CD16- and CD56+CD16+ natural killer cell subpopulations*. Blood, 2007. **110**(5): p. 1530-9.
281. Rambaldi, B., et al., *Impaired T- and NK-cell reconstitution after haploidentical HCT with posttransplant cyclophosphamide*. Blood Adv, 2021. **5**(2): p. 352-364.
282. Willem, C., et al., *Impact of KIR/HLA Incompatibilities on NK Cell Reconstitution and Clinical Outcome after T Cell-Replete Haploidentical Hematopoietic Stem Cell Transplantation with Posttransplant Cyclophosphamide*. J Immunol, 2019. **202**(7): p. 2141-2152.
283. Eissens, D.N., et al., *CD3+/CD19+-depleted grafts in HLA-matched allogeneic peripheral blood stem cell transplantation lead to early NK cell cytolytic responses and reduced inhibitory activity of NKG2A*. Leukemia, 2010. **24**(3): p. 583-91.
284. Pfeiffer, M.M., et al., *Reconstitution of natural killer cell receptors influences natural killer activity and relapse rate after haploidentical transplantation of T- and B-cell depleted grafts in children*. Haematologica, 2010. **95**(8): p. 1381-8.
285. Arase, H., et al., *Direct recognition of cytomegalovirus by activating and inhibitory NK cell receptors*. Science, 2002. **296**(5571): p. 1323-6.
286. Sun, J.C., J.N. Beilke, and L.L. Lanier, *Adaptive immune features of natural killer cells*. Nature, 2009. **457**(7229): p. 557-61.
287. Adams, N.M., et al., *Cytomegalovirus Infection Drives Avidity Selection of Natural Killer Cells*. Immunity, 2019. **50**(6): p. 1381-1390 e5.
288. Park, K.H., et al., *Delayed NK Cell Reconstitution and Reduced NK Activity Increased the Risks of CMV Disease in Allogeneic-Hematopoietic Stem Cell Transplantation*. Int J Mol Sci, 2020. **21**(10).
289. Hassan, N., et al., *CMV reactivation initiates long-term expansion and differentiation of the NK cell repertoire*. Front Immunol, 2022. **13**: p. 935949.
290. Della Chiesa, M., L. Muccio, and A. Moretta, *CMV induces rapid NK cell maturation in HSCT recipients*. Immunol Lett, 2013. **155**(1-2): p. 11-3.
291. Della Chiesa, M., et al., *Impact of HCMV Infection on NK Cell Development and Function after HSCT*. Front Immunol, 2013. **4**: p. 458.

292. Djaoud, Z., et al., *Amplified NKG2C+ NK cells in cytomegalovirus (CMV) infection preferentially express killer cell Ig-like receptor 2DL: functional impact in controlling CMV-infected dendritic cells*. J Immunol, 2013. **191**(5): p. 2708-16.
293. Zaghi, E., et al., *Single-cell profiling identifies impaired adaptive NK cells expanded after HCMV reactivation in haploidentical HSCT*. JCI Insight, 2021. **6**(12).
294. Zuo, W., et al., *The Interaction of HLA-C1/KIR2DL2/L3 Promoted KIR2DL2/L3 Single-Positive/NKG2C-Positive Natural Killer Cell Reconstitution, Raising the Incidence of aGVHD after Hematopoietic Stem Cell Transplantation*. Front Immunol, 2022. **13**: p. 814334.
295. Foley, B., et al., *Cytomegalovirus reactivation after allogeneic transplantation promotes a lasting increase in educated NKG2C+ natural killer cells with potent function*. Blood, 2012. **119**(11): p. 2665-74.
296. Bjorkstrom, N.K., et al., *Rapid expansion and long-term persistence of elevated NK cell numbers in humans infected with hantavirus*. J Exp Med, 2011. **208**(1): p. 13-21.
297. Petitdemange, C., et al., *Unconventional repertoire profile is imprinted during acute chikungunya infection for natural killer cells polarization toward cytotoxicity*. PLoS Pathog, 2011. **7**(9): p. e1002268.
298. Guma, M., et al., *Human cytomegalovirus infection is associated with increased proportions of NK cells that express the CD94/NKG2C receptor in aviremic HIV-1-positive patients*. J Infect Dis, 2006. **194**(1): p. 38-41.
299. Hasan, M.Z., et al., *SARS-CoV-2 infection induces adaptive NK cell responses by spike protein-mediated induction of HLA-E expression*. Emerg Microbes Infect, 2024. **13**(1): p. 2361019.
300. Hendricks, D.W., et al., *Cutting edge: NKG2C(hi)CD57+ NK cells respond specifically to acute infection with cytomegalovirus and not Epstein-Barr virus*. J Immunol, 2014. **192**(10): p. 4492-6.
301. Malone, D.F.G., et al., *Cytomegalovirus-Driven Adaptive-Like Natural Killer Cell Expansions Are Unaffected by Concurrent Chronic Hepatitis Virus Infections*. Front Immunol, 2017. **8**: p. 525.
302. Bjorkstrom, N.K., et al., *Characterization of natural killer cell phenotype and function during recurrent human HSV-2 infection*. PLoS One, 2011. **6**(11): p. e27664.
303. Pradier, A., et al., *Human pegivirus-1 replication influences NK cell reconstitution after allogeneic haematopoietic stem cell transplantation*. Front Immunol, 2022. **13**: p. 1060886.
304. Britanova, O.V., et al., *Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling*. J Immunol, 2014. **192**(6): p. 2689-98.
305. Yoshida, K., et al., *Aging-related changes in human T-cell repertoire over 20years delineated by deep sequencing of peripheral T-cell receptors*. Exp Gerontol, 2017. **96**: p. 29-37.
306. Emerson, R.O., et al., *Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire*. Nat Genet, 2017. **49**(5): p. 659-665.
307. Krishna, C., et al., *Genetic and environmental determinants of human TCR repertoire diversity*. Immun Ageing, 2020. **17**: p. 26.
308. Buhler, S., et al., *Genetic T-cell receptor diversity at 1 year following allogeneic hematopoietic stem cell transplantation*. Leukemia, 2020. **34**(5): p. 1422-1432.

309. Calderin Sollet, Z., et al., *CMV serostatus and T-cell repertoire diversity 5 years after allogeneic hematopoietic stem cell transplantation*. Leukemia, 2023. **37**(4): p. 948-951.
310. Pan, Y.G., et al., *Vaccination reshapes the virus-specific T cell repertoire in unexposed adults*. Immunity, 2021. **54**(6): p. 1245-1256 e5.
311. Morton, S.U., et al., *Premature Infants Have Normal Maturation of the T Cell Receptor Repertoire at Term*. Front Immunol, 2022. **13**: p. 854414.
312. Leick, M., et al., *T Cell Clonal Dynamics Determined by High-Resolution TCR-beta Sequencing in Recipients after Allogeneic Hematopoietic Cell Transplantation*. Biol Blood Marrow Transplant, 2020. **26**(9): p. 1567-1574.
313. Pagliuca, S., et al., *Clinical and basic implications of dynamic T cell receptor clonotyping in hematopoietic cell transplantation*. JCI Insight, 2021. **6**(13).
314. Meier, J.A., et al., *T Cell Repertoire Evolution after Allogeneic Bone Marrow Transplantation: An Organizational Perspective*. Biol Blood Marrow Transplant, 2019. **25**(5): p. 868-882.
315. Pradier, A., et al., *T cell receptor sequencing reveals reduced clonal breadth of T-cell responses against SARS-CoV-2 after natural infection and vaccination in allogeneic hematopoietic stem cell transplant recipients*. Ann Oncol, 2022. **33**(12): p. 1333-1335.
316. Chabannon, C., et al., *Hematopoietic stem cell transplantation in its 60s: A platform for cellular therapies*. Sci Transl Med, 2018. **10**(436).
317. Beziat, V., et al., *NK cell responses to cytomegalovirus infection lead to stable imprints in the human KIR repertoire and involve activating KIRs*. Blood, 2013. **121**(14): p. 2678-88.
318. Grassmann, S. and J.C. Sun, *Thanks for the NK cell memories*. Nat Immunol, 2022. **23**(11): p. 1512-1514.
319. Alter, G., J.M. Malenfant, and M. Altfeld, *CD107a as a functional marker for the identification of natural killer cell activity*. J Immunol Methods, 2004. **294**(1-2): p. 15-22.
320. Goncharov, M., et al., *VDJdb in the pandemic era: a compendium of T cell receptors specific for SARS-CoV-2*. Nat Methods, 2022. **19**(9): p. 1017-1019.
321. Nolan, S., et al., *A large-scale database of T-cell receptor beta (TCRbeta) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2*. Res Sq, 2020.
322. Melendez-Munoz, R., 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**(3): p. 577-586.
323. Goodridge, J.P., et al., *Remodeling of secretory lysosomes during education tunes functional potential in NK cells*. Nat Commun, 2019. **10**(1): p. 514.
324. Pollock, N.R., G.F. Harrison, and P.J. Norman, *Immunogenomics of Killer Cell Immunoglobulin-Like Receptor (KIR) and HLA Class I: Coevolution and Consequences for Human Health*. J Allergy Clin Immunol Pract, 2022. **10**(7): p. 1763-1775.
325. Gaimann, M.U., et al., *Early life imprints the hierarchy of T cell clone sizes*. Elife, 2020. **9**.
326. Suessmuth, Y., et al., *CMV reactivation drives posttransplant T-cell reconstitution and results in defects in the underlying TCRbeta repertoire*. Blood, 2015. **125**(25): p. 3835-50.

327. Mocarski, E.S., et al., *Human cytomegalovirus in a SCID-hu mouse: thymic epithelial cells are prominent targets of viral replication*. Proc Natl Acad Sci U S A, 1993. **90**(1): p. 104-8.
328. Kim, T.S. and E.C. Shin, *The activation of bystander CD8(+) T cells and their roles in viral infection*. Exp Mol Med, 2019. **51**(12): p. 1-9.
329. Soderstrom, A., et al., *T cell receptor excision circles are potential predictors of survival in adult allogeneic hematopoietic stem cell transplantation recipients with acute myeloid leukemia*. Front Immunol, 2022. **13**: p. 954716.
330. Wikstrom, M.E., et al., *Acute GVHD results in a severe DC defect that prevents T-cell priming and leads to fulminant cytomegalovirus disease in mice*. Blood, 2015. **126**(12): p. 1503-14.
331. Montague, Z., et al., *Dynamics of B cell repertoires and emergence of cross-reactive responses in patients with different severities of COVID-19*. Cell Rep, 2021. **35**(8): p. 109173.
332. Dekojova, T., et al., *Dynamic Changes of Inhibitory Killer-Immunoglobulin-Like Receptors on NK Cells after Allogeneic Hematopoietic Stem Cell Transplantation: An Initial Study*. J Clin Med, 2020. **9**(11).
333. Fischer, J.C., et al., *Relevance of C1 and C2 epitopes for hemopoietic stem cell transplantation: role for sequential acquisition of HLA-C-specific inhibitory killer Ig-like receptor*. J Immunol, 2007. **178**(6): p. 3918-23.
334. Stern, L., et al., *Immunoprofiling reveals cell subsets associated with the trajectory of cytomegalovirus reactivation post stem cell transplantation*. Nat Commun, 2022. **13**(1): p. 2603.
335. Horowitz, A., et al., *Regulation of Adaptive NK Cells and CD8 T Cells by HLA-C Correlates with Allogeneic Hematopoietic Cell Transplantation and with Cytomegalovirus Reactivation*. J Immunol, 2015. **195**(9): p. 4524-36.
336. Ito, S., et al., *CMV reactivation is associated with a lower incidence of relapse after allo-SCT for CML*. Bone Marrow Transplant, 2013. **48**(10): p. 1313-6.
337. Green, M.L., et al., *CMV reactivation after allogeneic HCT and relapse risk: evidence for early protection in acute myeloid leukemia*. Blood, 2013. **122**(7): p. 1316-24.
338. Elmaagacli, A.H., et al., *Early human cytomegalovirus replication after transplantation is associated with a decreased relapse risk: evidence for a putative virus-versus-leukemia effect in acute myeloid leukemia patients*. Blood, 2011. **118**(5): p. 1402-12.
339. Manjappa, S., et al., *Protective effect of cytomegalovirus reactivation on relapse after allogeneic hematopoietic cell transplantation in acute myeloid leukemia patients is influenced by conditioning regimen*. Biol Blood Marrow Transplant, 2014. **20**(1): p. 46-52.
340. Diaz-Salazar, C. and J.C. Sun, *Coordinated Viral Control by Cytotoxic Lymphocytes Ensures Optimal Adaptive NK Cell Responses*. Cell Rep, 2020. **32**(12): p. 108186.
341. Prockop, S.E., et al., *Third-party cytomegalovirus-specific T cells improved survival in refractory cytomegalovirus viremia after hematopoietic transplant*. J Clin Invest, 2023. **133**(10).
342. Pei, X.Y., et al., *Comparable anti-CMV responses of transplant donor and third-party CMV-specific T cells for treatment of CMV infection after allogeneic stem cell transplantation*. Cell Mol Immunol, 2022. **19**(4): p. 482-491.
343. Verron, Q., et al., *NK cells integrate signals over large areas when building immune synapses but require local stimuli for degranulation*. Sci Signal, 2021. **14**(684).

344. Al Absi, A., et al., *Actin Cytoskeleton Remodeling Drives Breast Cancer Cell Escape from Natural Killer-Mediated Cytotoxicity*. *Cancer Res*, 2018. **78**(19): p. 5631-5643.
345. Karlsson, F., et al., *Molecular pixelation: spatial proteomics of single cells by sequencing*. *Nat Methods*, 2024.
346. Shyr, D.C., et al., *HLA-haplotype loss after TCRalphabeta/CD19-depleted haploidentical HSCT*. *Bone Marrow Transplant*, 2021. **56**(3): p. 733-737.
347. Vago, L., et al., *Loss of mismatched HLA in leukemia after stem-cell transplantation*. *N Engl J Med*, 2009. **361**(5): p. 478-88.
348. Vago, L., et al., *Genomic loss of mismatched human leukocyte antigen and leukemia immune escape from haploidentical graft-versus-leukemia*. *Semin Oncol*, 2012. **39**(6): p. 707-15.
349. McGranahan, N., et al., *Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution*. *Cell*, 2017. **171**(6): p. 1259-1271 e11.
350. Schaafsma, E., et al., *Pan-cancer association of HLA gene expression with cancer prognosis and immunotherapy efficacy*. *Br J Cancer*, 2021. **125**(3): p. 422-432.
351. Zhao, X., et al., *Targeting neoantigens for cancer immunotherapy*. *Biomark Res*, 2021. **9**(1): p. 61.
352. Das, J. and S.I. Khakoo, *NK cells: tuned by peptide?* *Immunol Rev*, 2015. **267**(1): p. 214-27.
353. Lunemann, S., et al., *Sequence variations in HCV core-derived epitopes alter binding of KIR2DL3 to HLA-C *03:04 and modulate NK cell function*. *J Hepatol*, 2016. **65**(2): p. 252-8.
354. Ziegler, M.C., et al., *HIV-1 induced changes in HLA-C*03 : 04-presented peptide repertoires lead to reduced engagement of inhibitory natural killer cell receptors*. *AIDS*, 2020. **34**(12): p. 1713-1723.
355. Jia, H., et al., *NK cell exhaustion in the tumor microenvironment*. *Front Immunol*, 2023. **14**: p. 1303605.
356. Mortezaee, K., *HHLA2 immune-regulatory roles in cancer*. *Biomed Pharmacother*, 2023. **162**: p. 114639.
357. Palmer, W.H., et al., *Polymorphic KIR3DL3 expression modulates tissue-resident and innate-like T cells*. *Sci Immunol*, 2023. **8**(84): p. eade5343.
358. Wei, Y., et al., *KIR3DL3-HHLA2 is a human immunosuppressive pathway and a therapeutic target*. *Sci Immunol*, 2021. **6**(61).
359. Gielis, S., et al., *Identification of Epitope-Specific T Cells in T-Cell Receptor Repertoires*. *Methods Mol Biol*, 2020. **2120**: p. 183-195.
360. Schober, K., et al., *Reverse TCR repertoire evolution toward dominant low-affinity clones during chronic CMV infection*. *Nat Immunol*, 2020. **21**(4): p. 434-441.
361. Oliveira, G., et al., *Phenotype, specificity and avidity of antitumour CD8(+) T cells in melanoma*. *Nature*, 2021. **596**(7870): p. 119-125.
362. Ruckert, T., et al., *Clonal expansion and epigenetic inheritance of long-lasting NK cell memory*. *Nat Immunol*, 2022. **23**(11): p. 1551-1563.
363. Judge, S.J., W.J. Murphy, and R.J. Canter, *Characterizing the Dysfunctional NK Cell: Assessing the Clinical Relevance of Exhaustion, Anergy, and Senescence*. *Front Cell Infect Microbiol*, 2020. **10**: p. 49.

364. Lakshmikanth, T., et al., *Human Immune System Variation during 1 Year*. Cell Rep, 2020. **32**(3): p. 107923.