Elisa Zaghi

Personal Identifier: F1573378

Master degree in Experimental and Applied Biology

PATTERNS OF NATURAL KILLER CELL IMMUNE-RECONSTITUTION AND IMMUNOLOGICAL MEMORY IN PATIENTS AFFECTED BY HEMATOLOGIC MALIGNANCIES AND UNDERGOING HAPLO-IDENTICAL HEMATOPOIETIC STEM CELL TRANSPLANTATION

A thesis submitted for the degree of

DOCTOR OF PHYLLOSOPHY

The Open University

School of Life, Health and Chemical Science

Affiliated Research Center:

IRCCS Istituto Clinico Humanitas
# LIST OF CONTENTS

1 ABSTRACT .................................................................................................................. 6

2 DECLARATION ........................................................................................................... 8

3 PREFACE .................................................................................................................... 9

4 LIST OF FIGURE AND TABLES ............................................................................. 10

5 ABBREVIATION .......................................................................................................... 12

6 INTRODUCTION .......................................................................................................... 15

6.1 Hematologic Malignancies .................................................................................... 15

6.1.1 Lymphomas ....................................................................................................... 15

6.1.2 Multiple myeloma .............................................................................................. 16

6.1.3 Acute and chronic leukemia .............................................................................. 16

6.1.4 Myeloproliferative and myelodysplastic disorders .......................................... 17

6.2 Hematopoietic stem cell transplant ....................................................................... 17

6.2.1 Donor type ......................................................................................................... 17

6.2.2 Hematopoietic stem cell source ....................................................................... 18

6.2.3 Pre-conditioning protocols ............................................................................... 19

6.3 Haploidentical hematopoietic stem cell transplantation ....................................... 20

6.3.1 The experience of graft manipulation in h-HSCT ............................................. 21

6.3.2 T cell depletion .................................................................................................. 21

6.3.3 T cell repletion .................................................................................................. 22

6.3.4 G-CSF–primed grafts ....................................................................................... 22

6.3.5 Selective T-cell depletion .................................................................................. 23

6.3.6 Post-transplant high dose Cyclophosphamide ............................................... 23

6.4 Factors determining h-HSCT patient clinical outcomes ......................................... 24

6.4.1 Immune Reconstitution .................................................................................... 24

6.4.2 Graft versus Host Disease .............................................................................. 26

6.4.3 Infections .......................................................................................................... 27

6.5 HCMV ..................................................................................................................... 28
6.5.1 HCMV infection and latency ........................................................................................................29
6.5.2 Interplay between HCMV and the immune system .................................................................29
6.6 Natural Killer cells .........................................................................................................................30
6.6.1 NK cell ontogenesis ......................................................................................................................31
6.6.2 Heterogeneity of NK cell subsets .................................................................................................33
   6.6.2.1 cCD56br NK cells ..................................................................................................................34
   6.6.2.2 cCD56dim NK cells .............................................................................................................35
   6.6.2.3 unCD56dim NK cells ...........................................................................................................35
6.6.3 NK cell receptor repertoire and functions ....................................................................................36
6.6.4 NK cell education .........................................................................................................................37
6.6.5 NK cell immune reconstitution after h-HSCT ............................................................................38
6.6.6 Role of NK cells in h-HSCT ........................................................................................................39
6.6.7 NK cells and GvHD .....................................................................................................................39
6.6.8 NK cell shaping upon HCMV infection/reactivation .................................................................41
   6.6.8.1 HCMV immune escape .........................................................................................................41
   6.6.8.2 CD56negCD16pos NK cell subset ........................................................................................42
   6.6.8.3 Memory-like NK cell subset ..................................................................................................43
   6.6.8.4 Memory-like NK cells in h-HSCT patients ..........................................................................46
7 AIM AND RATIONALE ......................................................................................................................48
8 MATERIALS AND METHODS ...........................................................................................................50
8.1 Patients’ recruitment .........................................................................................................................50
8.2 MNC isolation from peripheral blood and bone marrow ............................................................52
8.3 Cell culture ......................................................................................................................................53
   8.3.1 K562 .........................................................................................................................................53
   8.3.2 721.221.G .................................................................................................................................53
   8.3.3 HUVEC .....................................................................................................................................54
   8.3.4 NK cell isolation from PBMCs ..................................................................................................54
   8.3.5 Stromal isolation from chips ......................................................................................................55
8.4 Polychromatic flow cytometry staining, phenotyping and cell sorting ........................................55
8.5 Functional assays ............................................................................................................................57
### 8.6 NK cell subset culture with cytokine stimulation

8.7 RNA isolation and cDNA synthesis

8.8 Real-Time PCR

8.9 RNA sequencing

8.10 Data Analysis

8.10.1 Flowjo

8.10.2 Unsupervised high-dimensional cytometry data analysis

8.10.3 Cluster identification and analysis with PhenoGraph

8.10.4 RNA-seq analysis

8.10.4.1 Pre-processing and quality assessment

8.10.4.2 Mapping, counting, and normalization

8.10.4.3 Variability within the experiments

8.10.4.4 Differential gene expression analysis

8.10.4.5 Pathway analysis

8.10.5 Statistical analysis

### 9 RESULTS

9.1 The early recovery of unCD56<sup>dim</sup>/NKG2A<sup>pos</sup> NK cells in h-HSCT patients

9.1.1 Preferential unCD56<sup>dim</sup> NK cell subset expansion in the first weeks after h-HSCT

9.1.2 unCD56<sup>dim</sup> lymphocytes are bona fide NK cells

9.1.3 unCD56<sup>dim</sup> cells expanded early after h-HSCT show an impaired cytotoxicity

9.1.4 Kinetics and phenotype of highly proliferating unCD56<sup>dim</sup> and cCD56<sup>hi</sup> NK cells

9.2 HCMV infection/activation and NK cells in h-HSCT patients: a flow cytometric point of view

9.2.1 HCMV influence on immune-reconstituting NK cell distribution after h-HSCT

9.2.2 High dimensional single-cell analysis optimization

9.2.3 HCMV reactivation affects the NK cell subset distribution in the t-SNE map

9.2.4 NK cell cluster expansion and their correlation with HCMV

9.2.5 Characterization of NK cell clusters expanded in not reactivated patients

9.2.6 Characterization of NK cell clusters expanded in patients experiencing HCMV infection/activation

9.2.7 HCMV infection/activation after h-HSCT shapes NK cell phenotype
9.2.8 Identification of a peculiar NK cell subset expanded upon HCMV infection/reactivation 88

9.3 HCMV infection/reactivation and NK cells in h-HSCT patients: molecular signature .......... 91

9.3.1 FACS-sorted NK cell subsets from h-HSCT patients ................................................. 92

9.3.2 Gene expression pattern between CD158b1b2jneg and CD158b1b2jpos NK cells of h-HSCT patients experiencing HCMV infection/reactivation .............................................................. 94

9.3.3 D-/R- vs R0: primary HCMV exposure during life does not impact on NK cell features after h-HSCT ........................................................................................................................................ 96

9.3.4 R1 vs R2: anti-viral treatment can influence the expansion of CD56neg NK cells .......... 97

9.3.5 Gene expression pattern between not reactivated patients and recipients experiencing HCMV infection/reactivation .............................................................................................................. 98

9.3.5.1 GSEA ................................................................................................................................. 99

9.3.5.2 IPA .................................................................................................................................. 104

9.3.6 PD-1 and LAG3 as potential immune-checkpoints to rouse ml-NK cells .................... 106

10 DISCUSSION ...................................................................................................................... 108

10.1 The early expansion of anergic unCD56dim/NKG2Apos NK cells represents a therapeutic target in h-HSCT ............................................................................................................................................. 108

10.2 Adaptive ml-NK cells in h-HSCT patients ........................................................................ 110

11 BIBLIOGRAPHY ............................................................................................................... 115
1 ABSTRACT

Haploidentical hematopoietic stem cell transplantation (h-HSCT) represents a promising therapeutic approach to cure patients affected by hematologic malignancies. Despite the positive results in terms of overall/disease free survival and reduced tumor relapse, the HLA mismatch between donor and recipient has not been yet fully exploited and the clinical outcome of h-HSCT patients is still hampered by life-threatening side effects, including relapse and viral infections.

In this context, immune cell reconstitution (IR) is certainly key in determining a positive h-HSCT clinical outcome. We demonstrated that Natural Killer (NK) cells represent the first innate lymphocytes recovering after h-HSCT, thus highlighting their role in ensuring a prompt alloreactivity early after the transplant as well as a rapid protection against opportunistic viral infections.

We characterized the predominant and temporary expansion, early after h-HSCT, of a donor-derived unconventional subset of CD56\textsuperscript{dim}CD16\textsuperscript{neg} (unCD56\textsuperscript{dim}) NK cells that is poorly represented in healthy donors. The expression of the inhibitory CD94/NKG2A receptor on all unCD56\textsuperscript{dim} NK cells is associated, at least in part, with NK cell impairment. Indeed, only masking CD94/NKG2A, we observed a recovery of the cytolytic capacity of this NK cell subset. Thus, this inhibitory receptor could represent a potential immunotherapeutic target to improve NK cell alloreactivity after h-HSCT, ensuring a prompt graft versus leukemia (GvL) effect and preventing disease relapse.

However, given the NK cell impairment and the late recovery of adaptive immune cells are responsible for a prolonged immunodeficiency, the increased risk in developing opportunistic infections is high early after h-HSCT. Among them, the human Cytomegalovirus (HCMV) infection/reactivation represents one of the major causes of morbidity and mortality after h-HSCT. In turn, HCMV infection/reactivation can greatly impact NK cell maturation and effector-functions, by providing a rapid expansion of mature, long-lived and hyper-functional NK cells showing memory-like (ml) properties, with unknown origin and unspecific phenotype.

With the aim to investigate the impact of HCMV infection/reactivation on NK cell IR after h-HSCT, we identified a subset of NK cells showing a peculiar CD158b1b2\textsuperscript{pos}NKG2A\textsuperscript{neg}NKG2C\textsuperscript{pos}NKp30\textsuperscript{low} phenotype that is expanded only in h-HSCT patients experiencing HCMV reactivation. Interestingly, this latter NK cell population showing
memory-like functional features is maintained even after the resolution of the infection and its frequency positively correlates with HCMV viral load and with the numbers of reactivation events. A deeper analysis of our putative ml-NK cell molecular fingerprint revealed that these cells are characterized by an impaired effector-functions, although retaining the capability to produce IFN-γ at transcriptional level.

These findings show in a human setting *in vivo* the expansion and the kinetic of those NK cells that "remember" the HCMV challenge in patients experiencing viral reactivation. Our data are important to better understand the ability of NK cells to control this life-threatening infection after h-HSCT as well as to exploit their alloreactive responses against tumor burden.
2 DECLARATION

The work described in this dissertation was performed at the IRCCS Clinical and Research Institute Humanitas, between February 2016 and January 2020. I declare that this dissertation has not been submitted in part or in whole to any other academic institution. The work reported here was entirely carried out by the author, unless otherwise indicated. Part of the results included in this dissertation have been already published in Haematologica journal.


3 PREFACE

Publications obtained during the course of this thesis:


Publications not related to the current thesis:

4 LIST OF FIGURE AND TABLES

Figure I.1: main interleukin functions in NK cell physiology
Figure I.2: NK cells and GvHD
Figure I.3: adaptive ml-NK cell development
Figure I.4: HCMV impact on NK cells
Figure M.1: Baltimore protocol used to treat patients affected by hematologic malignancies
Figure M.2: flow cytometric analysis of immune reconstituting NK cells in patients after h-HSCT
Figure M.3: RNA-seq workflow
Figure M.4: high dimensional single-cell analysis of immune reconstituting NK cells in patients after h-HSCT
Figure R.1: kinetic of NK cell subset IR after h-HSCT
Figure R.2: unCD56dim NK cell marker expression
Figure R.3: cytotoxicity of unCD56dim NK cells expressing CD94/NKG2A in patients after h-HSCT
Figure R.4: cytotoxicity of NK cell subsets with the masking anti-CD94/NKG2A mAb (Y9)
Figure R.5: CD56 expression of FACS-sorted NK cell subsets upon IL-15+IL-18 stimulation
Figure R.6: culture of sorted NK subsets from HCs with cytokines
Figure R.7: patients-derived FACS-sorted NK cell subset proliferation upon IL-15+IL-18 stimulation.
Figure R.8: frequencies of NKp46 and NKG2A positive cells in in vitro derived NK subsets
Figure R.9: immune reconstituting NK cell subset distribution in h-HSCT patients experiencing HCMV reactivation.
Figure R.10: NK cell subsets displayed on t-SNE maps
Figure R.11: NK cell cluster distribution on t-SNE maps after h-HSCT
Figure R.12: characterization of NK cell clusters enriched in the absence of HCMV reactivation at 3-4 months after h-HSCT
Figure R.13: characterization of NK cell clusters enriched upon HCMV reactivation at 8-12 months after h-HSCT
Figure R.14: correlation analysis of NK cell marker expression and HCMV viremia
Figure R.15: NK cell marker expression at different time points after h-HSCT.
Figure R.16: CD158b1b2jposNKG2AnegNKG2aposCNKp30low NK cells are expanded after HCMV infection/reactivation
Figure R.17: multiple reactivation events influence ml-NK cell phenotype
Figure R.18: kinetics of HCMV specific NK cell subsets in NR and R patients
**Figure R.19**: differentially expressed genes in sorted CD158b1b2jneg and CD158b1b2jpos NK cells of R patients

**Figure R.20**: gene expression profile of CD158b1b2jneg vs CD158b1b2jpos in R h-HSCT patients

**Figure R.21**: comparison between sorted CD158b1b2jpos NK cells in D-/R- and R0 patients

**Figure R.22**: differentially expressed genes in sorted CD158b1b2jpos NK cells from D-/R- and R0 patients

**Figure R.23**: comparison between sorted CD158b1b2jpos NK cells in R1 and R2 patients

**Figure R.24**: differentially expressed genes in sorted CD158b1b2jpos NK cells NR and R pts

**Figure R.25**: gene expression patterns in sorted CD158b1b2jpos NK cells from pts experiencing or not HCMV reactivation

**Figure R.26**: gene expression of signaling proteins in sorted CD158b1b2jpos NK cells derived from R patients resembling those of ml-NK cells

**Figure R.27**: gene expression patterns in sorted CD158b1b2jpos NK cells from patients (pts) experiencing HCMV reactivation

**Figure R.28**: IPA gene expression patterns in sorted CD158b1b2jpos NK cells from patients (pts) experiencing HCMV reactivation

**Figure R.29**: PD-1 and LAG3 checkpoint inhibitors

**Table 1**: patient (pt) clinical data

**Table 2**: Lists of antibodies used in flow cytometry panels

**Table 3**: FACS-sorted samples
5 ABBREVIATION

ADCC: antibody-dependent cell cytotoxicity;
Ag: antigen;
aGvHD: acute graft versus host disease;
ALL: acute lymphoblastic leukemia;
Allo-HSCT: allogeneic hematopoietic stem cell transplantation;
AML: acute myelogenous leukemia;
aNKR: activating Natural Killer cell receptor;
APC: antigen presenting cell;
ATG: antithymocyte globulin;
BM: bone marrow;
CD: cluster of differentiation;
cDNA: complementary DNA;
cGvHD: chronic graft versus host disease;
CFSE: carboxyfluorescein succinimidyl ester;
CLL: chronic lymphocytic;
CLP: common lymphoid progenitors;
CML: chronic myelogenous leukemia;
CSP: cyclosporine;
D/R: donor/recipient;
DCs: dendritic cells;
DMEM: Dulbecco's Modified Eagle medium;
DMSO: dimethyl sulfoxide;
EMEM: Eagle's Minimum Essential medium;
FBS: fetal bovine serum;
Fc: fragment crystallizable;
FCS: flow cytometry standard;
FLT3L: Fms-related tyrosine kinase 3 ligand;
G-CSF: granulocyte colony-stimulating factor;
γδ: gamma delta;
GSEA: gene set enrichment analysis;
GvHD: graft versus host disease;
GvL: graft-versus-leukemia;
HBSS: Hank's Balanced Salt Solution;
HCMV: human cytomegalovirus;
HC: healthy control;
HD: healthy donor;
h-HSCT: haploidentical hematopoietic stem cell transplantation;
HHV-5: human herpes virus 5;
HIV: human immunodeficiency virus;
HL: hodgkin lymphomas;
HLA: human leukocyte antigen;
HSC: hematopoietic stem cell;
HSCT: hematopoietic stem cell transplantation;
IFN: interferon;
Ig: immunoglobulin;
IL: interleukin;
ILC: innate lymphoid cell;
IMDM: Iscove's Modified Dulbecco's medium;
iMFI: median of fluorescence intensity;
iNKR: inhibitory Natural Killer cell receptor;
IR: immune reconstitution;
ITAM: immunoreceptor tyrosine-based activating motif;
ITIM: immunoreceptor tyrosine-based inhibitory motif;
KIR: killer Ig-like Receptor;
Lin: lineage;
MA: myeloablative;
mAb: monoclonal antibody;
MCMV: murine cytomegalovirus;
MDS: myelodysplastic syndromes;
MHC: human major histocompatibility complex;
ml-NK: memory-like Natural Killer cell;
MM: multiple myeloma;
MMF: mycophenolate mofetil;
MMUD: mismatched unrelated donor;
MPD: chronic myeloproliferative disorders;
MTX: methotrexate;
MUD: matched unrelated donor;
NaCl: sodium chloride physiological solution;
NCAM: neural cell adhesion molecule;
NCR: natural cytotoxic receptor;
NGS: next generation sequencing;
NHL: non-hodgkin lymphoma;
NK: Natural Killer;
NKp: NK cell progenitor;
NMA: non myeloablative;
OS: overall survival;
PB: peripheral blood;
PBMCs: peripheral blood mononuclear cells;
PLZF: promyelocytic leukemia zinc finger;
PT-Cy: post-transplant high-dose cyclophosphamide;
RAG: recombination activating gene;
RIC: reduced-intensity conditioning;
RNA-seq: RNA-sequencing;
RPMM: roswell park memorial institute;
SCF: stem cell factor;
SCID: severe combined immune-deficiency;
SYK: spleen tyrosine kinase;
TBI: total body irradiation;
TCD: T-cell depleted;
TCRep: T cell replete;
Th: T-helper;
TKIs: tyrosine kinase inhibitors;
TNF: tumor necrosis factor;
Treg: regulatory T cell;
TRM: transplant-related mortality;
TRT: transplant-related toxicity;
t-SNE: t-stochastic neighbor embedding;
UCB: umbilical cord;
WHO: world health organization.
6 INTRODUCTION

6.1 Hematologic Malignancies

The term “hematologic malignancies” gathers different neoplastic diseases of the hematopoietic and lymphoid tissues. They comprise diseases of both myeloid and lymphoid origin beginning in blood-forming tissues or in cells of the immune system, respectively. In 2008 the World Health Organization (WHO) published a classification of neoplasms of the hematopoietic and lymphoid tissues, which was subsequently updated and revised in 2016 considering the major advances in the field and their clinical and biologic implications. This classification represents a worldwide consensus on the diagnosis of blood malignancies [1, 2]. Based on different morphological, immunophenotypic, genetic and clinical features the hematologic malignancies include: i) Hodgkin and non-Hodgkin lymphoma; ii) Multiple myeloma; iii) Acute and chronic leukemia; iv) Myeloproliferative and myelodysplastic disorders.

6.1.1 Lymphomas

The first group of hematologic malignancies to be described were lymphomas, heterogeneous diseases that develop from mature lymphocytes in the lymphatic system and might spread to solid tissues as lungs and liver [3]. Lymphomas classification is based on the lineage they derived from (B, T, or Natural Killer (NK) cells) and the disease subtypes are defined through their morphology, immunophenotype, and genetic/clinical features [4]. Historically, the first neoplasm identified was Hodgkin Lymphomas (HL) studied by T. Hodgkin in 1832 in patients suffering of lymph nodes and spleen enlargement, cachexia and fatal termination [5]. Other lymphatic malignancies that account for the 90% of the cases, were then called Non-Hodgkin Lymphomas (NHL) and this distinction has been maintained to this day. HL has a bimodal disease distribution where the peak of incidence is reached in patients in their early 20s, it declines through the middle age and it increases again with advanced age [6]. NHL is more common in aged people (over 65 years) and consists in many biologically distinct lymphoid malignancies. Indeed, the revised WHO classification defined more than 50 subtypes of NHL showing different clinical presentations [2].
6.1.2 Multiple myeloma

Multiple myeloma (MM) represents the 10% of hematologic malignancies and it is characterized by a clonal expansion of plasma cells. It is considered an elderly malignant disease since the median age of onset is 66 years. MM evolves from a pre-malignant condition defined as monoclonal gammopathy of undetermined significance, which is mostly asymptomatic, associated to a risk of MM progression of 10% per year. The most common symptoms are bone pain and fatigue, more pronounced in the 75% of patients experiencing also anemia. Many factors including stage of tumor burden, cytogenetic abnormalities and response to therapy might affect the severity of MM. So far, no permanent cures are available to treat MM, but the combination of chemotherapy with new agents, such as thalidomide or lenalidomide, resulted in longer progression-free survival benefits [7, 8].

6.1.3 Acute and chronic leukemia

Leukemia is an uncontrolled proliferation of hematopoietic stem cells in the bone marrow (BM). The development of multidisciplinary approaches allowed the identification of four subtypes of leukemia more frequently encountered: acute lymphoblastic (ALL), acute myelogenous (AML), chronic lymphocytic (CLL), and chronic myelogenous (CML). Acute leukemia, in particular ALL, is commonly diagnosed in childhood, while the other subtypes are more frequent in adults. The onset of the acute diseases is suspected when blood count reveals a surplus of blast cells (mainly self-renewing stem cell or progenitors), which have to be confirmed and characterized by immunophenotyping and cytogenetic testing. Peripheral blood (PB) smear or BM specimen overpopulated of blast cells that are considered the precursors of hematopoietic cells, are suspicious of acute disease. Contrary, chronic leukemias are more common in adults and are characterized by a clonal expansion of at least 5,000 circulating B lymphocytes per μL (CLL) or by the rising white blood cells levels followed by disease progression (CML). Recently, tyrosine kinase inhibitors (TKIs) are considered as revolutionary drugs for the CML disease course [9]. Of note, treatment including chemotherapy, radiation, monoclonal antibodies, or hematopoietic stem cell transplantation depends on patient co-morbidities, age, and leukemia subtypes [10].
6.1.4 Myeloproliferative and myelodysplastic disorders

Chronic Myeloproliferative Disorders (MPD) are characterized by effective hematopoiesis showing increased levels of one or more of the circulating myeloid cell types as erythrocytes, platelets, and granulocytes.

The myelodysplastic syndromes (MDS), on contrary, result in decreased levels of one or more of the circulating myeloid cell types possibly associated with dysplastic signs and increased medullary and peripheral blasts. MDS are pre-cancerous diseases that can evolve, as MPD in AML. For both MPD and MDS, the median age at diagnosis is about 70 years. Both malignancies can be treated with supportive care measures, immunomodulatory agents, and epigenetic agents that, however, are not curative [3].

6.2 Hematopoietic stem cell transplant

Introduced in the early 1970 [11], hematopoietic stem cell transplantation (HSCT) has received great interest as a curative strategy to treat a wide range of acquired and inherited malignant and nonmalignant hematologic diseases together with genetic disorders including hemoglobinopathies and inborn errors of metabolism [12]. HSCT is defined as the intravenous infusion of healthy hematopoietic stem and progenitor cells from a donor in order to replace the entire hematopoietic and immune system of the recipient, re-establishing BM and immune function [13, 14]. Many protocols have been developed in clinic over the time and they differ in terms of: i) donor type; ii) hematopoietic stem cells (HSC) source; iii) conditioning regimen before transplant [15].

6.2.1 Donor type

One of the most important factors influencing both the selection of healthy donors and HSCT outcomes is represented by the polymorphism of the Human Leukocyte Antigen (HLA) genes [16].

HLA is the human major histocompatibility complex (MHC), whose antigens can be classified in class I (HLA-A, -B, -C) and class II (HLA-DRB1, -DQB1, -DPB1). The genes for HLA are located on chromosome 6 and are inherited as haplotypes [13].

Taking into account histocompatibility issues, there are three main HSCTs according to the donor type:
- Autologous, when stem cells are withdrawn from the same patient in need;
- Syngenic, when the donor is a genetically identical twin of the recipient. HLA-matched identical sibling is therefore defined as the best source of HSCs that unfortunately is available only for the 25% of patients needing an allograft [17];
- Allogeneic (allo-HSCT), when HSCs are obtained from related or unrelated donor source. In this case, one of the major serious complication, is represented by Graft versus host disease (GvHD), which is the manifestation of a reaction of donor’s immune system against recipient tissues, perceived as foreign [18].

In the absence of syngenic donors, other HSC sources for allo-HSCT have been evaluated [19]:
- matched unrelated donors (MUD) are unrelated individuals that satisfy match criteria as siblings;
- mismatched unrelated donors (MMUD), individuals with major or minor mismatches in HLA type;
- mismatched related or haploidentical donors (h-HSCT donors) which are partially matched family members.

In selecting the best donor, it has been reported that donor-specific HLA antibodies are associated with high risk of graft failure in non-matched siblings, and it might translate in reduced overall survival (OS) [20-22].

6.2.2 Hematopoietic stem cell source

Donor HSC sources can be BM, mobilized peripheral blood stem cells (obtained by leukapheresis), and placental blood collected through the umbilical cord (UCB) [19].

BM is usually harvested from the posterior iliac crests, but anterior iliac crest and sternum can be used when more cells are required. Once BM has been collected, it can be immediately infused of stored at 4°C or cryopreserved without loss of cell viability [23].

PB stem cells can be easily withdrawn and they are widely used alternatively to BM harvest. First evidence in animal models showed the presence of circulating hematopoietic progenitors in blood and prompt the use of this HSC source also in humans [24].

However, the paucity of hematopoietic precursors in blood at steady state rendered leukapheresis necessary to reach adequate collection of cells after being mobilized from BM through the administration of granulocyte colony-stimulating factor (G-CSF) [23].
Many studies have been conducted defining advantages and drawbacks between BM and PB transplantation [24], with particular attention to cell immune reconstitution and the occurrence of side effects as GvHD. Despite PB-derived HSCs engraft very quickly in respect with BM-derived stem cells, much debates concerning the long-term safety after allo-HSCT still remain.

Lastly, UCB contains high numbers of HSCs able to quickly reconstitute the hematopoietic system of the recipient. Cells from UCB can be safely collected at the time of birth from the umbilical cord vein and cryopreserved. UCB overcomes HLA barriers thanks to the relative naivety of the newborn immune system, thus a full HLA matching is not required [13]. Unfortunately, UCB provide almost a tenth of the CD34^{pos} HSCs present in the adult BM [25]. Given the high advantage of a reduced HLA compatibility, different approaches to overcome the low stem cell content have been used, including the infusion of two units of cord blood into a single recipient [26], and the *ex vivo* expansion of cord blood units [27].

### 6.2.3 Pre-conditioning protocols

In order to allow donor stem cell engraftment and to minimize the risk of rejection, thus ensuring a successful HSCT, host stem cells have to be eradicated from BM niche or suppressed in their proliferative growth and functions. Therefore, different pre-transplant regimens, presenting the advantage to counteract malignant cells overcoming therapy resistance, are currently in use. The term “conditioning” refers to the preparative regimen administered to patients before the intravenous infusion of donor allogeneic stem cells into bloodstream.

Many factors have to be taken into account to choose the right conditioning regimen. Firstly, the type of donor: in autologous HSCT, it usually consists in chemotherapy alone; while in allo-HSCT often Antithymocyte globulin (ATG) is included in the pre-therapy [28]. Moreover, besides the diseases patients are afflicted with and the type of donor, also age, co-morbidities, and specific organ-toxicity are influencing factors in governing pre-conditioning regimen selection [28].

The classification of conditioning regimens is based on the duration of the treatment-induced cytopenia and the demand of stem cell support: myeloablative (MA), non myeloablative (NMA), and reduced-intensity conditioning (RIC) [15].
MA conditioning consists in a high-dose intensity strategy that causes an irreversible pancytopenia in patients, thus making donor-derived HSC support necessary. MA strategies include often total body irradiation (TBI) alone or in combination with chemotherapy inducing an effective immune-suppression in recipients prior to HSCT, which better allows engraftment of donor hematopoietic stem cells [14, 29] and it functions as anticancer therapy too [7]. Although MA conditioning provides rapid donor cell engraftment, it is also associated with transplant-related toxicity (TRT) and mortality (TRM), limiting the use of this regimen to younger patients. High-dose regimens, however, are still used for patients with more aggressive malignant diseases needing strong anti-tumor responses [14].

NMA and RIC treatments have been developed in order to reduce toxicity, while maintaining the anti-tumor effect and gradually replace the host lympho-hematopoietic compartment [30]. NMA is considered a regimen that causes limited cytopenia, lower toxicity than MA one, and does not require donor cell support [15]. RIC regimens induce a wide and reversible cytopenia preventing the high toxicity of MA conditioning allowing the extension of the HSCT procedure, in particular h-HSCT, to elderly patients, avoiding also transplant complications [16, 31].

One of the barriers to RIC is the increased risk of graft rejection due to the low intensity of the regimen. Hence, to prevent rejection, post-transplant immune-suppressive drugs have been introduced along with low-dose TBI-based pre-conditioning. The effects of fludarabine, cyclosporine (CSP), and mycophenolate mofetil (MMF) have been deeply investigated [30, 32-35].

6.3 Haploidentical hematopoietic stem cell transplantation

In the absence of HLA-identical sibling as stem cell source, h-HSCT related donors represent a good alternative immediately available and accessible worldwide. Indeed, virtually all patients have at least one HLA-partially matched donor among family members [17]. According to an estimate, there are 2,7 h-HSCT related donors available for each patient in need [36]. Moreover, h-HSCT related donors allow to easily have disposability of stem cells for additional therapies after the transplantation if required.

First attempts of h-HSCT were initially associated to graft rejection, TRM, and severe GvHD mainly resulting from delayed immune reconstitution (IR) after transplantation which
leaded also to an increased risk in developing opportunistic infections [37, 38]. Despite HLA-mismatched between h-HSCT donor and recipient is responsible for GvHD onset, it leads to an increased Graft-versus-Leukemia (GvL) effect eradicating malignant residual cells, thus favoring good clinical outcomes in patients. The improvements in conditioning regimens combined with new strategies able to reduce the onset of these life-threatening side effects and to overcome immunologic HLA-barriers render h-HSCT one of the preferred platforms of transplantation in the recent years [39].

6.3.1 The experience of graft manipulation in h-HSCT

Basic graft manipulation is a procedure performed to define cellular composition of stem cell sources and it can include depletion of erythrocytes and volume reduction. Further complex graft manipulation techniques have been extensively used after the introduction of immune magnetic-cell separator devices and they are used to optimize cellular composition and to provide a specific transplant regimen. Notably, many h-HSCT platforms have been developed over time to limit donor-derived T cell alloreactivity, main responsible for GvHD hampering patient clinical outcomes, using either a T-cell depleted (TCD) or replete (TCRep) graft and effective GvHD prevention [40-44].

6.3.2 T cell depletion

The first successful h-HSCT with TCD was used to cure children with severe combined immune-deficiency (SCID) and it was shown to induce favorable clinical outcomes reducing GvHD without limiting immunologic reconstitution [45]. This h-HSCT platform highlighted the effect of TCD in preventing GvHD also in absence of any post-transplant prophylaxis. In contrast, when this clinical approach was transferred to patients affected by acute leukemia, it exhibited high rate of graft rejection due to an unbalance between host and donor T cells [46]. A major advance came with h-HSCT using mega-doses (on average 10x10^6 CD34^pos cells/kg body weight) of TCD, G-CSF mobilized CD34^pos PB stem cells following MA conditioning regimen based on TBI. The engraftment rate was reported to be higher than 90%, while acute and chronic GvHD (aGvHD and cGvHD) incidence was lower than 10%.

Despite this strategy ensured high engraftment rate associated with GvL effect and reduced GvHD, the MA regimens and the small number of T lymphocytes transplanted resulted in TRM
and delayed IR with subsequent occurrence of opportunistic infections [40, 47]. It was in the context of TCD h-HSCT that the alloreactivity of other immune cells, as NK cells, emerged [48, 49].

To overcome the pitfalls associated with TCD h-HSCT, several novel methods have been developed and display encouraging results with good engraftment, effective GvHD control, and favorable outcomes [50].

6.3.3 T cell repletion

Recently, the h-HSCT field saw the development of unmanipulated grafts, contrary to \textit{ex vivo} TCD, combined with strategies to limit residual donor T-cell alloreactivity. The TCRep h-HSCT presents two principal challenges: at one end donor T cell content enables an enhanced GvL effect; on the other the same T lymphocytes are responsible for GvHD onset. Nowadays GvHD prophylaxis includes G-CSF–primed grafts [51, 52], \textit{in vivo} selective TCD, and post-transplant high-dose cyclophosphamide (PT-Cy) in combination with other immunosuppressive agents. These novel approaches provide some degree of \textit{in vivo} depletion both impacting both donor-derived and host T cells. Contrary, pre-conditioning with TBI affects only recipient immune system maintaining donor cells intact, hence allowing them to exert the desired GvL responses (reviewed in [39, 53]).

6.3.4 G-CSF–primed grafts

The use of G-CSF \textit{per se} enables an increased interleukin (IL)-10 production along with a T-helper (Th)-2 activation resulting in an overall reduction of inflammatory status [51, 54]. Moreover, it has been demonstrated that the G-CSF priming enriches the BM with high number of CD34\textsuperscript{pos} progenitor cells at the expense of T lymphocytes, thus reducing the side effects resulting from their alloreactivity in the recipient [55].

Clinical evidence showed that G-CSF–primed graft followed by robust GvHD prophylaxis after h-HSCT reduces the incidences of TRM and improves patient survival [56]. In transplantation platforms where the ATG and post-transplant CSP, MMF, and short-term methotrexate (MTX) were administrated as GvHD prophylaxis, the 99\% of patients achieved sustained, full donor chimerism. The cumulative incidence of aGvHD was 45,8\% (grades II-IV), 13,4\% (grades III-IV), and cGvHD was 53,9\% indicating the overcoming of HLA barrier
without in vitro TCD for h-HSCT [56]. Another study proposed a similar protocol in 15 patients with high-risk leukemia undergoing h-HSCT [57]. Their results, in line with other groups [58-60], showed that G-CSF facilitates and accelerate the engraftment, while reducing the incidence of GvHD. Indeed, aGvHD incidence was only 6.3% in G-CSF-primed donors compared to 33% in unstimulated BM transplants [57]. More recently, a long (5 years) follow-up of a cohort of patients with different hematologic malignancies confirmed the previous encouraging results with G-CSF-primed donor [52].

### 6.3.5 Selective T-cell depletion

In the h-HSCT setting, the research field moved from an in vitro TCD to a TCRep with increased post-transplant GvHD prophylaxis. Novel graft manipulation by αβ T cell receptor (TCR) depletion has been explored [61-63]. This strategy preserves NK cells, γδT lymphocytes, Innate Lymphoid Cells (ILC), and other innate immune actors, herein enabling faster IR and maturation with consequent related immune protection [64]. Indeed, NK and γδT cells have been shown to recover quickly in selective αβ TCD compared to CD34⁺ selection [65, 66].

### 6.3.6 Post-transplant high dose Cyclophosphamide

Another TCRep h-HSCT platform consists in the use of PT-Cy pioneered by Baltimore group. It became a promising life-saving alternative strategy thanks to its safety and efficacy in lowering GvHD [67].

The administration of high doses (50 mg/kg/day) of Cy at days +3 and +4 after h-HSCT induces a state of immune tolerance between donor and recipient. Indeed, PT-Cy inhibits alloreactive donor T cell proliferation [68], sparing HSCs expressing high level of aldehyde dehydrogenase that confers Cy resistance [69]. Moreover, also regulatory T cells (Treg) cells highly express this enzyme conferring further resistance to Cy-mediated killing soon after transplantation, thus contributing to PT-Cy effect and reducing GvHD [70]. Indeed, low levels of Treg in patients after allogeneic transplantation has been associated with increased incidence of aGvHD [71, 72].

The selective alloreactive T cell elimination together with the concomitant maintenance of Treg cells ensure a peripheral tolerance between donor and recipient cells. Another mechanism of tolerance is due to intrathymic deletion of donor-derived T reactive clones during negative selection [73, 74].
However, this Cy-induced immune-tolerance is achieved only between 48 and 72 hours post transplantation; hence only precise time administration promotes GvHD protection and inhibits graft rejection [75]. For this reason the original Baltimore protocol provides the administration of immunosuppressive drugs, as tacrolimus and MMF, starting from day +5 soon after the PT-Cy, and suspended at days +180 and +35, respectively [50, 76].

Studies using PT-Cy after NMA conditioning showed rapid engraftment with donor chimerism and low risk of aGvHD (34% for grades II-IV; 6% for grades III-IV) [45]. The rates of GvHD development were comparable to that of patients treated with HLA-matched HSCT without PT-Cy [77]. Another great advantage in the use of PT-Cy after NMA conditioning is represented by the reduce toxicity of the treatment that allows also elderly people in their 6th or 7th decade of life to be treated, while previously not admitted to the HSCT procedures [50].

6.4 Factors determining h-HSCT patient clinical outcomes

6.4.1 Immune Reconstitution

Fist attempts of h-HSCT were associated to graft rejection and severe GvHD, particularly due to degrees of HLA-mismatch between donor and recipient. A delayed IR after transplantation, mostly in TCD h-HSCT, is responsible for the worsening of these poor clinical outcomes by favoring the concomitant occurrence of opportunistic infections. Contemporary, such HLA-mismatch guarantees an increased GvL effect, easily recognizing and eliminating tumor residual cells [37, 38]. Since the comprehension of the kinetics of IR following h-HSCT is key in govern clinical outcomes, several efforts have been made to rapidly achieve a fully competent immune system. As a matter of fact, many studies demonstrate that an efficient and fast immune recovery promotes h-HSCT clinical benefits [40, 64, 78-83] by improving OS, reducing relapse, and decreasing TRM [81, 84].

While innate immunity, which includes neutrophils, monocytes, macrophages, dendritic cells (DCs), and innate lymphocytes as ILCs, γδT lymphocytes, and NK cells, reconstitutes quickly, adaptive B- and T-cell lymphopoiesis takes longer, resulting in a global immune deficiency [85-87].

The first immune population that reconstitutes are neutrophils approximately 15 days after unmanipulated h-HSCT, depending also on the conditioning regimen administered [88]. Although their rapid recovery, neutrophil function is often impaired early after HSCT [89].
Moreover, a good neutrophil count and function early after transplant has been associated to a reduced GvHD, infection and cytokine release syndrome onset [90].

Contrariwise dendritic cells (DCs) IR takes longer after transplant. DC count at 1-3 months after h-HSCT appears highly low, while it reaches normal values one year later [91]. Moreover, it has been suggested that circulating DCs are mainly of donor origin, while tissue DCs remain of host origin after HSCT [85]. Albeit delayed DC IR might be involved in occurrence of infections early after transplantation, their functions in h-HSCT have not been fully clarified [91].

Three main innate lymphoid populations have been revised in the recent years as major players in defining early h-HSCT outcomes [92].

ILCs are a heterogeneous population that belongs to the lymphoid lineage. ILCs have been grouped in three distinct subsets, ILC1, ILC2, and ILC3, according to their transcriptional regulation and cytokine production, mirroring Th cell functions [93-96].

Concerning transplantation, it has been demonstrated that immune-reconstituting ILCs are donor-derived and take more than 6 months to be fully functional [97]. Recent evidence suggests a particular involvement of ILC3 cells in preventing GvHD reaction both in mice [98] and in humans, where high circulating levels of this subset in allo-HSCT patients positively correlate with a low incidence of intestinal GvHD [99, 100]. It is also proposed that ILC1 could mediate opportunistic infections soon after transplantation through their prompt interferon (IFN)-γ production in the tissue in the absence of mature and functional T and B cells [101]. However, ILC role in h-HSCT is still unclarified.

γδT cells are peculiar T lymphocytes that fill the gap between innate and adaptive immunity. Inversely to αβT cells, γδT lymphocytes reconstitute very quickly after h-HSCT [44, 61, 63, 66, 102], thus providing GvL effects and surveillance of opportunistic infections without causing GvHD in the first weeks post transplantation [103, 104].

The reconstitution of γδT cells after HSCT can be influenced by the different therapeutic approaches undertaken, including the conditioning regimen, the administration of immuno-suppressive agents, the GvHD prophylaxis, and the presence or absence of infections [104]. Notably, also viral infection/reactivation in patients undergoing HSCT has been reported to influence γδT cell IR and their phenotype [105-107].
An intense perturbation of TCR repertoire upon viral infection suggests that γδT lymphocytes are endowed with adaptive responses determined by specific clonal expansion [105]. These virus-induced features correlate with better clinical outcomes [108-110].

NK cell IR will be deeper discussed in NK cell immune reconstitution after h-HSCT paragraph.

6.4.2 Graft versus Host Disease

One of the main complications hampering allo-HSCT is still represented by donor-derived alloreactive T cells against host tissues resulting in GvHD onset. The major risk for the occurrence of GvHD is the presence of certain degrees of HLA-mismatch between donors and recipients, rendering the donor selection a hallmark in HSCT treatment, in particular in h-HSCT. Together with competent donor T cells, the different expression of tissue antigens (Ag)s between donors and recipients, and the recipient suppressed/impaired immune system are required factors that worsen GvHD [111].

Generally, GvHD is defined acute when it occurs prior to 100 days post-HSCT; while the onset of chronic GvHD is later [112]. Skin, gastrointestinal tract, and liver complications are considered as the main clinical manifestations of aGvHD [113]. The later onset of cGvHD might be de novo generated or progressive to aGvHD, therefore the implement of strategies that aims to decrease aGvHD might also prevent cGvHD onset.

The development of GvHD may be summarized in three stages [114]. The first step begins with initial host tissue damage, mainly due to HSCT pre-conditioning regimens, that in turn leads to the activation of antigen presenting cells (APCs). The consequent increased production of pro-inflammatory cytokines and of the expression of co-stimulatory molecules on APCs concur to enhance the inflammatory milieu of GvHD. The central role in GvHD reaction is played by alloreactive donor T cells that proliferate and differentiate in response to APCs of recipient. Together these events lead to a complex cascade of both cellular and inflammatory effectors that amplify tissue damages, further inflammations and subsequent GvHD. Different GvHD prophylaxis have been introduced based also on the pre-conditioning regimen patients were treated with. The combination of CSP and MTX was considered the “gold standard” procedure for MA conditioning [115], while later also tacrolimus with MTX have been extensively used
On contrary, CSP alone or in combination with MMF is the main GvHD prophylaxis therapy adopted in the setting of RIC regimens (reviewed in [117]).

Many methods of graft manipulation have been explored to overcome and/or to partner aGvHD prophylaxis as the infusion of expanded Treg cells together with the graft and the elimination of alloreactive T lymphocytes both in vivo and in vitro before the transplantation procedures (see The experience of graft manipulation in h-HSCT paragraph).

6.4.3 Infections

Opportunistic infections in h-HSCT patients mainly occur as a consequence of a delayed IR resulting in a maintained immune deficiency after transplantation. Human cytomegalovirus (HCMV) infection/reactivation represents one of the most dangerous infection threatening lives of h-HSCT patients.

In immune competent subjects, primary HCMV infection is generally asymptomatic or shows mild flu-like symptoms, while its infection/reactivation remains one of the main hampering side-effects in h-HSCT patients [118-122].

Several studies reported that more than 50% of h-HSCT patients experience HCMV reactivation in the first 100 days after transplantation. This, it seems to be more related to TCD than TCRep h-HSCT [42, 123-126]. Although novel preemptive therapies have been developed in the last years [127], infection complications after transplantation still occur [128].

A careful selection of donor is recommended, mostly in h-HSCT, since harmful effects due to the HCMV-serostatus mismatch between donor and recipient have been observed [128]. In particular, a primary infection in donors HCMV-seronegative/recipients HCMV-seronegative is very rare [126, 129]. The HCMV-seropositive status of the donor results in improved OS in HCMV-seropositive patients receiving MA conditioning, underlying a further involvement of pre-conditioning regimen [120]. In this regard, a protective effect of HCMV-seropositive donor is limited to HCMV-seropositive recipient potentially due to the transfer of donor immunity has been observed [130]. Contrariwise, HCMV-seronegative donors combined with HCMV-seropositive recipients are associate with the highest risk of HCMV infection/reactivation and related disease complications as pneumonitis, fever, colitis, and BM failure. Indeed, it has been suggested that the reactivating HCMV virus strains are generally of recipient origin, while their control is mediated by donor-derived immune cells [131]. However, other studies failed in
demonstrating any significant effect of donor HCMV-serostatus, thereby the impact of donor HCMV-serostatus remains controversial and should be deepen [132, 133].

Lack or delayed recovery of lymphocyte immune functions leads to a prolonged HCMV viremia in the recipients concurring to worsen clinical outcomes [134]. HCMV infection/reactivation has, in turn, a strong impact on IR pathways of both adaptive [135, 136] and innate immune system [107, 137]. It is conceivable that innate immune cells, recovering prior to the adaptive ones, concur in the immune-surveillance against HCMV in the first months after h-HSCT [107, 137].

6.5 HCMV

HCMV, also known as Human Herpes Virus 5 (HHV-5), is a double-stranded (ds) DNA virus belonging to β-Herpesviridae family. HCMV, as the other herpesviruses, is able to establish latent infection in healthy subjects, while in certain clinical settings, as h-HSCT, it can undergo reactivation aggravating patient outcomes [118, 119, 122].

HCMV infection induces both innate and adaptive anti-viral immune responses. In immune competent subjects HCMV infection is found in 60%-100% of individuals and it is usually asymptomatic. Indeed, in only 10% of adults, it can cause a mononucleosis syndrome characterized by fever, malaise, atypical lymphocytosis, pharyngitis and, rarely, cervical adenopathy or hepatitis [118, 138]. The viral transmission may occur in several different ways, all requiring close contact with virus-bearing secretions, such as saliva and urine, which can contain large quantities of virus for prolonged periods of time. In perinatally, intrapartum, and antenatally newborns congenital HCMV infections can also be transmitted transplacentally from the mother to the fetus during pregnancy. In these cases viral infection can cause birth defects and disabilities in approximately the 10% of congenitally infected children [139]. Besides congenital viral infections, HCMV is responsible for life-threatening disease in strongly immunocompromised hosts as h-HSCT patients, organ transplant recipients, and individuals with human immunodeficiency virus infection. In these patients lungs, gastrointestinal tract, liver, retina, and central nervous system functionality can be negatively affected [118, 138]. During transplantation, in almost the 78% of cases, the virus can be transmitted from HCMV-seropositive donors frequently to HCMV-seronegative recipients. However, HCMV-seropositive
recipients can either reactivate latent HCMV or being re-infected with new strains of HCMV when receiving organs from HCMV-seropositive donors [118, 132, 138].

6.5.1 HCMV infection and latency

HCMV shows a broad cell tropism in the host allowing the infection of virtually any organ in humans [118].

The HCMV entry process requires multiple glycoproteins (g) that bind to different cellular receptors [118, 119, 122]. In particular, gB allows the initial attachment to the cell and the further virus fusion; while gH–gL dimer is essential for the entrance into all cell types [138]. Once entered in the host, herpesvirus tegument and capsid are transported to the nucleus of infected host cells in order to inject viral DNA [140]. To enable viral DNA replication the host cellular protein synthesis is suppressed and other cellular functions are affected [140]. Herpesvirus DNA is then packaged again into new capsids and several mature virions are formed and released from the infected cell.

Following this acute replication phase, a tissue-localized persistent infection can occur, which can end with a multi-site lifelong latency characterized by the absence of detectable production of new virions [119, 141]. However, HCMV retains the capacity to reactivate, driving new antigenic stimulation, and leading to secondary immune responses [141, 142]. Unfortunately, the precise mechanisms of latency and the persistency of the virus remain elusive. Disclosing these processes would provide clues to prevent viral reactivation, risky phenomenon in immune-compromised subjects as in h-HSCT setup.

An important site of latency is represented by CD34\textsuperscript{pos} myeloid progenitor in the BM and their derived CD14\textsuperscript{pos} monocytes in the bloodstream [119, 143, 144]. Of note, HCMV infects only circulating myeloid cells, while it is absent in polymorphonuclear leukocytes, T cell or B cell fractions, although all of them derive from the same CD34\textsuperscript{pos} progenitor cells [119].

6.5.2 Interplay between HCMV and the immune system

The HCMV is considered a non-heritable factor that contributes more to immune system variation. Indeed, HCMV can shape the immune system, acting on both innate and adaptive compartment: innate immune cells are strongly activated during the acute systemic phase of HCMV infection, while during the persistent phase, HCMV acts on both innate and adaptive
immunity. The third phase of viral infection, that corresponds to latency with restricted viral gene expression, continues to shape and inflate host defenses [141]. Animal models, in particular murine ones, have been extensively used to study CMV infection phases and immune responses [141]. The key regulators of the first phase of acute infection are inflammatory cytokines, mainly type I Interferon (IFN-I) produced by NK cells, tissue-resident ILC1 [141, 145]; and splenic stromal cells within few hours post-infection [118]. The result of these phenomena is the restriction of viral replication to the site of infection and the initiation of adaptive immune responses driven by infiltrating T and B cells [118, 138]. The persistent infection induces CD8^{pos} and CD4^{pos} T cell CMV-specific response [141, 143], as well as γδT cells expansion that might play important role in the control of CMV infection [92, 146]. Despite robust cell activation, CMV has acquired different modulatory strategies to evade from both innate and adaptive immune cell recognition, thus establishing viral latency and a widespread dissemination in the host [141, 147] (see paragraph of HCMV immune evasion).

6.6 Natural Killer cells

NK cells represent approximately 5-15% of circulating lymphocytes in PB and they belong to the innate arm of the immune system [148-150]. Indeed, they naturally recognize “non-self” target cells without a prior antigen sensitization, through the engagement of activating and inhibitory NK cell receptors (aNKRs and iNKRs) [151]. NK cell activation and the nature of their functions are tightly regulated by the balance between aNKRs and iNKRs together with environmental cues and “licensing” processes occurring during NK cell development [152] (better discussed in NK cell education paragraph). Once activated, NK cells exert their roles in host defense via cytotoxic effector-functions as well as secreting cytokines and interacting with both innate and adaptive immune [153-155].

Recently, NK cells have been included in the cytokine producer ILC family characterized by the lack of rearranged Ag-specific receptors as well [96]. The lack of univocal marker identification renders NK cells study more challenging [150]. Mature NK cells are therefore defined as Lineage (Lin)^neg (i.e. markers specifically expressed by T cells, B cells, and monocytes), CD45^{pos} lymphocytes expressing the neural cell adhesion molecule-1 (CD56) and/or the low-affinity Fcγ receptor III (CD16) [156].
NK cells are the first innate lymphocyte subpopulation to recover after h-HSCT, thus representing primary controllers of cancer growth and acting as immune-supervisors of invading pathogens, especially in immune-compromised patients [49]. NK cell ontogenesis, IR after h-HSCT, and functions during physiopathologic conditions will be deepen in the following paragraphs.

6.6.1 NK cell ontogenesis

NK cells develop from CD34$^{\text{pos}}$ HSCs and they are primarily found in the bloodstream. However, the exact sites of their differentiation steps are still a matter of debate, since NK cell precursors have been found in different anatomic districts of the human body (i.e. thymus, spleen, liver, and lungs) [150, 157-159]. Several evidences demonstrated that the first phases of NK development occur in the BM niche where the interaction with stromal cells and cytokines support NK cell growth. Contrary, later stages of NK cell maturation can be observed also in secondary lymphoid tissues (SLTs), PB, and liver [160-162]. This latter represents one of the main tissue of NK cell development in the fetus and this ability is, in part, retained even after birth [163].

HSCs commitment toward lymphoid/myeloid lineage is the first step driving NK cell differentiation followed by the expansion of Common Lymphoid Progenitors (CLPs) [164]. It has been proposed that the development of NK cells is a multistep process and the different NK cell developmental intermediates can be identified based on the differential marker expression [148]. Indeed, further three stages of NK cell progenitors (NKPs), namely “pro-NK cells” (stage 1); “pre-NK cells” (stage 2); and “immature (i) NK cells” (stage 3), precede the two terminal stages of NK differentiation characterized by discrete levels of CD56 and CD16 [148, 165].

NK cells differentiation is finely tuned by different cytokine signals [165, 166]. Fms-related tyrosine kinase 3 ligand (FLT3L) and stem cell factor (SCF) preserve HSC survival and proliferation. In fact, mice lacking their receptors, FLT3 and c-kit receptor (CD117) respectively, present a decrease in CLP frequency [167]. Moreover, the presence of these two cytokines induces the expression of IL-15 receptor complex on NKPs. Finally, IL-15 is high released by BM-derived stromal and myeloid cells, strongly suggesting a direct effect of this cytokines in situ [168]. Indeed, IL-15 marks the transition from stage 2 to stage 3 promoting NK cell differentiation, functional maturation, and survival in human [169]. Moreover, IL-15 facilitates
the generation and proliferation of NK cells and its trans-presentation regulates human NK cell homeostasis both in vitro and in vivo [170].

The stage 4 accounts for conventional CD56\textsuperscript{bright}/CD16\textsuperscript{neg-low} (cCD56\textsuperscript{br}) NK cells exerting regulatory functions via the production of chemokines and cytokines and though the cross-talk with others self-cells as DCs and monocytes/macrophages [155, 171, 172]. The cCD56\textsuperscript{br} NK population contains low levels of lytic granules (i.e. perforins and granzymes) that result in a reduced cytolytic potential compared to the more mature CD56\textsuperscript{dim}/CD16\textsuperscript{pos} (cCD56\textsuperscript{dim}) NK cells (stage 5) [149, 173].

The cCD56\textsuperscript{br} NK cells outnumber cCD56\textsuperscript{dim} NK cell subpopulation in SLT, while this ratio is inverted in the bloodstream, where the latter subset of terminally differentiated cells reaches the 90% of total circulating NK cells [174]. The cCD56\textsuperscript{dim} NK cell subset is defined as the terminal stage of NK cell maturation originated from cCD56\textsuperscript{br} NK cells [148]. This developmental relationship is confirmed by the high cytotoxicity, the reduced/absence of proliferation, and the shorter telomeres that characterize cCD56\textsuperscript{dim} NK cells [175, 176]. Finally, cCD56\textsuperscript{dim} NK cells belonging to stage 5 could be further subdivided according to the later expression of CD57, marker enabling to discriminate replicative senescence of cells [177].

It has been shown that other cytokines, along with IL-15, are involved in NK cell maturation, activation, and survival, in particular in the latest stages.

IL-2 was originally associated to T cell proliferation and activation [178]. Only few years later also the increase of NK cell activity was correlated to IL-2 stimulation via the binding with IL-2 receptor (IL-2R) expressed on NK and other immune cells [179]. Indeed, IL-2 stimulation

**Figure I.1: main interleukin functions in NK cell physiology**

NK cell development and activities are finely governed by interleukins (ILs). ADCC: antibody-dependent cell cytotoxicity; NCR: natural cytotoxic receptor; IFN-\(\gamma\): interferon-\(\gamma\); TNF: tumor necrosis factor.
induced a complex cascade of signaling, including other cytokine secretion by NK cells that enhances intracellular signal transducer and activator of transcription 3 (STAT3) /AKT signaling, and upregulates various Natural Cytotoxicity Receptors (NCRs) and NKG2D receptors [180]. Low doses of IL-2 enable cCD56^{br} NK cell proliferation due to their constitutive expression of high-affinity heterotrimeric IL-2R, peculiar feature of this latter NK cell subset [181]. IL-2 was the first cytokine used for clinical trial to boost NK cells in an in vivo setting [182]. Growing evidence in the field of transplantation have reported significant clinical results when h-HSCT was supplemented with IL-2 [183, 184].

IL-18 is a member of the IL-1 family and it represents an immune stimulatory cytokine regulating both the innate and the adaptive immunity. It co-stimulates NK cell CD16-mediated cytolytic activity, IFN-γ production, and activation of Tumor necrosis factor (TNF) signaling, highlighting its role in control of infections [185-187].

IL-12 is a pro-inflammatory type I cytokine that governs the activation of immune responses against infected and malignant cells. IL-12 exerts its properties through the stimulation of NK cells and T lymphocyte driving the production of IFN-γ [188], improving the antibody-dependent cell cytotoxicity (ADCC) [189], and enhancing their cytolytic properties [190].

IL-21 is a type I cytokine able to influence both innate and adaptive immunity causing lymphoid proliferation, in particular of NK, CD8^{pos} T, and B cells [191]. Moreover, it leads to JAK1/JAK3 pathway activation resulting in an increased production of IFN-γ [192]. In combination with IL-15, IL-21 it can promote the expansion of CD56^{pos}CD16^{pos} NK cells derived from human BM [193].

The above-mentioned cytokines, due to their role in implement NK cell numbers, functions, and persistence, have been increasingly used for cancer immunotherapy [184] (Figure I.1).

### 6.6.2 Heterogeneity of NK cell subsets

Mature circulating NK cell subsets are generally defined by the relative expression of CD56 and CD16, as previously mentioned [194].

The expression of CD56 antigen increases gradually during NK cell development from stage 2 to stage 4 (cCD56^{br} NK cells) and then decrease from stage 4 to stage 5 (cCD56^{dim} NK cells), but there is no direct evidence explaining the biological meaning of different degrees of expression of this molecule on NK cells [148, 195]. Recently it has been proposed that CD56 antigen is
involved in the maturation process of NK cells, probably mediating a required developmental synapse between iNK and stromal cells [196]. On the contrary, CD16 expression is acquired during the transition from stage 4 to stage 5 [148, 195].

Besides conventional NK cell subsets, other NK cell subpopulations can be identified on the basis of the relative CD56 and CD16 surface expression: unconventional CD56dimCD16neg (unCD56dim) CD56brightCD16pos (unCD56br) and CD56negCD16pos (CD56neg) ([194, 197]. These subsets not only differ in phenotype, but also in functional capabilities, hinting they might play different roles within the innate immune system ([198]. A more detailed characterization of NK cell subsets is provided in the following paragraphs.

### 6.6.2.1 cCD56\^br NK cells

Conventional CD56\^br NK cells represents about 10% of total peripheral blood circulating NK [181], while these cells are mainly enriched in secondary lymphoid organs (SLTs), like tonsils and lymph nodes. They are defined by the high expression levels of L-selectin (CD62L) and CC-chemokine receptor 7 (CCR7), receptors required for lymphocyte homing to SLTs, whose expression suggests the ability to traffic to these compartments in vivo [181].

The progression from “iNK” cells to mature cCD56\^br NK cells is marked by the acquisition of CD94/NKG2A expression, which confers HLA-I dependent inhibition [148]. cCD56\^br NK cells are show low to absent expression of KIRs and ILT-2, while high levels of the hematopoietic stem cell marker CD117 and of the activating receptor NKp46 compared to other NK cell subsets [181, 199].

cCD56\^br NK cell subset appear to be the only lymphocyte population with constitutive expression of the high-affinity heterotrimeric IL-2R (IL-2Rαβγ) showing a high proliferative response to picomolar concentrations of IL-2 [199]. Moreover, cCD56\^br NK cells display higher-level expression of cytokine receptors, as IL1R-I and IL18R, with respect to the other NK cell subpopulations [199].

Functional studies demonstrated that cCD56\^br NK cells are poorly cytotoxic but appear to be the primary source of immunoregulatory cytokines, including IFN-γ, TNF-α, TNF-β, and GM-CSF, in response to monocyte-derived cytokine stimulation [181].
A new poorly characterized subset of NK cells, with intermediate phenotype and properties between cCD56\textsuperscript{br} and cCD56\textsuperscript{dim}, has been identified [198]. This population of CD56\textsuperscript{bright}CD16\textsuperscript{pos} (here, for brevity, unCD56\textsuperscript{br}) NK cells transiently expand after HSCT, constituting up to 61% of the cCD56\textsuperscript{br} population at 6 month after transplant, have been proposed to be an intermediate stage of differentiation between cCD56\textsuperscript{br} and cCD56\textsuperscript{dim} NK cell subsets [200]. Indeed, unCD56\textsuperscript{br} NK cells exhibit intermediate expression levels of CD25, CD117, CD8, CX3CR1, NKG2D, KIRs, and granzyme A between the conventional mature NK cell subsets, in accordance with evidence in literature predicting CD117 loss and KIRs acquisition during stage 4 to stage 5 transition [148, 195, 200].

6.6.2.2 cCD56\textsuperscript{dim} NK cells

The majority of peripheral blood NK cells is constituted by cCD56\textsuperscript{dim} NK cells [181]. Their high-expression levels of CD16 and their high content of cytolytic granules, composed by perforin and granzymes, make them efficient mediators of ADCC, hence demonstrating a greater natural cytotoxicity than cCD56br NK cells [181, 201].

The complete maturation of cCD56\textsubscript{dim} NK cells involves the progressive loss of CD94/NKG2A and the acquisition of combination of KIR molecules, in order to form a relatively stable repertoire of receptors which is mainly genetically determined [181]. During the transition from cCD56br to cCD56\textsubscript{dim} NK cells, a downregulation of NKp46 and CD117 activating receptors also occurs [199]. cCD56\textsubscript{dim} NK subset displays a stronger expression of C-X-C motif chemokine receptor 1 (CXCR1), CX3CR1 and CD11a compared to cCD56br NK cells, that is reflected on migratory ability towards acute inflammatory sites [181, 199]. cCD56\textsubscript{dim} NK cells lack a significant proliferative response even to high concentrations of IL-2 and produce a negligible amount of IFN-\(\gamma\) and other NK-cell-derived cytokines following monocyte-derived cytokine stimulation [181].

6.6.2.3 unCD56\textsuperscript{dim} NK cells

Recently a new unconventional NK cell subset, defined by CD56\textsuperscript{dim}CD16\textsuperscript{neg-low} expression (hereafter unCD56\textsuperscript{dim}), has been discovered [194, 202]. This subset is significantly more abundant in the bone marrow than in the peripheral blood (10% and 5% on total NK cells, in BM and in PB respectively) [203].
Concerning their phenotypic profile, unCD56\textsuperscript{dim} NK cells displays intermediate levels of NKG2D expression with respect to cCD56\textsuperscript{br} and cCD56\textsuperscript{dim} NK cell subsets, while similar levels of Granzyme-B, Perforin and NKp46 expression compared to cCD56\textsuperscript{dim} NK cells [203]. Unlike this latter NK cell subset, unCD56\textsuperscript{dim} NK cells exhibit low expression levels of KIRs and CD57 senescence marker, but also high-expression levels of CD94/NKG2A, CD25, CD122, CD127, and CD27, whose expression has been associated with earlier differentiation stages [202, 203]. Although unCD56\textsuperscript{dim} NK cell subset is poorly represented in healthy individuals, they have a potential to be strong effectors in innate immunity [202]. Indeed, these cells exhibit high proliferative potential and high IFN-\gamma producing ability upon cytokine stimulation [204, 205]. In addition, unCD56\textsuperscript{dim} NK cells isolated from h-HSCT donors, display the highest degranulation and cytotoxic potential against both K562 target cells and autologous leukemic blasts [204, 206]. This NK cell subpopulation displays distinct phenotypic and functional features of NK cells but remains to be clarified which is their role in the NK cell differentiation and maturation pathway.

6.6.3 NK cell receptor repertoire and functions

NK cell effector-functions are tightly governed by integration of both aNKR and iNKR signaling. Self-HLA-I molecules are expressed on almost all healthy cell surfaces and they act as ligands for iNKRs, ensuring the self-tolerance. Primary tumor-transformed and viral infected cells show a decreased expression of self HLA-I molecules that drives NKR pattern variability among NK cells [207, 208] and enable them to spare self-cells from the killing (“missing-self hypothesis”) [209]. In the meanwhile, the upregulation of ligands for aNKR on target cells drives NK cell activation which manifests with the release of cytotoxic granules and/or through the production of inflammatory cytokines [210].

The main classes of receptors specific for HLA-I molecules, hence mediating self-recognition, include Killer Ig-like Receptors (KIR), which are able to recognize different HLA-A, -B, and -C allotypes [211] and C-type lectin receptors (CD94/NKG2) binding the non-classical HLA-E molecules [212, 213].

KIRs (also known as CD158) are highly polymorphic family of receptors, including both iNKRs and aNKRs, that serve not only as regulators of NK cell activation, but are also actors in their development and tolerance [214]. KIR members share homology in the extracellular domain, while aKIRs and iKIRs differ in their cytoplasmatic tails. In detail, iKIRs are
characterized by a long intracellular domain containing Tyrosine-Based Inhibitory Motif (ITIM) that recruits SHP-1/2 tyrosine phosphatases and interrupts the downstream NK cell activation pathway [211, 215].

On contrary, a short cytoplasmatic tail interacts with adaptor signaling molecules, as DAP-12, carrying an Immunoreceptor Tyrosine-Based Activating Motif (ITAM) [215].

C-type lectin receptors are mainly represented by CD94/NKG2 heterodimers: similarly to KIRs, the inhibitory CD94/NKG2A is characterized by long intracellular tail, while CD94/NKG2C lacks the ITIM and it is correlated to DAP-12 to drive NK cell activation [216].

Other NKR s, acting as important mediators of NK cell cytotoxicity against target cells, belong to the family of NCRs and CD16. NCRs interact with a broad range of ligands of different origin and control NK cell activation. Human NCRs include NKp30, NKp46, NKp80, and NKp44. While the former are expressed on both resting and activated NK cells, NKp44 is upregulated upon activation and it is not expressed in basal conditions [217, 218].

CD16 (FcγRIIIA) is an immunoglobulin (Ig) receptor that recognizes and binds to the Fc portion of IgG antibodies. Through adaptor molecules containing the activation ITAM, as CD3ζ and FceRIγ, CD16 mediates the so-called ADCC. Hence, the preferential expression of CD16 on cCD56dim cells renders this NK cell subset highly cytotoxic with respect to cCD56br NK cells [153, 219].

6.6.4 NK cell education

NK cell education/licensing is defined as the mechanism ensuring NK cell responsiveness against non-self targets while retaining tolerance versus autologous cells [210]. This NK cell homeostatic process occurs during NK cell development through the sequential acquisition of iNKR and their binding to self-HLA-I molecules [220].

The interaction is contemporary responsible for the NK cell recognition of downregulated or lack HLA-I molecules on tumor-transformed or viral-infected cells, targets of NK cell functional activation [220-222]. Moreover, the balance between aNKR s and iNKR s strictly governs NK cell alloreactivity, which relies on the number of iNKR s expressed and the avidity of binding with HLA-I complex [223-225].

Albeit a rigorous NK cell receptor regulation is necessary to direct NK cell effector-functions, it also drives NK cell maturation [226]. In fact, the expression of iNKR s is involved in NK cell differentiation process: cCD56br NK cells usually show high levels of CD94/NKG2A,
while almost lack KIR molecules that are known to be involved in self-tolerance and licensing responses [221]. On contrary, cCD56dim NK cells progressively acquire KIR expression at the expense of CD94/NKG2A, thus representing the main subset of “educated” NK lymphocytes [227, 228]. Indeed, a fine regulation of NK cell responsiveness through KIR-KIR ligand binding during development is necessary. It has been hypothesized that when an NK cell, although expressing a certain iKIR, does not emit or emits a weak inhibitory signal, the cell expresses a new KIR and so on until it reaches the target threshold [229]. For that reason, educated cells have to express at least one iKIR; while both aNKRs and iNKRs are further involved in NK cell education [230].

6.6.5 NK cell immune reconstitution after h-HSCT

NK cells are able to mount an efficient alloreactive reaction against target cells without any antigen presentation, thus accelerating the host response to tumor-transformed or viral-infected cells, in particular after HSCT when adaptive immune cell recover is still delayed. Therefore, the analysis of NK cell IR pathway represents certainly a key in determining the clinical outcome of transplanted patients. In addition, the chance to trace NK cell IR steps in HSCT setting provides a unique model for the disclosure of NK cell maturation in vivo and the development of novel therapies able to exploit NK cell alloreactivity. In the context of HSCT, NK cells are the first donor-derived lymphocytes to recover defining their role in defending immune-compromised recipients [49]. Indeed, a delayed NK cell IR has been associated with poor prognosis in transplanted patients [231, 232].

NK cell donor chimerism is completed within a month after allo-HSCT irrespective of the graft source, prior to adaptive immune cells [49, 88, 226].

Besides their prompt expansion, NK cell full maturation, involving differentiated phenotype and functional competency, takes longer [78, 175]. Similar results have been described also in HLA-matched HSCT where the immune reconstituting NK cells remain immature for more than 6 months and their prolonged functional defects might be combined to more adverse events such as HCMV reactivation, GvHD, relapse and death at one year post transplantation [233]. It is generally accepted that immune reconstituting NK cells derive from the differentiation of donor CD34pos progenitors rather than already mature NK cells retained in
the graft. In this regard, it has been demonstrated that PT-Cy treatment few days after h-HSCT eliminates the majority of mature alloreactive NK cells infused with the TCRp graft [234]. Different works agree that the first NK cell subset to recover is the less differentiated cCD56^br NK cells, while the usual CD56/CD16 distribution in patients return similar to that of healthy subject only several months later [78, 235-238].

6.6.6 Role of NK cells in h-HSCT

In h-HSCT setting, given to certain degrees of HLA-mismatch between donor and recipients, NK cells can mediate a potent GvL effect prior to and in absence of T cell participation. NK cells are further involved in limiting GvHD onset and in preventing graft rejection [239, 240]. The emergence of these concepts has highlighted the importance of innate immunity in determining positive clinical outcomes and its possible exploitation for the development of novel therapeutic strategies [17, 49, 241].

As already mentioned, a prompt NK cell repopulation following HSCT has been associated to reduced incidence of relapse and infections together with increased OS [231, 242, 243]. Conversely, low numbers of NK cells the first weeks after allo-HSCT has been correlated to increased non relapse mortality (NRM) and shorter OS concomitant to the occurrence of viral infections [232, 242].

6.6.7 NK cells and GvHD

The GvHD represents one of the main clinical complications affecting patient lives after h-HSCT. It is certainly true that NK cells are involved in the determination of clinical parameters, but their role in GvHD onset is not fully clarified. It has been proposed that alloreactive NK cells can directly suppress GvHD by killing donor T cells responsible for the disease [244]. Indeed, the upregulation of stress molecules on T cell surfaces may be responsible for NK cell activation through the NKG2D receptor recognition [245].

Moreover, NK cells are recruited in GvHD target organs, where also activated donor-derived T cells traffic. The contemporary presence of NK cells may offer the possibility to kill alloreactive T cells directly in the GvHD sites [246]. In addition, recipient tissue-resident NK cells, persisting to pre-conditioning treatment, can exert GvHD protection by T cell elimination.
as reported in murine model [247]. T cell killing seems to depend on both perforin release [244, 247] and FAS-mediated induction of apoptosis [244] (Figure I.2).

In addition to a direct effect on T cells, NK cells may also participate to the GvHD-responsible T cell clearance through the indirect depletion of APCs that, in turn, would stimulate alloreactive T cells [248] and by producing immune-regulatory cytokines as IL-10 [249]. In agreement, several evidence demonstrates that high NK cell counts are correlated to a decrease onset of aGvHD [49, 231, 239, 242, 244]. On the contrary, the release of the pro-inflammatory IFN-γ by NK cells could promote tissue damage, thus being potentially involved in GvHD onset [248].

Notably, NK cell receptor expression and their consequent functional responses could influence the disease occurrence. Recently, it has been shown that h-HSCT patients developing GvHD display a more differentiated and activated NK cell phenotype than patients without GvHD [250]. In this regard, aNKR CD94/NKG2C is significantly reduced in patients suffering from severe GvHD after allo-HSCT, thus indicating its likely involvement in disease prevention [251]; conversely a more recent study showed that CD94/NKG2A<sup>pos</sup> NK cells regulate T cell functions <i>in vitro</i>, inhibiting T cell proliferation and activation, thus likely attenuating aGvHD also <i>in vivo</i> [252] (Figure I.2).

Looking more in deep to NK cell subpopulation, cCD56<sup>br</sup> NK cells appeared diminished in the first two months after allo-HSCT in patients suffering from aGvHD, identifying their putative
role as early prognostic factor for disease development [237, 240]. Moreover, high levels of cCD56br NK cells soon after h-HSCT have been demonstrated to drive survival advantage in patients suffering hematologic malignancies [80].

Although the NK cell role in participating in GvHD onset and development is still controversial, it is known that this process can in turn affect NK cell biology. Indeed, aGvHD leads to a decreased IFN-γ production, thus influencing NK cell activity in mouse models [253]. Moreover, GvHD could slow NK cell IR after transplantation both in mice and humans [240, 254]. The dysfunctional GvHD-induced status of NK cells might results in the reduction of GvL effect and in the impairment of viral clearance by NK cells [255]. Remarkably, also pre- and post-conditioning treatments in h-HSCT might affect NK cell features. In particular, GvHD prophylaxis with MMF has been shown to reduce NK cell proliferation and effector-functions, thus conditioning the NK cell potential control of disease and infections [256, 257].

6.6.8 NK cell shaping upon HCMV infection/reactivation

6.6.8.1 HCMV immune escape

HCMV contains a ds-DNA genome coding for hundreds of genes, including the ones used to evade both adaptive and innate immune recognition [258-260]. Indeed, more than 40 genes are engaged in modulating immune responses in the host upon infection [261, 262]. Although host immune system throughout life has developed mechanisms of viral defense, HCMV has evolved different processes to escape from immune system [141, 142, 263-266].

As other herpes viruses, HCMV can drive the downregulation of HLA-I surface expression on infected cells interfering with Ag presentation, thus avoiding cytotoxic T cell killing [267, 268]. HCMV is able to both impede peptide translocation and to stimulate degradation of HLA-I molecules before they reach the cell membrane [262]. However, this mechanism leads to the NK cell activation through both the missing iNKR-HLA binding and the concomitant engagement of aNKRs, rendering HCMV more susceptible to NK cell-mediated lysis. Contemporary, HCMV virus has developed a sophisticated array of strategies to evade the innate immunity too [264, 269, 270]. There are about 12 gene products known to be involved in NK cell modulation [262]. The major mechanism inducing HCMV-mediated inhibitory signaling that has been proposed sees the over-expression of HLA-E on the surface of infected cells, facilitating the interaction with the CD94/NKG2A iNKR and the escape
from NK cell killing [271-273]. Moreover, since NK cell activation is tuned by a fine balance between aNKRs and iNKRs, it is also important for the virus to prevent the signaling by activating receptors [274]. For a further description of NK cell receptor repertoire, refer to NK cell receptor repertoire and functions paragraph.

For instance, protein and RNA-based processes occurring during HCMV infection act to prevent NKG2D ligand expression [264, 275]; while US12-family genes, in particular US18 and US20, have been shown to determine the downregulation of B7-H6 membrane protein levels, thus preventing NK cell activation through NKp30 receptor. Moreover, previous work showed that HCMV protein pp65 (UL83) is able to bind NKp30 inhibiting NK cell-mediated killing, suggesting further involvement of HCMV in manipulating NK cell activity [260, 276].

6.6.8.2 CD56negCD16pos NK cell subset

During the co-evolution between the host and viruses, HCMV has been significant in modeling host immune system [141]. Growing evidence showed that HCMV infection/reactivation influences NK cell IR and development after h-HSCT by shaping NK cell subset distribution and promoting the expansion of an unusual CD56negCD16pos (CD56neg) NK cell subpopulation, commonly poorly represented under homeostatic conditions [277, 278].

CD56neg NK cell expansion has been documented also upon the occurrence of other infections, as HIV-1 and HCV [197, 279], where it has been described as a dysfunctional subset [280, 281] with aberrant reduction of aNKR expression [277], contributing to a state of defective control of infections.

Although the origin of CD56neg is still debated, it is conceivable that its expansion can be associated with chronic and systemic inflammation usually occurring along with viral infections [278, 282-287].

In regard of their ontogeny, it has been proposed that CD56neg NK cell subset could arise from a failure in NK cell development, hypothesis supported by the experimental evidence showing that interleukin stimulation can restore conventional NK cell subset distribution and classical NKR repertoire [284, 288]. However, CD56neg NK cell population shares phenotypic features with mature NK cell subsets as the expression of CD16 and KIRs, thus suggesting that this particular NK cell subset does not derive from malfunctional iNK cell (stage 3) development. Moreover, CD56neg NK cells can express CD57, marker of terminal differentiation.
[289, 290] and exert killing capacity albeit reduced compared to cCD56^{dim} NK cells [281, 282], thus representing a more terminal differentiation stage of NK cell development [197, 280].

In agreement, an alternative hypothesis defines CD56^{neg} NK cells as already mature, but hypofunctional lymphocytes that have recently engaged target cells, suggested from the deregulation of perforin content with respect to cCD56^{dim} NK cell subset [279]. However, the origin, the function, and the impact of CD56^{neg} NK cells in determining the clinical outcomes of transplantation are not yet been clarified.

### 6.6.8.3 Memory-like NK cell subset

The term “immunological memory” stands for the property of the immune system to “store” information about a stimulus, to “remember” it, and to mount a stronger and faster effective response upon a secondary encounter with the same stimulus/pathogen, also many years after the first exposure [291].

Seminal evidence show that not only adaptive T and B lymphocytes can exert secondary responses, but surprisingly also innate immunity does.

First evidence in mice showed that a secondary exposure to a certain chemical hapten would induce a severe hapten-specific hypersensitivity on the contact site. This Ag-specific reaction is induced also in recombination activating gene (RAG)-deficient animals lacking B and T cells, thus first highlighting the role of NK cells in inducing immunological memory [152, 292]. These surprising findings were successively reinforced by the discovery of a peculiar NK cell subpopulation that can be detected upon a
second exposure to murine cytomegalovirus (MCMV) showing certain similarities with memory responses [293]. This NK cell subset is able to robustly expand during MCMV infection upon the specific binding of Ly49H receptor with m157 viral ligand and to exert primary response to viral infection. Moreover, adoptive transfer experiments demonstrated not only the capability of Ly49H<sup>pos</sup> NK cells to proliferate and expand, but also to follow the other conventional memory phases of immunity: “contraction”, “memory maintenance” and enhanced “secondary/recall” responses [293] (Figure I.3 memory).

Briefly, after the first encounter with a pathogen Ag the “contraction” phase takes place. It is characterized by the sever reduction in all tissues of activated cells undergoing apoptosis [294, 295]. Subsequently, during the “memory maintenance” phase a stable population of long-lived cells resides in the tissues and patrols against pathogens [293, 296]. Finally, the re-encountering of the same Ag drives a “recall” wave of effector-functions which is stronger and more efficient than the primary response [297-299].

MCMV is usually studied as the example of NK cell mediated surveillance of viral infection and consequent NK cell adaptive features [300, 301]. However, burgeoning evidence has suggested that those phenomena driving the appearance of memory-like (ml)-NK cell may occur also in humans and non-human primates [152].

Opposite to infections in mice [302], a univocal ml-NK cell phenotype in humans has not been defined yet and a hypothetical viral ligand remains uncertain. Notably, in HCMV-seropositive individuals, an expanded population of long-lived NK cells expressing NKG2C and CD57 has been observed in response to HCMV-infected fibroblasts [303, 304] and it seems to require monocytes producing IL-12 after HCMV infection in order to proliferate and to expand [305]. This unusual NK cell subset expansion has been later clarified also in an in vivo setting after HCMV infection [290]. Although the NKG2C<sup>pos</sup> NK cell subset undergoes clonal-like expansion resembling adaptive responses to HCMV infection, there is no evidence of a direct NKG2C-viral protein recognition [270, 304, 306-308]. A recent work demonstrated that NKG2C<sup>pos</sup> NK cells differentially recognize HCMV UL40-encoded peptides contributing to the expansion of ml-NK cells based on the strength of peptide recognition [309]. Nevertheless, it is conceivable that variations in HCMV-encoded proteins could influence the expansion of adaptive NK cells showing different phenotype. To further confirm this concept, novel insights demonstrated that adaptive NK cells undergo an epigenetic reprogramming through a specific
reconfiguration of adaptor molecules including tyrosine kinase SYK, the intracellular adaptor EAT-2, as well as the transmembrane adaptor protein FceRIγ (also known as FceRγ) [310-312]. Indeed, the reduced expression of at least one of the mentioned signaling proteins is observed in the 50% of the HCMV-seropositive donors. Of note, the transcription factor promyelocytic leukemia zinc finger (PLZF) is downregulated in the majority of ml-NK cells. It can bind to the promoters of the genes encoding for FceRγ, SYK, and EAT-2, likely explaining the shaping in these NK cell gene expression upon HCMV infection/reactivation [310, 311]. The reduced levels of PLZF also causes a low expression IL-12 and IL-18 receptors [313] resulting in the lack of responsiveness to inflammatory cytokines [208].

Importantly, the downregulation of FceRγ, SYK, and EAT-2 in mature cCD56dim NK cells correlated with the expression of NKG2C, thus indicating that, albeit heterogeneous, the defined NK cell population resembles to ml-NK cells with respect to surface marker expression [310]. Moreover, cCD56dim NK cells exerts their cytotoxicity primary through CD16 receptor coupled to CD3ζ and FceRγ adapters to a PTK activation pathway [314]. However, CD16 remain functionally coupled to CD3ζ even after the downregulation of FceRγ adaptor molecule retaining NK cell activity [311] (Figure I.4).

To further corroborate the presence of innate cells able to manifest adaptive features, ml-NK cells revealed genome-wide DNA methylation patterns shared with those of cytotoxic CD8pos T cells, strongly suggesting epigenetic influence in HCMV-driven NK cells [310]. Specially, the increased IFN-γ release by ml-NK cells has been correlated with a stable demethylation of conserved non-coding sequence 1 of the IFNG locus, similar to memory Th-1 cells [315].
6.6.8.4 Memory-like NK cells in h-HSCT patients

The prolonged immunodeficiency following transplantation causes increased risk in developing viral infections/reactivation, mainly HCMV [124]. Contrary to solid organ transplantation, HCMV infections following HSCT are more frequently caused by the reactivation of latent virus present in the HCMV-seropositive recipients than primary infection. Indeed, only the 30% of HCMV-seronegative recipients develop primary HCMV infections, while the 80% of HCMV-seropositive patients experience reactivation of the latent virus [118, 138]. Due to the high percentage of HCMV infection/reactivation after h-HSCT [42, 123-126], memory-like responses against the virus have to be considered.

In fact, recent studies in h-HSCT setting revealed that HCMV infection/reactivation could be beneficial rather than detrimental for patients. Indeed, HCMV accelerates NK cell maturation and influences NK cell receptor repertoire. Despite the mechanisms behind this are not completely understood, the rapid NK cell development may be exploited as an advantage in h-HSCT setting where terminally differentiated cCD56\textsuperscript{dim} NK cells exert an important anti-leukemia role [316].

In particular, NK cells in HCMV reactivated patients are characterized by a cCD56\textsuperscript{dim}NKG2C\textsuperscript{pos} CD57\textsuperscript{pos}NKG2A\textsuperscript{neg}KIR\textsuperscript{pos} signature [317, 318], opposite to HCMV not reactivated recipients showing a more immature phenotype expressing high levels of the CD94/NKG2A inhibitory receptor [319]. Despite the expansion of NKG2C\textsuperscript{pos} NK cells seems to be representative of viral infection/reactivation, also patients receiving cord blood from NKG2C-deficient donors (carrying homozygous deletion of the NKG2C gene) show faster NK cell maturation upon HCMV reactivation similar to recipients transplanted from NKG2C\textsuperscript{pos} donors. NK cell compartment in these patients has been characterized by CD56\textsuperscript{dim}CD94\textsuperscript{pos}NKG2A\textsuperscript{neg}NKG2C\textsuperscript{neg}aKIR\textsuperscript{pos} phenotype triggering NK cell cytotoxicity [320, 321].

It is conceivable that NK cell subsets expressing CD94/NKG2C, aKIR or other markers still unknown may play complementary roles in the response to HCMV, thus favoring viral protection.

Importantly, the frequency of these mature NK cells persists after one year from h-HSCT in patients experiencing HCMV reactivation and they display properties of the adaptive immunity
Indeed, ml-NK cells show strongest effector-functions, in terms of IFN-γ and TNF-α production [317], when re-encountering the same antigen [319, 322] or following a proper activation with specific pro-inflammatory cytokines [323] (Figure I.3).

Although a significant heterogeneity characterizes ml-NK cell phenotype among infected individuals, their rapid maturation following HCMV infection/reactivation could favor not only the control of infection, but also NK cell alloreactivity against residual tumor cells [108], thus rendering ml-NK cells an attractive tool for immunotherapy [92, 109, 324, 325].

Some evidence in humans suggest that ml-NK cells could be transplantable from donor to recipient: an expansion of NKG2C$^{\text{pos}}$ NK cells was observed not only in recipients experiencing HCMV reactivation, but also when both donor and recipient were HCMV-seropositive even in the absence of detectable HCMV viremia [317, 326]. Interestingly, when NKG2C$^{\text{pos}}$ NK cells are transplanted from HCMV-seropositive donor, they exert robust and stronger functions (in term of IFN-γ production) upon a secondary HCMV event in respect to NKG2C$^{\text{pos}}$ NK cells derived from an HCMV-seronegative donor. The results further suggest that also HCMV antigens in the recipient might contribute to NKG2C$^{\text{pos}}$ NK cell expansion [326]. Similarly, high levels of NKG2C$^{\text{pos}}$ NK cells in patients before kidney transplantation were associated to reduce post-transplant HCMV viremia, indicating that the pre-existence of adaptive NK cells in the recipient may confer certain degrees of protection against viral reactivation [327].

Recently, it has been reported that patients with acute leukemia (both myeloid and lymphoblastic) undergoing h-HSCT and experiencing HCMV reactivation display the expansion of NK cells producing high level of IFN-γ following in vitro co-culture with K562 erytroleukemia cell line [328].

Furthermore, consistent with murine findings, the adoptive transfer of donor-derived NK cells (pre-activated with IL-12, IL-15, IL-18 before the infusion) in patients with refractory AML resulted in a NK cell expansion in the recipient [323, 329] displaying stronger IFN-γ production than recipient-derived NK cells when ex vivo stimulated with K562 target cells.
7 AIM AND RATIONALE

h-HSCT represents a promising approach to cure patients affected by high-risk hematologic malignancies in the absence of an HLA-matched related donor. The introduction of new h-HSCT platforms, as the use of PT-Cy and RIC regimens, has significantly improved the clinical outcome of patients. However, the immune-mediated GvL effect has not been fully exploited yet and it is still hampered by several life-threatening side effects such as GvHD and opportunistic infections. In this regard, the quality of the IR is emerging as a key determinant for h-HSCT patient clinical outcome.

Our group recently demonstrated that NK cells are the first lymphoid population recovering after h-HSCT, thus playing a crucial role in early immunity after the transplant [238]. Therefore, in this project of translational immunology, we aim to dissect the NK cell subset IR and we intend to investigate their alloreactive responses soon after h-HSCT.

Firstly, we focused on the characterization of CD56\textsuperscript{dim}CD16\textsuperscript{neg-low} (unCD56\textsuperscript{dim}) NK cells poorly represented in healthy individuals, while highly expanded early after h-HSCT starting from the 2\textsuperscript{nd} week from the transplant. We hypothesized that proliferating unCD56\textsuperscript{dim} subset in patients could govern NK cell development, figuring as an intermediated additional stage of maturation. Moreover, giving their transient expression of inhibitory CD94/NKG2A receptor differently from physiologic conditions, we aim to disclose unCD56\textsuperscript{dim} NK cell effector-functions and their putative role in performing GvL effect.

Emerging evidence suggests that opportunistic infections, in particular those driven by HCMV reactivation, can affect NK cell IR after h-HSCT, promoting the expansion of a dysfunctional CD56\textsuperscript{neg} NK cell subset [319, 330] and driving the development of hyper-responsive mature NK cells endowed with adaptive immune traits, named "memory-like" NK cells [322].

Since their origin, their univocal phenotype and their impact in determining the clinical outcome of h-HSCT have not been clarified yet, we sought to deeply define the influence of HCMV infection/reactivation in modulating NK cell phenotype and effector-functions against viral infected targets. We hypothesize that HCMV-driven ml-NK cell anti-viral potential can be exploited also against residual cancer cells improving GvL and enhancing tumor protection. Our
final purpose is to assess the ml-NK cell fingerprint and understand whether these cells are *de novo* generated or they can be transferred from donor to recipient, thus allowing the selection of the best donor before h-HSCT.
8 MATERIALS AND METHODS

8.1 Patients’ recruitment

In this study we enrolled 50 patients affected by hematologic malignancies and treated according to our published h-HSCT protocol approved by the Institutional Review Boards of Clinical and Research Institute Humanitas [125]. Briefly, recipients received RIC starting from the 6th day before HSC infusion to suppress their immune system and to allow an adequate engraftment [42]. Cy was administered by intravenous infusion at days +3 and +4 to selectively deplete alloreactive T cells, allogeneic Ag-stimulated cells that clonally expand after transplantation. Starting from the day +5, pharmacologic prophylaxis with tacrolimus and MMF was administered to prevent GvHD and it was then suspended at day +180 and at day +35, respectively [42, 50, 76] (Figure M.1).

Patients were recruited at the Department of Hematology and Bone Marrow Transplant Unit, Humanitas Cancer Center, Humanitas Research Hospital, Rozzano, Milan, Italy (Dr. L. Castagna, Dr. S. Bramanti, and Dr. J. Mariotti).

All patients and donors signed an informed consent form in accordance with the Declaration of Helsinki specifying that the donation of specimens for this project would not have influenced diagnosis, therapy and prognosis of the disease in any way.

We collected blood samples from both donors and recipients before the transplantation and from the recipients every week for the first month post-h-HSCT and then every month till

![Figure M.1: Baltimore protocol used to treat patients affected by hematologic malignancies](image)

Cy: Cyclophosphamide; GCSF: Granulocyte Colony Stimulating Factor (week 3); HSC: Hematopoietic stem cells; MMF: Mycophenolate Mofetil (week 5); NMA: non myeloablative.
one year after the transplant, ranging from day +14 to day +386, adding up to a median of 6 samples per patient.

Table 1: patient (pt) clinical data and sample collection

<table>
<thead>
<tr>
<th>N</th>
<th>k-ESCT</th>
<th>Sex: D/R</th>
<th>R: age</th>
<th>D: age</th>
<th>Relation</th>
<th>pathtology</th>
<th>HCMV infection/reactivation</th>
<th>HCMV therapy (days)</th>
<th>Multiple/Single</th>
<th>Reactivation samples (time points) used</th>
<th>experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>#3</td>
<td>FF</td>
<td>20</td>
<td>51</td>
<td>father</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>#4</td>
<td>MF</td>
<td>22</td>
<td>52</td>
<td>father</td>
<td>HL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>#5</td>
<td>FM</td>
<td>44</td>
<td>47</td>
<td>mother</td>
<td>NHL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>#6</td>
<td>MM</td>
<td>45</td>
<td>47</td>
<td>brother</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>#8</td>
<td>FF</td>
<td>33</td>
<td>61</td>
<td>mother</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>#9</td>
<td>MM</td>
<td>45</td>
<td>47</td>
<td>mother</td>
<td>NHL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>#10</td>
<td>MM</td>
<td>57</td>
<td>62</td>
<td>brother</td>
<td>NHL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>#11</td>
<td>MF</td>
<td>34</td>
<td>29</td>
<td>brother</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>#12</td>
<td>MF</td>
<td>41</td>
<td>23</td>
<td>son</td>
<td>NHL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>#13</td>
<td>MM</td>
<td>54</td>
<td>46</td>
<td>mother</td>
<td>NHL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>#14</td>
<td>MF</td>
<td>57</td>
<td>24</td>
<td>son</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>#15</td>
<td>MF</td>
<td>57</td>
<td>62</td>
<td>brother</td>
<td>NHL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>#16</td>
<td>FF</td>
<td>24</td>
<td>69</td>
<td>mother</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>#17</td>
<td>FM</td>
<td>24</td>
<td>45</td>
<td>mother</td>
<td>HL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>#18</td>
<td>MM</td>
<td>25</td>
<td>51</td>
<td>father</td>
<td>NHL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>#19</td>
<td>FM</td>
<td>47</td>
<td>38</td>
<td>sister</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>#25</td>
<td>MM</td>
<td>62</td>
<td>35</td>
<td>son</td>
<td>NHL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>#26</td>
<td>FF</td>
<td>42</td>
<td>49</td>
<td>sister</td>
<td>NHL</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>#27</td>
<td>MF</td>
<td>40</td>
<td>48</td>
<td>cousin</td>
<td>HL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>#28</td>
<td>MF</td>
<td>21</td>
<td>52</td>
<td>father</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>#33</td>
<td>MM</td>
<td>57</td>
<td>47</td>
<td>brother</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>#34</td>
<td>F M</td>
<td>33</td>
<td>64</td>
<td>mother</td>
<td>AML</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>#35</td>
<td>MM</td>
<td>50</td>
<td>22</td>
<td>brother</td>
<td>HL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>#37</td>
<td>MM</td>
<td>54</td>
<td>55</td>
<td>mother</td>
<td>NHL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>#38</td>
<td>FM</td>
<td>44</td>
<td>27</td>
<td>cousin</td>
<td>HL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>#42</td>
<td>MF</td>
<td>46</td>
<td>35</td>
<td>brother</td>
<td>ALL</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>#47</td>
<td>MF</td>
<td>40</td>
<td>45</td>
<td>brother</td>
<td>NHL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>#48</td>
<td>FM</td>
<td>24</td>
<td>51</td>
<td>mother</td>
<td>ALL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>#49</td>
<td>MF</td>
<td>56</td>
<td>25</td>
<td>brother</td>
<td>ALL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>#50</td>
<td>MM</td>
<td>49</td>
<td>56</td>
<td>cousin</td>
<td>MDS</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>#52</td>
<td>MM</td>
<td>58</td>
<td>53</td>
<td>brother</td>
<td>HL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>#56</td>
<td>FF</td>
<td>28</td>
<td>29</td>
<td>sister</td>
<td>HL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>#57</td>
<td>MF</td>
<td>68</td>
<td>46</td>
<td>father</td>
<td>NHL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>#58</td>
<td>MF</td>
<td>70</td>
<td>59</td>
<td>brother</td>
<td>AML</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>35</td>
<td>#59</td>
<td>MM</td>
<td>70</td>
<td>32</td>
<td>son</td>
<td>MDS</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>36</td>
<td>#60</td>
<td>MM</td>
<td>20</td>
<td>47</td>
<td>father</td>
<td>NHL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>#62</td>
<td>FM</td>
<td>59</td>
<td>21</td>
<td>son</td>
<td>NHL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>38</td>
<td>#63</td>
<td>MM</td>
<td>25</td>
<td>28</td>
<td>brother</td>
<td>HL</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>39</td>
<td>#65</td>
<td>MM</td>
<td>69</td>
<td>57</td>
<td>son</td>
<td>MDS</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>#66</td>
<td>MF</td>
<td>68</td>
<td>30</td>
<td>daughter</td>
<td>AML</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>41</td>
<td>#79</td>
<td>FM</td>
<td>37</td>
<td>40</td>
<td>sister</td>
<td>AML</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>42</td>
<td>#76</td>
<td>MF</td>
<td>56</td>
<td>33</td>
<td>son</td>
<td>NHL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>43</td>
<td>#77</td>
<td>MF</td>
<td>26</td>
<td>62</td>
<td>father</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>#78</td>
<td>FF</td>
<td>48</td>
<td>41</td>
<td>cousin</td>
<td>MDS</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>#84</td>
<td>MF</td>
<td>52</td>
<td>25</td>
<td>son</td>
<td>MFI</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>46</td>
<td>#88</td>
<td>FM</td>
<td>50</td>
<td>44</td>
<td>mother</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>47</td>
<td>#89</td>
<td>FF</td>
<td>30</td>
<td>42</td>
<td>sister</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>48</td>
<td>#105</td>
<td>FF</td>
<td>47</td>
<td>49</td>
<td>sister</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>49</td>
<td>#106</td>
<td>MF</td>
<td>31</td>
<td>21</td>
<td>brother</td>
<td>HL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>#118</td>
<td>MM</td>
<td>47</td>
<td>44</td>
<td>brother</td>
<td>ALL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

d: donor; r: recipient; f: female; m: male; 0: no reactivation; 1: clinical reactivation; 2: subclinical reactivation; m: multiple reactivation; s: single reactivation; ALL: acute lymphoid leukemia; AML: chronic lymphoid leukemia; MDS: myelodysplastic syndrome; MFI: myelofibrosis; HL: hodgkin lymphoma; NHL: non-hodgkin lymphoma; BM: bone marrow; PB: peripheral blood; m: months.

Last column underlines which experiments pt samples have been used for: unCD56dim: results of the paragraph 9.1; CMV: results of the paragraph 9.2; RNAseq: results of the paragraph 9.3.
Moreover, we collected graft sample (leukapheresis or BM) the transplantation day. Blood samples from age- and sex-matched healthy controls (HCs) were used as controls. More detailed patient clinical features are shown in Table 1.

In addition, according to our institutional protocol, patients receiving h-HSCT were monitored for the HCMV infection/reactivation and were treated with a prophylactic therapy followed by a preemptive one in case of viral reactivation. Thus, 500 mg/m² of Acyclovir, three times per day were administered starting from the 6th day before the transplant. Then, HCMV viral load was assessed by Real-time PCR (CMV R-GENE, Argene, Biomérieux, Florence, Tuscany, Italy) in the PB of both donor and recipients before and every week after the transplant to identify those patients experiencing HCMV infection/reactivation.

Patients showing a viral load higher than 4000 IU/mL, at least at one time point after the transplant, were considered as HCMV reactivated recipients and were treated with preemptive therapy based on Foscarnet or Gancyclovir (and/or Vangancyclovir), 90 mg/kg and 5 mg/kg respectively, twice a day for two weeks. The choice between the two drugs was done considering renal function and hematologic parameters, owing to the medullary toxicity of Gancyclovir and renal toxicity of Foscarnet [331].

### 8.2 MNC isolation from peripheral blood and bone marrow

PB/BM samples were withdrawn in litium heparin as anticoagulant from anonymous HCs, or from h-HSCT recipients and their corresponding healthy donors (HDs). PB samples were diluted with sodium chloride physiological solution (NaCl) in a 1:1 ratio (volume-to-volume) and mixed gently. BM samples were collected from HDs, diluted in PBS/- with 2mM of EDTA, and processed similarly to PB samples.

Peripheral blood mononuclear cells (PBMCs) were collected from patients, HDs, and HCs; while BM-derived mononuclear cells (MNCs) were obtained only from HDs. PBMCs and MNCs were isolated by density-gradient centrifugation using Lympholyte®-H Cell Separation Media (Cedarlane, Burlington, North Carolina, USA). Briefly, 30 mL of diluted blood were stratified on 15 mL of Lympholyte®-H Media in 50mL conical tube (Corning, Reynosa, Tamaulipas, México) and centrifuged for 30 minutes (min) at 400 relative centrifugal force (rcf) at room temperature (RT) without brake.
The interphase ring, containing PBMCs and MNCs between Lympholyte®-H and plasma, was collected and washed with NaCl or PBS/- with 2mM of EDT, respectively. When necessary, a further centrifuge at 200 rcf for 10 min was run to remove platelets. The obtained cells were then counted and, eventually, frozen in cryovials (Globe Scientific INC., Paramus New Jersey, USA) in a final volume of 1 mL of Fetal Bovine Serum (FBS) (Lonza, Basel, Switzerland) containing 10% Dimethyl Sulfoxide (DMSO) (Sigma Aldrich, St. Louis, Missouri, USA), as cryoprotectant.

8.3 Cell culture

8.3.1 K562

The human leukemic cell line K562 (ATCC® CCL-243™) is a lymphoblastoid immortalized cell line derived from a patient affected by CML in blast crisis. The lack of expression of any MHC-I molecule makes this cell line a highly sensitive in vitro target for NK cell cytotoxicity assays. K562 cells were obtained from Clinical and Research Institute Humanitas Cells Bank and tested for the absence of mycoplasma infection before the usage. Cells were cultured in Iscove's Modified Dulbecco's Media (IMDM) (Lonza) supplemented with: 10% FBS (Lonza); 1% Penicillin-Streptomycin (Invitrogen, Paisley, UK); 1% Ultra Glutamine (Lonza) in cell flasks with filter cups (Corning Inc., New-York, USA) at a density of $10^5$-$10^6$ cells/mL in a humidified atmosphere at 37°C with 5% CO₂. Cells were split every 2/3 days to allow cell expansion, using fresh medium.

8.3.2 721.221.G

Mycoplasma free B-lymphoblastoid cell line 721.221-G was obtained from Clinical and Research Institute Humanitas Cells Bank. These cells derive from the LCL 721.221 (ATCC® CRL-1855™) cell line transfected with the complementary (cDNA) encoding for HLA-G1, which allows the surface exposure of the HLA-E molecule [332], a major ligand for the CD94/NKG2 receptors. 721.221-G cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 Complete Medium (Lonza) consisting of RPMI supplemented with 10% FBS (Lonza); 1% Penicillin-Streptomycin (Invitrogen, Carlsbad, California, USA); 1% Ultra Glutamine (Lonza), and were maintained in the same condition culture as K562 cell line.
8.3.3 HUVEC

The EndoGRO human humbelical vein endothelial cells (HUVEC) (Millipore, Burlington, MA) were cultured in endothelial cell growth medium-2 (EGM-2) (Lonza, Basel, Switzerland) supplemented with: 5% FBS (Lonza); 1% Penicillin-Streptomycin (Invitrogen, Paisley, UK); 1% Ultra Glutamine (Lonza) in cell flasks with filter cups pre-coated with Collagen I (Corning, New York, USA) at a density of 10^5-10^6 cells/mL in a humidified atmosphere at 37°C with 5% CO₂. HUVEC were treated with IFN-γ (100 ng/mL) for three days in order to induce the expression of PDL1 and HLA-II and further used for as target cells in functional assays.

8.3.4 NK cell isolation from PBMCs

Frozen cells were firstly thawed in RPMI complete medium (Lonza), then washed in Hank’s Balanced Salt Solution without Calcium (Ca), Magnesium (Mg) and Phenol Red (HBSS -/-) (Lonza), both containing 50 U/mL Benzonase nuclease (Sigma) to avoid cell clumping.

In order to isolate NK cells, PBMCs were thawed the day before the usage and were maintained overnight in a humidified atmosphere, at 37°C with 5% CO₂, in RPMI (Lonza) complete medium supplemented with 100 U/mL of recombinant human (rh)IL-2 (Miltenyi Biotec, Bergisch Gladbach, Germany) at 1 x 10^6 cell/mL.

The day after PBMCs were centrifuged at 400 rcf for 5 min and re-suspended in RoboSep Buffer (Stem cell technologies, Vancouver, BC, Canada) at 5 x 10^7 cell/mL. Then, the EasySep™ Human NK Cell Enrichment Kit (Stem cell technologies) for NK cell negative selection was used according to the manufacturer’s instructions. Briefly, PBMCs were labeled with the antibody-mix (CD3, CD4, CD14, CD19, CD20, CD36, CD66b, CD123, HLA-DR, glycophorin A and dextran), bound by magnetic particles, and separated from NK cells using an EasySep™ magnet.

Alternatively NK cells were obtained by CD3, CD20, and CD14 manual depletion purifying samples from T lymphocytes, B cells, and monocytes, respectively. Therefore, PBMCs were marked with anti-CD3, anti-CD20, anti-CD14 antibodies APC-conjugated for 20 min. Cells were washed with HBSS -/- (Lonza) supplemented with 2% FBS (hereafter FACS Buffer), re-suspended in Buffer Beads (PBS+0,5%BSA+2mM EDTA), and incubated with anti-APC Microbeads (Miltenyi, Biotec) for 15 min at 4°C. LD columns (cat. 130-042-901) were used for negative NK cell selection.
After both procedures, NK cells were poured off into a new tube, counted and 0.5 x 10^6 NK cells were used for flow cytometry purity check. Successively, purified NK cells were cultured as described for PBMCs for the indicated periods of time.

### 8.3.5 Stromal isolation from chips

From January 2019 we enrolled 15 patients undergoing a primary total hip replacement thanks to the collaboration with the Hip Orthopedic Unit of the Ortho Center (Clinical and Research Institute Humanitas, Dr G. Grappiolo and Dr M. Loppini). At the time of surgery, up to 20 mL of PB and 50 mL of BM were collected, mononuclear cells were immediately isolated by density gradient centrifugation, and cryopreserved for further applications as previously detailed (*MNC isolation from peripheral blood and bone marrow paragraph*). Moreover, bone fragments obtained from hip surgery were collected, and subjected to an enzymatic digestion, in RPMI complete medium with 3 mg/mL of collagenases D (Roche, Basel, Switzerland) for 3 hours [333], and to a mechanic dissociation by a gentleMACS dissociator (Miltenyi Biotec, Bergisch Gladbach, Germany). Bone chip samples were then passed through a 100 µm cell strainer and washed with NaCl. Bone chip cell suspension was incubated with Ammonium-Chloride-Potassium (ACK) solution for 5 min to lyse erythrocytes and then seeded in MESENCELL medium (StemCell Technologies, Vancouver, Canada). Cells were daily observed to the microscope to detect the appearance of adherent fibroblastoid-like cells. Fresh medium was added every 2 days. Cells were splitted when the 70-80% of confluence was reached and maintained in culture for 4-5 passages [334].

### 8.4 Polychromatic flow cytometry staining, phenotyping and cell sorting

All the flow cytometry experiments were performed in batch on frozen cells to minimize variability and were acquired at BD FACS Fortessa and Symphony flow cytometer (San Jose, California, USA); while FACS-sorting experiments were performed on 2-4 samples at a time because of the instrument availability through BD FACS Aria™ III.

The anti-human monoclonal antibodies (mAbs) listed in Tables 2 were used for the analysis of NK cell compartment of both h-HSCT patients and HDs or HCs.

The mAbs used in all flow cytometry panels were purchased from the following companies: BioLegend (St. Diego, California, USA), BD Biosciences (San Jose, California, USA), Beckman...
Coulter (Brea, California, USA), eBioscence (St. Diego, California, USA), R&D (Minneapolis, Minnesota, Stati Uniti) and Miltenyi (Bergisch Gladbach, Germany).

All antibodies were titrated on human PBMCs and used at the concentration giving the highest signal-to-noise ratio over background, as described [335-337].

In all experiments, cells were stained for 15 min at RT with LIVE/DEAD™ Fixable Aqua Dead Cell Stain Kit (Invitrogen) to eliminate dead cells, which may influence the analysis. At the end of the incubation, cells were washed with FACS Buffer and centrifuged for 5 min, at 400 rcf. Pellet cells were then stained for 20 min at RT with a combination of mAbs specific for surface proteins prepared in FACS Buffer, in a final volume of 100 ul (for a maximum of 10 x 10⁶ cells).

### Table 2: Lists of antibodies used in flow cytometry panels

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Clone</th>
<th>Staining</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-CD107a</td>
<td>H4A3</td>
<td>SUPERFICIAL</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>Anti-CD117</td>
<td>104D2D1</td>
<td>SUPERFICIAL</td>
<td>BioLegend</td>
</tr>
<tr>
<td>Anti-CD127</td>
<td>A619D5</td>
<td>SUPERFICIAL</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>Anti-CD14</td>
<td>M5E2 or 7AU1</td>
<td>SUPERFICIAL</td>
<td>BioLegend</td>
</tr>
<tr>
<td>Anti-NKG2A</td>
<td>REA110 or Z199</td>
<td>SUPERFICIAL</td>
<td>Miltenyi</td>
</tr>
<tr>
<td>Anti-CD16</td>
<td>3G8</td>
<td>SUPERFICIAL</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>Anti-CD19</td>
<td>LT19</td>
<td>SUPERFICIAL</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>Anti-CD20</td>
<td>2H7</td>
<td>SUPERFICIAL</td>
<td>BioLegend</td>
</tr>
<tr>
<td>Anti-CD3</td>
<td>OKT3 or SK7</td>
<td>SUPERFICIAL</td>
<td>BioLegend</td>
</tr>
<tr>
<td>Anti-NKG2D</td>
<td>1D11</td>
<td>SUPERFICIAL</td>
<td>BioLegend</td>
</tr>
<tr>
<td>Anti-NKp46</td>
<td>9E2</td>
<td>SUPERFICIAL</td>
<td>BioLegend</td>
</tr>
<tr>
<td>Anti-NKp30</td>
<td>P30-15 or Z2A</td>
<td>SUPERFICIAL</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>CD34</td>
<td>581</td>
<td>SUPERFICIAL</td>
<td>BioLegend</td>
</tr>
<tr>
<td>CD56</td>
<td>B159 or NCAM16-2</td>
<td>SUPERFICIAL</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>CD8</td>
<td>3B5</td>
<td>SUPERFICIAL</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>Anti-HLA-A2</td>
<td>BD7.2</td>
<td>SUPERFICIAL</td>
<td>BioLegend</td>
</tr>
<tr>
<td>Anti-Granzyme-β</td>
<td>GB11</td>
<td>INTRACELLULAR</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>Anti-K67</td>
<td>B56</td>
<td>INTRACELLULAR</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>Anti-Perforin</td>
<td>8G9</td>
<td>INTRACELLULAR</td>
<td>BioLegend</td>
</tr>
</tbody>
</table>

The Cytofix/Cytoperm™ kit (BD Biosciences) was used according to the manufacturer's protocol to detect intracellular IFN-γ, granzyme-β, perforin, and Ki-67 expression.

In the experiments concerning cell IR and unCD56dim phenotypic and functional analysis, total NK cells were defined as CD3neg/CD14neg/CD20neg/NKG2Dneg lymphocytes. The same gating
strategy was used for cell sorting experiments where cCD56<sup>br</sup>, unCD56<sup>dim</sup>, cCD56<sup>dim</sup> NK cell subsets were cultured with different cytokine stimulation (see paragraph *NK cell subset culture with cytokine stimulation*).

In further experiments assessing HCMV impact on h-HSCT patients, total NK cells were defined as CD14<sup>neg</sup>/CD3<sup>neg</sup>/Lineage<sup>neg</sup> (CD4, CD15, CD20, CD19, CD33, CD34, CD203c, FcɛRI) viable lymphocytes as defined in the *Results* section. The same gating strategy was used to FACS-sort NK cell subsets enriched in patients experiencing HCMV infection/reactivation used for Real-Time PCR and/or RNA-sequencing analysis.

Finally, NK cell subsets were always identified through the differential expression of CD56 and CD16. (Figure M.2).

### 8.5 Functional assays

For degranulation assays, PBMCs were thawed and left overnight in RPMI complete medium supplemented with 100 IU/mL of rhIL-2 at 37°C in 5% CO₂, then re-suspended at 1 x 10<sup>6</sup> cells/mL and cultured with K562 cell lines for 4 hours at PBMCs/target ratio 1:1 in the presence of anti-CD107a antibody and Golgi Plug (BD Bioscience).

For Y9 masking experiments, PBMCs were left 2 hours in RPMI complete medium after thawing, and re-suspended at 4 x 10<sup>6</sup> cells/mL in the presence of Y9 mAb (IgM, anti-CD94) in
5% CO₂ humidified atmosphere at 37°C for 30 min. For PD-1 and LAG3 blocking experiments, PBMC were thawed and left overnight in RPMI complete medium supplemented with rhIL-2 (100 IU/mL) and rhIL-12 (100 IU/mL) at 37°C in 5% CO₂; the day after, PBMCs were incubated in the presence of blocking mAbs (10 ug/mL): α-PD-1 (EH12.27H7; 329926; BL) or α-LAG3 (17B4; AG-20B-0012PF; AdipoGen) in 5% CO₂ humidified atmosphere at 37°C for 30 minutes [338]. Subsequently, PBMCs were seeded at 1 x 10⁶ cells/mL with 200 U/mL of rhIL-2 in the presence of anti-CD107a mAb and HLA-Epos 721.221.G or HUVEC target cells at a PBMCs/target ratio of 5:1 [332]. After 4 hours of co-culture, NK cell degranulation capability was assessed evaluating the CD107a expression and IFN-γ production by flow cytometry staining [339]. Y9 mAb was kindly provided by A. Moretta [340].

8.6 NK cell subset culture with cytokine stimulation

NK cell subsets (cCD56br, unCD56dim, cCD56dim) from HCs were FACS-sorted through BD FACS Aria™ III, stained with 2uM of 5-(and 6)-carboxyfluorescein diacetate succinimidyl ester (CFSE; Life Technologies, Carlsbad, California, USA) for 7 min at 37°C, then washed with double volume of cold FBS (Lonza) and RPMI complete medium. Stained cells were left at 4°C for 1 hour and successively seeded at 0.1 x 10⁶ cells/mL in a 96 multiwell plate and cultured in presence of cytokines. Cytokine mix stimulation consisted in: 50 ng/mL of rhIL-6 (Peprotech, Rocky Hill, New York, USA); 30 U/mL of rhIL-15 (Miltenyi Biotec); 100 ng/mL of rhIL-18 (MBL, Woburn, Massachusetts, USA); 25 ng/mL of G-CSF (Peprotech); 50 ng/mL of MCP1/CCL2 (Peprotech) added from the beginning of the cell culture; while 100 U/mL of rhIL-2 (Miltenyi Biotec) and 50 ng/mL of rhIL-7 (Peprotech) were added after 8 days of culture.

Alternatively, only rhIL-15 and rhIL-18 were used at the same concentrations. Cells were cultured in a humidified atmosphere at 37°C with 5% CO₂ and one-half of the medium was exchange every 3 days with fresh cytokine addition. The cell culture was followed for 14 days. NK cell subset proliferation was determined by the analysis CFSE dilution and proliferation index was assessed as previously described [341].

8.7 RNA isolation and cDNA synthesis

The FACS-sorted cells were collected in RNAse free tube (Bisosphere®, Nümbrecht, Germany) containing 200 μL of RPMI complete medium. 10 μl of the sorted cell suspension
were collected, to check sorted cell purity: only the samples with a purity higher than 95% were then processed for RNA extraction, as detailed below.

Thereafter, cells were washed with HBSS -/- (Lonza), centrifuged at 650g for 10 min at 4°C twice, and re-suspended in 49 µL of RLT buffer (Qiagen®, Hilden, Germany) and 1 µL of RNase inhibitor (Thermo Fisher Scientific®, Woolston, Warrington, UK). The sorted samples were then cryopreserved at -80°C till RNA extraction.

To avoid batch effect among samples, they were randomized during RNA isolation and sample preparation working sessions. All samples were processed with the same reagents lot number, when available.

For RNA extraction we used the MicroRNAeasy Kit™ (Qiagen®), according to the manufacturer’s instructions. Briefly, further RLT lysis buffer was added to the cryopreserved cell lysates in order to avoid the RNA degradation. The samples were centrifuged and ethanol 70% was added to induce the precipitation of the released RNA. Samples were transferred to the RNeasy MinElute® Spin Columns. In the following steps, subsequent to each solution added, the samples were centrifuged and the liquid collected in the tube underneath was discarded.

The RW1 buffer used as a stringent washing buffer, was added to the samples, preceding and subsequent to their incubation with DNAse (for 15 min at RT) to eliminate the co-purified DNA, thus obtaining a highly-purified RNA sample.

Next, RPE buffer was added, followed by ethanol 80%. Finally, the tube underneath was changed with a new RNAse-free 1.5 mL tube (Bisosphere®), and 20 µL of RNAse-free H2O was added to the column and left 1 min, then centrifuged for 1 min at maximum speed in order to elute the RNA. The extracted RNA was collected in the tube underneath and the column was discarded.

4 µL of each sample were transferred in a 0.5 mL RNAse-free tube for RNA quantification at Nanodrop 2000 (Thermo Fisher Scientific®, Waltham, Massachusetts, USA). The resting volume was kept at -80°C for further semi-quantitative Real-Time PCR or RNA-sequencing (-seq).

To perform Real-Time PCR experiments, complementary DNA (cDNA) was synthesized by RNA reverse transcription reaction. To preserve the integrity of the extracted RNA, all the steps of this process were performed on ice, and the used pipettes and surfaces were decontaminated with RNAse ZAP™ (Invitrogen).
RNA samples were diluted to 100 ng/15 µL with RNAse-free H$_2$O (Qiagen®) in RNAse-free PCR tubes (Thermo Fisher Scientific®). Afterwards, the reverse transcription mix was prepared in RNAse-, DNAse-free 0.5 mL tube by adding the following reagents provided within the High Capacity cDNA Reverse transcription Kit (Thermo Fisher Scientific®): RT buffer 10x, dNTP mix, RT random primers, RNAse inhibitor and MultiScribe® Reverse Transcriptase. 15 µL of the mix was added to each sample and mixed well pipetting. The tubes were then sealed and loaded into the thermal cycler.

According to the manufacturer’s instructions, the thermal cycler was set as follow:
- 25°C for 10 minutes;
- 37°C for 120 minutes;
- 85°C for 5 minutes;
- 4°C held until the samples were collected then kept at -20°C.

### 8.8 Real-Time PCR

Specific TaqMan probes to our target genes (coding for PLZF, EAT-2, SYK, FcεRγ) and the housekeeping gene GAPDH were used. All the following steps were performed on ice to preserve the integrity of cDNA samples and the other reagents.

Each sample of cDNA was diluted in RNAse-free H$_2$O (Qiagen®) to a final volume 37.5 µL sufficient to investigate the expression of the 5 genes in triplicates. The Real-Time PCR mix was prepared to investigate each gene separately and comprised: i) TaqMan™ Universal PCR Master Mix; ii) the target gene specific TaqMan probe/primer set (both provided by Thermo Fisher Scientific®); iii) RNAse-free H$_2$O (Qiagen®). The mix was aliquoted in an RNAse-, DNAse-free MicroAmp™ 96 multiwell plate (Applied biosystems®) prior to adding the cDNA samples. The plate was kept in the dark at 4°C until the running of the Real-Time PCR.

The high throughput 7900HT Fast Real-Time PCR system (Thermo Fisher Scientific®) was used for our experiments, consisting in different phases:
- 95°C for 15 minutes, enables the activation of the Taq polymerase by ending the effect of the polymerase inhibitor;
- 95°C for 15 seconds (melt and annealing phase);
- 60°C (elongation phase).
The last two stages were repeated 40 times to amplify the input cDNA, measuring meanwhile the emitted signal from the fluorescent reporter of the TaqMan probes. At the end of the run, data were analysed calculating the relative expression using the $2^{-\Delta\Delta Ct}$ method.

### 8.9 RNA sequencing

RNA sequencing (RNA-seq) is an application of Next Generation Sequencing (NGS) that enables a deep investigation of gene expression. Indeed, through RNA-seq technology it is possible to quantify the abundance level or relative changes of each transcript under specific conditions [342], thus allowing the comparative analysis of the global expression level in different conditions, as in this proposed thesis.

RNA-seq workflow can be divided in three sections: Experimental Biology, Computational Biology, and Systems Biology [343] (Figure M.3).

Indeed, the complexity of databases generated by RNA-seq technology requires powerful computational programs in constant development.

_Eperimental Biology_ consists in RNA extraction and reverse transcription (RNA isolation and reverse transcription paragraph) followed by library construction resulting in millions of short reads [344]. The amount and the quality of extracted RNA were assessed by Qubit4 (Invitrogen) instrument. Only RNA samples with an RNA integrity (RIN) >6 were successively analysed. Preparation and processing of libraries were performed through QuantSeq 3’ mRNA-Seq Library Prep Kit (Lexogen protocol). Illumina compatible libraries of sequences close to the 3’ end of poly-(A) RNA were generated and NextSeq500 System (Illumina, Cambridge, UK) was used to perform sequencing, producing an average of 60M reads per sample (single-end).

_Computational Biology_ uses the raw reads obtained by experimental procedures as the starting point for further bioinformatic analysis. It consists in the pre-processing and quality
assessment of the reads, followed by the alignment to a reference genome and the normalization of the data (see *Data analysis paragraph*).

*Systems Biology* approach allows the analysis of the variability within the experiments and of the differential gene expression. It implements pathway or network level analyses to gain biological insight to RNA-seq data (see *Data analysis paragraph*).

### 8.10 Data Analysis

#### 8.10.1 Flowjo

Flow Cytometry Standard (FCS) 3.0 files, acquired through both FACS Fortessa and Symphony A5 flow cytometers, were analysed using the FlowJo software (TreeStar Inc, Ashland, Oregon, USA), versions 9.9.6 and 10.2.0.

#### 8.10.2 Unsupervised high-dimensional cytometry data analysis

Flow cytometry data are usually analysed plotting cells in two-dimensional plots and gating them manually to identify cell population and their phenotype [345]. This manual strategy is still adequate to investigate few parameters on a single sample; however, several weaknesses arise when a more complex flow cytometry panel is applied [346]. Indeed, as the number of measured parameters increases, the number of two-dimensional plots increases exponentially, and the manual analysis become inefficient and always more subjective [345, 346].

To avoid these problems, novel computational methods have been developed in the recent years. They allow the automatically identification of cell populations in multi-dimensional flow cytometry data combining all marker information at the same time and assessing similarities among cells through a more unbiased and unsupervised approach [345, 346]. A computational method applied to the analyses of complex flow cytometry panels, resembling to that used to analyse high-dimensional single cell data, has been recently developed [347]. Briefly, flow cytometry raw data were pre-processed with FlowJo software version 9.9.6 and FCS files were analysed by standard manual gating strategy in order to remove debris and doubles, dead cells, and to identify NK cells (Figure M.1B). Afterwards, 2000 NK cells from each sample were imported in FlowJo software version 10.2.0 and bi-exponentially transformed for further analysis in R (version 3.3.3). Each sample was labeled with a unique computational barcode, for further identification, converted in comma
separated (CSV) files and concatenated in a single matrix using “cytof_exprsMerge”, an R function [348].

8.10.3 Cluster identification and analysis with PhenoGraph

In collaboration with the Bioinformatic Unit of Clinical and Research Institute Humanitas, concatenated data were analysed using the PhenoGraph algorithm coded in the cytofkit package (version 1.6.5) [348]. This graph-based clustering algorithm is able to convert single-cell data into a graph that represents the phenotypic relationships among cells [349]. It takes a matrix of \( n \) single-cell measurements as input and convert them into a k-nearest neighbour graph (k-NNG) output, in which each cell is a node connected to other similar cells, defined as neighbours [346, 349]. Phenotypically similar cells are grouped in clusters by a single user-defined constant that is the “k” value corresponding to the final number of clusters identified as “closest neighbours” [349-351]. This means that lower is the “k” parameter, higher is the final number of clusters, since it is able to detect very little differences among NK cells; contrary, a “k” parameter of “100” enables the generation of fewer clusters of identical NK cells reducing the effect of the noise. Indeed, PhenoGraph runs different values for the parameter “k” (sequentially set to 15, 30, 45, 60, 75, and 100) and the best choice of “k” depends upon the data.

Data obtained using PhenoGraph were reorganized and saved as new FCS files, one for each sample, containing information about cell clustering. These new FCS files were further

![Figure M.4: high dimensional single-cell analysis of immune reconstituting NK cells in patients after h-HSCT](image)

Heatmap showing the 28 clusters of phenotypically identical NK cells identified by PhenoGraph software (k score=45) and their corresponding markers expression (expressed as integrated MFI).
analysed using FlowJo version 10.2.0. Clusters representing <0.5% were excluded from the analysis. For the selected “k” value we created an ad hoc heatmap showing the identified clusters and their markers expression, calculated as integrated Median of Fluorescence Intensity (iMFI = frequency of positive cells for each marker multiplied for the corresponding MFI) of markers within all clusters [347] (Figure M.4).

Finally, PhenoGraph clusters were visualized using the t-stochastic neighbour embedding (t-SNE) technique, which is a non-linear dimensionality reduction algorithm well-suited for embedding high-dimensional data for visualization in a two-dimensional scatter plot (t-SNE map), named t-SNE1 and t-SNE2.

8.10.4 RNA-seq analysis

8.10.4.1 Pre-processing and quality assessment

The reads were stored in files, typically FASTQ format, which includes the nucleotide sequences and the quality values for each nucleotide [352]. Sample quality is the determinant of qualified data and derived biological insights; thus, quality assessment is the first step of bioinformatic pipeline before every other analysis of RNA-seq [343]. RNA quality control was performed with the Agilent 2200 Tape Station system and only RNAs having a RIN >6 were used for library preparation. All samples were sequenced on an Illumina NextSeq 500 at an average of 60M 75-bp single-end reads.

8.10.4.2 Mapping, counting, and normalization

Once high-quality data are obtained from preprocessing, the short reads were mapped to the reference genome (GRCh38.p12) using the STAR aligner with default parameters (version 2.7.0). Once the reads are mapped to the reference genome, they are counted to facilitate the next steps.

Gene-based read counts were then obtained using HTSeq count module (version 0.11) and GENCODE v29.gtf annotation [353]. The read counts were imported into R statistical software (http://www.r-project.org/) and differential gene expression analysis was performed using the edgeR package (version 3.22). The raw read counts can contain artifacts or errors. Therefore, for pair-wise comparisons, raw read counts were normalized using the TMM method (trimmed mean of log-ratio values) and genes that failed to achieve a counts per million (CPM) mapped reads
value greater than 1 in at least two libraries were not considered [354]. P-values were adjusted using the Benjamini-Hochberg method.

8.10.4.3 Variability within the experiments
A dendrogram and a multidimensional scaling plot were drawn to assess the main variability within the experiment. The hierarchical clustering of the whole sample set was obtained after a transformation of the count data as moderated log-counts-per-million. Euclidean distance was computed between samples, then the dendrogram was built upon the Ward criterion. Similarly, a multidimensional scaling plot was used to look at the experiment variability. The first dimension is expected to separate samples from the different biological conditions, meaning that the biological variability is the main source of variance in the data.

8.10.4.4 Differential gene expression analysis
Differential gene expression analysis was assessed on TMM normalized data by EdgeR (v3.24.3). Significant differential expression in each gene was tested using the QL F-test; we selected only genes with a False Discovery Rate (FDR, p-value adjusted considering Benjamini-Hochberg correction) <0.05.

For each comparison, genes were plotted in a MA-plot that allows to visualize the differences between two cohorts of samples, by transforming the data onto M (log-intensity ratio) and A (log-intensity mean average) scales [355]. Instead, volcano plot (significance, y axis, versus fold change (FC), x axis; significance and FC were reported in log scale (log_{10} and log_{2}, respectively)) was drawn only for the comparison between total not reactivated patients and HCMV reactivated ones. Gene expression values of selected genes (FDR <0.05) were plotted in heatmaps, one for each comparison.

8.10.4.5 Pathway analysis
Pathway analysis was performed both by Gene Set Enrichment Analysis (GSEA, v3.0) and on Ingenuity Pathway Analysis (IPA, v01-13) software. Lists of differentially expressed genes for each comparison and their log_{2}FC values were used for the identification of significantly enriched pathways. The gene set enrichment analysis was conducted in pre-ranked
mode with scoring scheme “classic” and 1,000 permutations. The maximum gene set size was fixed at 5,000 genes, and the minimum size fixed at 10 genes. GSEA was performed for interpreting gene expression data [356] with H (hallmark) and C2 (curated) gene sets from Molecular Signatures DataBase (MSigDB database v6.0, http://software.broadinstitute.org/gsea/msigdb/index.jsp). GMT file was downloaded with the gene symbol information.

IPA was used to interpret data in the context of upstream regulators and canonical pathways deregulated in patients experiencing HCMV reactivation compared to not reactivated ones. Dataset was filtered using a log2|FC| ≥1.5 and a p-value adjusted ≤0.05.

For GSEA only pathways with a |NES| ≥1.8 and q-value ≤0.02 were selected; for IPA only pathways with |z-score| ≥2 and log10 p-value ≥1.3 were further analysed.

8.10.5 Statistical analysis

Statistical analysis was performed using GraphPad PRISM (version 7.0, La Jolla California, USA) and SPICE 5.22 software. Paired Student’s t-test, non-parametric paired Wilcoxon rank test, and unpaired Mann-Whitney test were used to compare the different variables. One-way Anova with Bonferroni’s correction test was used to analyze the kinetics of NK cell subset proliferation and differentiation. Multiple t-test was used for the cluster frequencies upon PhenoGraph analysis. Results are represented as a mean ± Standard Deviation (SD) or as mean± Standard Error of the Mean (SEM) when repeated sampling were considered. P values are two-sided and were considered significant when p ≤0.05.
9 RESULTS

9.1 The early recovery of unCD56\textsuperscript{dim}/NKG2A\textsuperscript{pos} NK cells in h-HSCT patients

NK cells represent the first lymphoid cell compartment to recover after h-HSCT arising in the first 2 weeks from transplantation [148] and the single-cell analysis of the mismatched HLA-A2 in circulating NK cells reveals that their chimerism is completely donor-derived after only 28 days from h-HSCT [238]. In h-HSCT patients, the profound understanding of immune cell IR mechanisms is key in governing positive clinical outcomes and side effects occurring after transplantation. Therefore, we deeply investigated the kinetics of circulating NK cell subset IR at different time points post-h-HSCT. We performed a detailed analysis of unCD56\textsuperscript{dim} NK cells since transiently, but highly expanded, in our h-HSCT patients contrary to healthy subjects.

9.1.1 Preferential unCD56\textsuperscript{dim} NK cell subset expansion in the first weeks after h-HSCT

PBMCs from patients were stained with an extensive flow cytometry panel, the single cell suspension was acquired by BD FACS Fortessa and analysed by FlowJo software. Patient

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure.png}
\caption{kinetic of NK cell subset IR after h-HSCT}
\end{figure}
(A) Representative example of flow cytometry dot plots showing the kinetic of donor-derived NK cell subset distribution in a patient (pt #18) at different time points from h-HSCT. (B) Summary statistical graphs showing the frequency (median ± SD on total NK cells) of conventional NK cell subsets (cCD56\textsuperscript{br} and cCD56\textsuperscript{dim}) (upper panels) and unconventional NK cell subpopulations (unCD56\textsuperscript{br} and unCD56\textsuperscript{dim}) (lower panels) in the peripheral blood (PB) and bone marrow (BM) of healthy donors (HDs; n=30) compared to their relative recipients (n=30) at different time points after h-HSCT. Unpaired t-test vs HD-PB; * p <0.05.
samples were compared to both PB and BM of their relative HDs. NK cells were defined as single, viable, CD3<sup>neg</sup>/CD14<sup>neg</sup>/CD20<sup>neg</sup>/NKG2D<sup>neg</sup> lymphocytes, while the NK subsets were identified on the basis of the relative expression of CD56 and CD16 (Figure M.2). In both HDs and h-HSCT patients we detected four distinct NK cell subsets, whose relative frequencies are extremely different within the time points analysed (Figure R.1). We firstly noticed that conventional NK cell subset (cCD56<sup>br</sup> and cCD56<sup>dim</sup>) frequencies early after h-HSCT are opposite to those observed in physiologic conditions. In particular, cCD56<sup>br</sup> NK cells arise starting from the 3<sup>rd</sup> week from the transplant and outnumber cCD56<sup>dim</sup> NK cells till 8 weeks post-h-HSCT. Conversely, cCD56<sup>dim</sup> NK cell subset is almost undetectable during the first weeks after the transplant, while its percentage returns similar to healthy counterparts from the 5<sup>th</sup> month after h-HSCT (Figure R.1B, upper panels). The low levels of cCD56<sup>dim</sup> NK cells are counterbalance by the unconventional CD56<sup>dim</sup>CD16<sup>neg</sup> (unCD56<sup>dim</sup>) NK cell population, expanded early after h-HSCT, while poorly represented under homeostatic conditions. Indeed, the frequency of this latter NK cell subset is statistically increased in patients starting from the 2<sup>nd</sup> week post-h-HSCT with respect to relative HDs and, along with cCD56<sup>br</sup> NK cells, represents the most prevalent NK cell population soon after h-HSCT up to 8 weeks from transplantation. Moreover, we observed that the unconventional CD56<sup>br</sup>CD16<sup>neg</sup> (unCD56<sup>br</sup>) subset is poorly represented both in HDs and in h-HSCT patients at all the time points investigated (Figure R.1B, lower panels).

9.1.2 <i>unCD56<sup>dim</sup></i> lymphocytes are bona fide NK cells

The higher frequencies of cCD56<sup>br</sup> and, in particular, unCD56<sup>dim</sup> NK cells in patients early after...
the h-HSCT compared to HDs suggested us that these subsets could play a pivotal role in NK cell ontogenesis. According to literature, NK cell development is finely regulated by a plethora of NK cell markers that identify different stages of maturation [238]. While cCD56^br NK cells are accounted in the stage 4, unCD56^dim NK cell subset is still unclassified in a specific stage of differentiation. Moreover, given that unCD56^dim cells expand early after transplantation, we also hypothesized that this subset could represent an NK cell precursor stage of differentiation. We thus investigated the expression of markers of both precursor cells and of mature NK cells. Our data showed that unCD56^dim NK cells of both HDs and h-HSCT patients do not express NK cell precursors markers (CD34, CD117, and CD127), thus suggesting that they are cells in later stages of maturation (Figure R.2A). We then found that unCD56^dim NK cells express typical mature NK cell markers, including NKG2D and NKp30, while low levels of NKp46, making this NCR an additional feature enabling a further distinction of this latter subset from conventional NK cell subpopulations (Figure R.2B). We finally observed that unCD56^dim NK cells express high levels of perforin and granzyme-β, confirming their cytolytic potential (Figure R.2C).

Given these results, we hypothesize that unCD56^dim NK cells could figure as an intermediated additional stage of development between stages 3 and 4 or between stages 4 and 5 (cCD56^dim NK cells).

9.1.3 unCD56^dim cells expanded early after h-HSCT show an impaired cytotoxicity

We later assessed the functional ability of unCD56^dim NK cell subset expressing high levels of cytolytic content in order to understand whether these cells could cover a pivotal role not only in NK cell IR, but also in mediating alloreactive responses early after h-HSCT. Beside immune-reconstituting unCD56^dim NK cells in h-HSCT patients are fully armed to efficiently kill tumour target cells (Figure R.2C), their effector-functions are significantly impaired compared to healthy controls (HCs). Indeed, we found that the degranulation ability of unCD56^dim NK cells, defined as CD107a surface expression against K562 cell line, is extremely reduced at 2 months post-h-HSCT (Figure R.3A). Conversely, cCD56^br NK cells expressing high levels of cytotoxic markers (Figure R.2C), retain their lytic potential in patients early after h-HSCT (Figure R.3A).

Given that the activation status of NK cells is regulated by dominant inhibitory NK cell receptor signaling pathways over the activating ones, we screened the expression of different
inhibitory receptors likely explaining the disablement of unCD56\textsuperscript{dim} NK cell functional responses in h-HSCT patients. We observed that the expression of CD94/NKG2A is significantly increased in h-HSCT patients compared to HCs. In particular, at 3 weeks and 2 months after transplantation the 70\% and 80\% of unCD56\textsuperscript{dim} NK cells are CD94/NKG2A\textsuperscript{pos}, respectively. The levels of this iNKR are maintained high also in cCD56\textsuperscript{br} NK cells from both HCs and h-HSCT patients, while even upregulated at 2 and 6 months after the transplant (Figure R.3B).

To understand whether high expression of those iNKRs on early immune-reconstituting NK cell subsets could affect their cytotoxic potential, we performed masking experiments by blocking CD94/NKG2A in the presence or in the absence of tumour target cells expressing their putative ligands HLA-E \textsuperscript{357, 358}. To assess the degranulation potential of NK cell subsets of both in HCs and recipients at 3 weeks and 2 months after h-HSCT, we performed a degranulation assay with the blocking mAb in the presence or in the absence of target cells.

The treatment with anti-CD94/NKG2A monoclonal antibody (Y9 mAb) shows an increased degranulative ability only in patient-derived cCD56\textsuperscript{br} NK cells; while no significant differences are observed in HCs and in the other NK cell subsets from the h-HSCT patients (not shown). Both cCD56\textsuperscript{br} and unCD56\textsuperscript{dim} NK cells of HCs exert efficient effector-functions against HLA-E\textsuperscript{pos} 721.221.G target cells that are further enhanced in presence of the masking Y9 mAb only in cCD56\textsuperscript{br}, in line with their high expression of CD94/NKG2A. Contrary, both cCD56\textsuperscript{br} and unCD56\textsuperscript{dim} NK cells purified from recipients after 3 weeks post-h-HSCT show impaired cytotoxic responses that are partly reversed by the blocking of CD94/NKG2A treatment. Of note, the degranulative behaviour of cCD56\textsuperscript{br} and unCD56\textsuperscript{dim} NK cells at 2 months after h-HSCT is

Figure R.3: cytotoxicity of unCD56\textsuperscript{dim} NK cells expressing CD94/NKG2A in patients after h-HSCT
(A) Summary statistical graph showing the percentage of CD107\textsuperscript{a}\textsuperscript{pos} (median ± SEM) on cCD56\textsuperscript{br} (blue), cCD56\textsuperscript{dim} (black) and unCD56\textsuperscript{dim} (red) from healthy controls (HCs; n=5) and patients (h-HSCT pts; n=4) after 2 months from transplantation. The spontaneous degranulation of NK cells has been subtracted for the analyses performed in the presence of K562 cell line. (B) Summary statistical graphs showing the expression of NKG2A on unCD56\textsuperscript{dim} (red) and on cCD56\textsuperscript{br} (blue) NK cells from HCs (n=52) and recipients (n=20) at different time points up to 12 months after h-HSCT (medians ± SD). Unpaired t-test vs HCs. \*p <0.05, **** p <0.0001. w: weeks; m: months.
more similar to that of HCs, where NK cells retain their cytotoxic ability against 721.221.G targets and the presence of Y9 masking mAb is no more significantly efficient (Figure R.4). Considering cCD56\textsuperscript{dim} NK cells, as expected, the blocking of CD94/NKG2A inhibitory receptor has no effect on their cytotoxic potential in HCs, while it increases the CD107a expression at 2 months after h-HSCT when also this NK cell subset starts to recover and it can show a transient high expression of CD94/NKG2A \[49, 235, 238, 359\] (Figure R.4).

These set of experiments suggest that the high increased expression of CD94/NKG2A on unCD56\textsuperscript{dim} NK cells is responsible for their reduced cytotoxicity exerted against 721.221.G targets and it represents an inhibitory checkpoint, especially early after h-HSCT. Indeed, unCD56\textsuperscript{dim} together with cCD56\textsuperscript{br} NK cells outnumber other NK cell subsets in the first period after h-HSCT and are mainly CD94/NKG2A\textsuperscript{pos} (Figure R.3B).

9.1.4 Kinetics and phenotype of highly proliferating unCD56\textsuperscript{dim} and cCD56\textsuperscript{br} NK cells

We showed that unCD56\textsuperscript{dim} and cCD56\textsuperscript{br} subsets recover prior to terminally differentiated cCD56\textsuperscript{dim} NK cells and they also express high levels of Ki-67, marker of proliferation, contrary to their counterparts in HDs (not shown;[238]). The high rate of NK cell proliferation in h-HSCT is associated to the so call “cytokine-storm” occurring usually after
transplantation and inducing immune cell differentiation and activation [234, 360]. The very early expansion of unCD56dim NK cells prompted us to hypothesize that this subset could play a key role in IR and in generating either cCD56br or cCD56dim NK cells.

To disclose the maturation/differentiation relationships intervening among NK cell subpopulations, we FACS-sorted unCD56dim, cCD56br and cCD56dim NK cells from HCs (Figure R.5A) and cultured them with different cytokines. We selected those cytokines (cytokines mix: IL-15+IL-18+IL-6+GCSF+MCP1+IL-2+IL-7) better mimicking the lymphopenic milieu occurring after allogeneic HSCT [360] and IL-15+IL-18 able alone to drive the general effect observed with the previous condition. Sorted NK cell subsets were labelled with CFSE allowing us to analysed the phenotype of proliferating cells diluting CFSE at several time points (4-8-14 days) (Figure R.5B). We firstly observed that terminally differentiated cCD56dim NK cells do not proliferate and retain their original phenotype upon both cytokine mix (not shown) and IL-15+IL18 stimulation (Figure R.5B, black). The co-culture with the cytokine mix induces strong proliferation of both unCD56dim and cCD56br NK cells after only 4 days (Figure R.6A). The analysis of CD56 and CD16 expression in proliferating cells showed that both unCD56dim and cCD56br NK cells are able to generate themselves but can also differentiate in cCD56br and unCD56dim NK cells, respectively (Figure R.6A).

Indeed, we noticed that unCD56dim NK cells are able to maintain the phenotype of parental cells, but a certain percentage they also upregulate the expression of CD56 and at the end of the culture (day 14) the majority of them are cCD56br NK cells, suggesting a sort of plasticity

Figure R.5: CD56 expression of FACS-sorted NK cell subsets upon IL-15+IL-18 stimulation
(A) Representative example from a healthy controls of flow cytometry dot plots showing the purity of FACS-sorted NK cell subsets, overlaid with CD3neg/CD19neg NK cells expressing CD56 and CD16 (gray background). (B) Representative example of flow cytometry dot plots showing the expression of CD56 on FACS-sorted and CFSE-diluting NK cell subpopulations after 8 days in culture with IL-15+IL-18.
Contrary, FACS-sorted cCD56<sup>br</sup> NK cells preferentially retain their original phenotype, while only few of them reduce the expression of CD56 becoming unCD56<sup>dim</sup> (Figure R.6A, right).

With the aim of understanding which cytokine(s) mainly governs the observed alterations in proliferating NK cell phenotype, we further tested the above-mentioned cytokines alone or in different combinations. The lonely cytokine combination that recapitulates the changes in unCD56<sup>dim</sup> and cCD56<sup>br</sup> NK cell subsets is IL-15+IL-18 (Figure R.6B). Interestingly, the ability of proliferating unCD56<sup>dim</sup> NK cells in modifying the expression of CD56 and giving origin to cCD56<sup>br</sup> subset is increased when FACS-sorted unCD56<sup>dim</sup> NK cells are cultured with IL-15+IL-18 compared to cytokine mix stimulation (Figure R.6B, left). Hence, for further analysis we focused on IL-15+IL-18 stimulation.

We then calculated the proliferation index of FACS-sorted unCD56<sup>dim</sup> and cCD56<sup>br</sup> NK cell subsets from HCs at 4, 8, and 14 days of culture with IL-15+IL-18 and we noticed higher proliferative capability of cCD56<sup>br</sup> compare to unCD56<sup>dim</sup> NK cells at all the time points considered.
even significant after 14 days of stimulation (Figure R.6C). Finally, we observed that neither FACS-sorted unCD56\textsuperscript{dim} nor cCD56\textsuperscript{br} NK cells are able to originate mature cCD56\textsuperscript{dim} NK cells (Figure R.6).

Similar results were obtained in h-HSCT patients, allowing us to perform the next analyses only on healthy volunteers from whom we can isolate higher number of PBMCs and ease FACS-sorting experiments (Figure R.7).

**Figure R.7:** patients-derived FACS-sorted NK cell subset proliferation upon IL-15+IL-18 stimulation
Summary statistical graphs showing the percentages on total NK cells (mean ± SEM) of derived-cCD56\textsuperscript{br}, unCD56\textsuperscript{dim}, and cCD56\textsuperscript{dim} phenotype upon IL-15+IL-18 stimulation of FACS-sorted unCD56\textsuperscript{dim} (upper panels) and cCD56\textsuperscript{br} (lower panels) cells from healthy controls ( – - - 4) and h-HSCT patients ( – – – – ). Unpaired t-test; * p <0.05.

In order to comprehend if derived NK cells acquire other features typical of the deriving FACS-sorted subsets, we analysed the surface levels of NKp46 and CD94/NKG2A, two markers differentially expressed on unCD56\textsuperscript{dim} and cCD56\textsuperscript{br} NK cells (Figure R.2B and R.3B, respectively). Our results demonstrated that almost all unCD56\textsuperscript{dim} generates cCD56\textsuperscript{br}/NKp46\textsuperscript{pos} NK cells, while a low percentage of cells retains the original unCD56\textsuperscript{dim}/NKp46\textsuperscript{neg-low} phenotype (Figure R.8A, left). Opposite, only a minor fraction of FACS-sorted cCD56\textsuperscript{br}/NKp46\textsuperscript{pos} gives rise unCD56\textsuperscript{dim}/NKp46\textsuperscript{neg-low} NK cells over the considered days (Figure R.8A, right). The expression of CD94/NKG2A on derived subsets is similar to their NK cell subsets of origin. Derived cCD56\textsuperscript{br} NK cells from both unCD56\textsuperscript{dim} and cCD56\textsuperscript{br} maintain higher, but not significant, CD94/NKG2A expression compared to derived unCD56\textsuperscript{dim} NK cells (Figure R.8B).
These results point out that proliferating unCD56\textsuperscript{dim}/NKp46\textsuperscript{neg-low} and cCD56\textsuperscript{br}/NKp46\textsuperscript{pos} NK cells are plastic subpopulations, able to generate themselves and, in a certain degree, to give origin to cCD56\textsuperscript{br}/NKp46\textsuperscript{pos} and to unCD56\textsuperscript{dim}/NKp46\textsuperscript{neg-low}, respectively. They do not modify the levels of CD94/NKG2A, while they acquire NKp46 expression of the derived subsets upon IL-15+IL-18 stimulation.

These first set of experiments allowed us to uncovered a neglected NK cell subset highly expanded in the first weeks post-h-HSCT, namely unCD56\textsuperscript{dim} NK cells. Their early expansion and phenotypic features urged us to hypothesize that unCD56\textsuperscript{dim} NK cells could figure as an intermediated additional stage of development before cCD56\textsuperscript{br} NK cells or between cCD56\textsuperscript{br} and cCD56\textsuperscript{dim} NK cell subsets. Indeed, proliferating unCD56\textsuperscript{dim} NK cells are further characterized by a sort of plasticity, enabling these cells to generate themselves and cCD56\textsuperscript{br} as well. Moreover, we showed that the transient but very high expression of CD94/NKG2A on unCD56\textsuperscript{dim} NK cells is responsible for their reduced cytotoxicity soon after the transplant, contrary to their observed degranulation capability exerted in healthy subjects.

9.2 HCMV infection/reactivation and NK cells in h-HSCT patients: a flow cytometric point of view

As mentioned in the Introduction section, h-HSCT platforms showed remarkable positive clinical outcomes inducing immunologic tolerance between donor and recipient cells [50, 76]. However, to date, the immune-mediated GvL response is not fully exploited and several life-
threatening side effects, as the onset of GvHD and the occurrence of opportunistic viral infections, can still manifest hampering h-HSCT patient life. Of evidence, more than 50% of h-HSCT recipients show a HCMV infection/reactivation especially favoured by a prolonged immunodeficiency occurring after transplantation [124]. Therefore, our second aim was to uncover the HCMV influence on NK cell subset development and their effector functions.

9.2.1 HCMV influence on immune-reconstituting NK cell distribution after h-HSCT

To investigate the likely impact of HCMV infection/reactivation after h-HSCT, blood samples from 34 donor/recipient (D/R) pairs were collected at different time points in the first year after h-HSCT (Figure M.1, M.2A).

HCMV viral load in blood of Ds and Rs, before and every two weeks after h-HSCT, was assessed by Real-time PCR (CMV R-GENE, Argene) to identify those patients experiencing HCMV reactivation. We identified 14 patients that did not experienced HCMV reactivation and showed a viral load always lower than 500 IU/ml and 20 patients characterized by a viral load higher than 4000 IU/ml, at least once after the transplant, thus, considered as HCMV reactivated recipients. Further patient clinical information including their HCMV serological status (D-/R-, D-/R+, D+/R-, D+/R+) or the day of viral infection/reactivation occurrence are described in Material and methods section and are listed in Table 1.

PBMCs were isolated by density-gradient centrifugation and stained with a 21-color flow cytometry panel acquired at BD FACSymphony (BD Bioscience) in order to follow the dynamics of NK cell IR post-h-HSCT and to evaluate their receptor repertoire over the time (Figure M.1) (Table Abs). For this set of experiments NK cell gating strategy was refined, thus NK cells were identified as viable CD14<sup>neg</sup>/CD3<sup>neg</sup>/Lin<sup>neg</sup> (CD4, CD15, CD20, CD19, CD33, CD34, CD203c, FcεRI) lymphocytes (Figure M.2).

Although the flow cytometry panel has been modified and made more complicated in order to allow a deeper marker investigation, novel data are in line with our previous results [238] (Figure R.1). Indeed, soon after h-HSCT, patients either lack or have very low frequencies of circulating cCD56<sup>dim</sup> NK cells, whereas unCD56<sup>dim</sup> and cCD56<sup>br</sup> arise earlier, outnumbering the more differentiated NK cells. Subsequently, starting from the 2<sup>nd</sup> month after h-HSCT the
percent of unCD56^{dim} and cCD56^{br} NK cells progressively decreases, while the frequency of mature cCD56^{dim} NK cells increases, reaching a distribution similar to that of HCs after only 4 months from transplantation.

Then, investigating the immune-reconstituting NK cell distribution in the two cohorts of samples patients not experiencing (NR; n=14) and recipients undergoing viral infection/reactivation (R; n=20), we noticed that HCMV greatly affects NK cell subset IR, hence influencing the development and the maturation of NK cells (Figure R.9). In detail, the frequency of cCD56^{br} NK cells in R patients significantly decreases early after h-HSCT (at 1 and 3-4 months) with respect to NR recipients (Figure R.9, blue). Similarly, we observed that also unCD56^{br} NK cell subset, although little represented, shows a faster reduction in their frequencies early after h-HSCT in R patients compared to NR ones (Figure R.9, light blue).

As previously demonstrated the percentages of unCD56^{dim} NK cells in presence of HCMV infection/reactivation is not affected over the time post-h-HSCT (Figure R.9, red) [238]. The fast reduction of cCD56^{br} and unCD56^{br} NK cell frequencies is counterbalanced by the greater expansion of cCD56^{dim} NK cells and by a
significant increase of an unusual CD56\textsuperscript{neg}CD16\textsuperscript{pos} (CD56\textsuperscript{neg}) NK cell subset upon viral infection/reactivation (Figure R.9, black and green, respectively). Interestingly, CD56\textsuperscript{neg} NK cells arise soon after HCMV infection/reactivation (starting from the 3\textsuperscript{rd} month after h-HSCT) and their frequency remains stable till one year after transplantation. As expected, this subset is barely detectable in NR patients, strongly suggesting that viral infection/reactivation is able to influence NK cell IR and maturation after h-HSCT.

We further investigated the possible differences existing within patient cohorts. Among NR patients, the analysis of their HCMV-serostatus revealed that 7 out of 14 D/R couples are HCMV-seronegative (D-/R-). This means that both the donor and the relative recipient had never encountered the virus before. The D-/R- status is the rarest condition, considering the epidemiology of the infection in healthy individuals. The remaining 7 recipients are characterized by both HCMV-seronegative and HCMV-seropositive status, while they received HSCs from a HCMV-seropositive donor. Hence, they could potentially undergo HCMV reactivation, however the virus is still latent (R0) after h-HSCT (Table 1). Among D-/R- or R0 patients, we did not observed any significant difference in the NK cell IR dynamics, suggesting that primary HCMV exposure during life does not significantly impact on NK cell development after h-HSCT (not shown).

From clinical data on patients experiencing HCMV infection/reactivation, beside their HCMV-serostatus, we observed that 10 out of 18 recipients showed clinical viral reactivation and therefore have been subjected to anti-viral treatment with Ganciclovir or Valganciclovir (R1); the remaining 7 recipients underwent a subclinical reactivation and are able to clear the infection/reactivation without the need of anti-viral treatment (R2). Comparing NK cell distribution of these two subgroups, we noticed that at 1 month post-h-HSCT the frequency of cCD56\textsuperscript{br} NK cells derived from R2 recipients decreases faster than that of R1 patients. Moreover, cCD56\textsuperscript{dim} NK cell subset in R1 patients has a tendency to increase in its frequency compared to R2 recipients that, in turn, show significantly higher expansion of CD56\textsuperscript{neg} NK cell subset at 3-4 and 10-12 months after h-HSCT (not shown).

These results suggest that the development of hypo-responsive CD56\textsuperscript{neg} NK cells might be affected by anti-viral treatment, which, while controlling the HCMV viremia, could limit the need of expansion of mature NK cells efficiently prone to clear the viral infection/reactivation.
9.2.2 High dimensional single-cell analysis optimization

We observed that HCMV infection/reactivation is able to modify NK cell subset distribution by accelerating NK cell maturation after h-HSCT and driving the expansion of CD56<sup>neg</sup> NK cells. We, thus, wonder if HCMV could be responsible for further changes in NK cell phenotype and features. To gain more insight into NK cell diversity, high-dimensional flow cytometry data were concatenated and analysed with PhenoGraph algorithm. This unsupervised computational method evaluates, at single cell level, differences/similarities of marker expression automatically defining clusters of phenotypically identical cells [347].

In collaboration with Bioinformatic Unit of Clinical and Research Institute Humanitas, we performed PhenoGraph analysis running different values for the parameter “k” (sequentially set to 15, 30, 45, 60, 75, and 100) to evaluate which of them identified biologically relevant clusters and avoided overlapping phenotypical features of NK cells. For this purpose we generated a heatmap for each “k” value showing the integrated Median of Fluorescence Intensity (iMFI) of markers within all clusters (not shown). The “k45” was the most suitable for our analysis and it was further validated considering our initial biological knowledge on marker expression in different NK cell subsets. It enabled the identification of 28 phenotypically distinct clusters of identical NK cells (Figure M.4).

9.2.3 HCMV reactivation affects the NK cell subset distribution in the t-SNE map

PhenoGraph clustering (k45) obtained on all concatenated samples (2000 events/sample) was applied to each sample. In a first analysis, we investigated the NK cell subset distribution at different time points considering the HCMV infection/reactivation onset: 1-2 months (before viral reactivation), 3-4 months (soon after the pick of viremia), and 8-12 months (late after viral reactivation) post-h-HSCT. This analysis was carried out on 10 HCMV not reactivated patients and on 18 HCMV reactivated recipients, selected because all experienced the HCMV reactivation after 2 months from transplantation.
The localization of the NK cell subsets on the t-SNE plots shows that cCD56\textsuperscript{br}, unCD56\textsuperscript{br} and unCD56\textsuperscript{dim} NK cell subsets are mainly localized in the upper part of the t-SNE map (Figure R.10A), while cCD56\textsuperscript{dim} and CD56\textsuperscript{neg} NK cell subsets (black and green, respectively) are geographically distinct from the other subsets expanded early after h-HSCT, being mainly localized in the lower part of the plot (Figure R.10B). This observation suggests that more differentiated cCD56\textsuperscript{dim} and CD56\textsuperscript{neg} NK cells share similar features and that they are phenotypically different from the other immature NK cell subsets. Moreover, we confirmed that early reconstituting NK cell subsets (cCD56\textsuperscript{br}, unCD56\textsuperscript{br}, and unCD56\textsuperscript{dim}) arise prior to cCD56\textsuperscript{dim} in both cohorts of patients, while their frequencies decrease concomitantly to the increase of cCD56\textsuperscript{dim} NK cells at 3-4 months after h-HSCT.

The t-SNE plot analyses also confirmed that in R recipients, the reconstitution of cCD56\textsuperscript{dim} NK cell occurs earlier than in NR patients and it is coupled with the expansion of CD56\textsuperscript{neg} NK subset (Figure R.10, lower panels).

These results, in line with those obtained by manual gating, further validate the use of this high-dimensional single-cell analysis of NK cells for high-dimensional data.
9.2.4 NK cell cluster expansion and their correlation with HCMV

The automatic identification of clusters through Phenograph enabled us to deeply delineate NK cell features and to highlight which of those are associated with HCMV infection/reactivation, likely resembling the adaptive ml-NK cells.

To identify clusters preferentially enriched in NR patients and in the R ones, the mean of frequency of the 28 identified clusters was calculated for each samples at the three different time points previously described (1-2, 3-4, and 8-12 months after h-HSCT). As expected, at 1-2 months post-h-HSCT (before HCMV infection/reactivation) the frequency of clusters is similar between the two cohorts of patients. Indeed, HCMV infection/reactivation usually occurs from the 3\textsuperscript{rd} month after transplantation, thus changes in clusters and in NK cell phenotype are conceivable to develop in the following time points (Figure R11A). We only noticed an expansion of cluster 6 in not reactivated patients (19,3±12,8\% in NR vs 8,8±8,1\% in R), while cluster 14 is enriched in recipients that will experience viral reactivation (7, 1±5,6\% in NR vs 13,1±10,3\% in R).
The percentage of clusters at 3-4 months after h-HSCT (soon after viral reactivation) showed that five clusters are mainly represented in not reactivated patients: cluster 2 (7.5±5.6% in NR vs 2.1±3.6% in R), cluster 3 (3.9±3.4% in NR vs 1.3±2.0% in R), cluster 6 (11.4±7.7% in NR vs 6.1±6.2% in R), cluster 7 (9.7±8.8% in NR vs 2.6±2.6% in R), and cluster 8 (9.3±3.6% in NR vs 3.4±2.1% in R). On the contrary, other six clusters are preferentially expanded in recipients undergoing viral infection/reactivation: cluster 17 (3.7±2.7% in R vs 1.1±0.7% in NR), cluster 18 (5.7±3.7% in R vs 1.1±0.9% in NR), cluster 19 (8.6±5.1% in R vs 2.1±1.1% in NR), cluster 20 (7.8±4.8% in R vs 1.3±0.7% in NR), cluster 21 (5.4±4.1% in R vs 1.6±1.1% in NR), and cluster 27 (4.4±2.8% in R vs 1.4±1.2% in NR) (Figure R11B).

Finally, cluster distribution was assessed at 8-12 months after h-HSCT (late after HCMV infection/reactivation). This analysis allowed us to better investigate on persistent HCMV-driven changes in NK cell properties since, at almost one year from h-HSCT, NK cell IR is complete, HCMV infections are resolved, and persistent 

![Figure 11: NK cell cluster distribution on t-SNE maps after h-HSCT](image)

*Summary statistical graphs showing the frequency on total NK cells (mean ± SD) of NK cell clusters in NR patients (pts) (blue; n=10) and in R recipients (orange; n=24) at 1-2 (A), 3-4 (B), and 8-12 (C) months after h-HSCT. Multiple t-test; ** p <0.01; *** p <0.001. t-SNE plots showing the distribution of the clusters expanded in the absence of HCMV reactivation (D) (2, 3, 6, 7, and 8) and upon viral reactivation (E) (18, 19, 20, 21, and 27) in NR pts (upper panels) and in R recipients (lower panels) at different time points after h-HSCT. NK cell clusters are overlaid with the total NK cell distribution (gray background).*

infection/reactivation). This analysis allowed us to better investigate on persistent HCMV-driven changes in NK cell properties since, at almost one year from h-HSCT, NK cell IR is complete, HCMV infections are resolved, and persistent ml-NK cells should be highly expanded.

At this time point, our data showed that five clusters are preferentially enriched in not reactivated patients compared to reactivated recipients: cluster 2 (7.2±4.8% in NR vs 1.1±0.9% in R); cluster 3 (5.1±4.7% in NR vs 1.9±2.1% in R); cluster 7 (4.9±2.8% in NR vs 1.3±1.4% in R); cluster 8 (6.4±2.7% in NR vs 2.9±1.4% in R); and cluster 9 (12.73±6.07% in NR vs 8.75±4.59% in R) (Figure R11C, D).
Moreover, further five clusters resulted to be expanded in patients experiencing HCMV infection/reactivation with respect to not reactivated recipients: cluster 18 (8,1±4,6% in R vs 2,1±1,2% in NR); cluster 19 (9,5±5,7% in R vs 2,4±1,3% in NR); cluster 20 (6,6±5,3% in R vs 1,4±0,7% in NR); cluster 21 (4,7±3,6% in R vs 1,1±0,4% in NR); and cluster 27 (9,6±6,6% in R vs 3,4±0,5% in NR) (Figure R11C, E). These clusters are the same expanded also at 3-4 months post-h-HSCT (except for cluster 17), suggesting that modifications in NK cell features appear soon after HCMV infection/reactivation onset and they can persist over the time after transplantation.

9.2.5 Characterization of NK cell clusters expanded in not reactivated patients

We next investigated the phenotype of NK cells belonging to the clusters expanded in NR recipients. A heatmap showing the expression of each marker was created for the three time points considered: 1-2 (not shown), 3-4 (Figure R12A left) and 8-12 months (not shown). We focused our analysis at 3-4 months after h-HSCT when all the five clusters expanded in NR recipients (2, 3, 6, 7, and 8) reach the highest significant difference from R patients. Moreover, these clusters are enriched also in the other time points considered: cluster 6 mainly represented also at 1-2 months (before viral reactivation), while clusters 2, 3, 7, and 8 expanded also at 8-12 months (late after viral reactivation) after h-HSCT. We first observed that all NK cells belonging to these clusters are characterized by the expression of CD94/NKG2A regardless of the HCMV infection/reactivation status, thus suggesting that NK cells are not fully mature. Indeed, the majority of cells express also high levels of CD56 and NKp46, while low levels of CD57 and CD158b1b2j, markers of terminally differentiated cells. Moreover, they are comprised within the subsets of cCD56<sup>br</sup> and unCD56<sup>br</sup> NK cells. In contrast, as expected, we did not observed any expansion of CD56<sup>neg</sup> NK cells either in NR or in R patients (Figure R12A, right panel). The analysis of the expression of each single marker between NR and R patients in these clusters demonstrated that only NKp30 expression is significantly downregulated in R patients in clusters 2, 3, and 8 (Figure R12B).
9.2.6 Characterization of NK cell clusters expanded in patients experiencing HCMV infection/reactivation

We, then, evaluated the phenotype of NK cells belonging to the five clusters enriched in patients experiencing HCMV infection/reactivation. Similarly to the previous analysis, we generated a heatmap at three different time points (before/soon after/late after the viral infection/reactivation), and we focused more at 8-12 months after h-HSCT to find out persistent HCMV-driven changes in NK cell features (Figure R.13A, left panel) already present at 3-4 months post-transplantation. Differently from the phenotype of NK cells in NR patients, we noticed that all NK cells in clusters 18, 19, 20, 21, and 27 are defined by the expression of CD158b1b2j
(KIR2DL2/2DL3/2DS2) regardless of the HCMV infection/reactivation status, prompting us to hypothesize that all NK cells belonging to these clusters are terminally differentiated and fully licensed [165, 279]. This hypothesis was also supported by the observation that the majority of these NK cells are comprised within the subset of mature cCD56<sup>dim</sup> NK cells: cluster 18 (68.01%), cluster 19 (73.93%), cluster 20 (45.95%), cluster 21 (66.51%); and cluster 27 (67.67%) (Figure R.13A, right panel). Our analysis also revealed a differential expression of several other NK cell receptors between the two cohorts of patients. In particular, the surface levels of NKG2A, NKp30, and NKp46 are significantly reduced following HCMV reactivation, whilst NKG2C is more expressed in R patients compared to the NR ones. CD57 shows a kinetic of expression similar to that of NKG2C, although statistical significance is reached only at 10-12 months after h-HSCT (Figure R.13B and R.15).
Taken together these data demonstrated that HCMV infection/reactivation upon h-HSCT can not only shape NK cell subset distribution by inducing the expansion of CD56\textsuperscript{neg} NK cell subset, but also modulate NK cell receptor repertoire. We next performed the same cluster analysis comparing D-/R- and R0 patients in order to understand whether the HCMV-serostatus of both donors and recipients could affect the NK cell phenotype in patients that do not show any viral reactivation event following h-HSCT. Our results did not highlighted significant differences in cluster distribution and NK cell features (not shown), indicating that it is the HCMV reactivation in the recipients to drive the shaping in NK cell phenotype, rather than a previous viral infection. We did not observed differences in cluster characterization between R1 and R2 reactivated patients as well, suggesting that the anti-viral treatment following h-HSCT has not a significant impact on NK cells (not shown).

These observations further confirm our initial classification and distinction of samples in only the two cohorts of NR and R patients.

9.2.7 HCMV infection/reactivation after h-HSCT shapes NK cell phenotype

Since the aim of this study is to assess adaptive ml-NK cell features and their expansion in presence of HCMV infection/reactivation in h-HSCT patients, we next focused on a deeper characterization of those NK cells showing HCMV-correlated changes. In particular, to investigate the clinical relevance of our findings, we wondered if the differences in the expression levels of the markers observed in clusters enriched in R patients could eventually depend on the viral load. Therefore, we correlated the expression levels of NKp30, NKp46, NKG2A, NK2C, CD158b1b2j and CD57 (markers highlighted from the heatmap in Figure R.13) on total NK cells at 8-12 months with the highest peak of viremia reached during the reactivation time. The results obtained showed that the increase in viral load is accompanied by a proportional decrease in NKp30, NKp46, and NKG2A expression and by an increase in NKG2C, CD158b1b2j and CD57 frequencies (Figure R.14), thus suggesting that HCMV infection/reactivation is able to modulate the expression of these markers on NK cells.
We next evaluated the kinetic of those NK cell markers differentially expressed between the two cohorts of samples at different time points after h-HSCT in all the patients recruited. We noticed that the majority of markers analysed (NKp46, CD158b1b2j, NKG2A, and NKG2C) shows different expression levels starting from the 3rd month after h-HSCT, soon after HCMV infection/reactivation occurrence. In particular, NKG2A and NKp46 are significantly reduced in R patients, while CD158b1b2j and NKG2C show a significant increased expression in the same cohort of samples. The expression of NKp30 is always higher in NR patients, while an increase in CD57 marker expression was observed only at 8-12 months after h-HSCT following HCMV infection/reactivation (Figure R15).

Figure R.14: correlation analysis of NK cell marker expression and HCMV viremia
Graphs showing the correlation (black line) and the relative confidence interval at 95% (dashed line) between NK cell marker expression (NKp30, NKp46, NKG2A, NKG2C, CD158b1b2j, and CD57) and HCMV pick of viremia (IU/ml) in NR (○; n=8) and in R (●; n=12) patients at 8-12 months after h-HSCT.
Identification of a peculiar NK cell subset expanded upon HCMV infection/reactivation

Our phenotypic analysis showed that NK cells expanded after HCMV infection/reactivation in h-HSCT patients are defined by a high expression of CD158b1b2j and NKG2C and by a reduced expression of NKG2A and NKp30. Combining these results, we manually gated on NK cells characterized by a CD158b1b2j pos NKG2A neg NKG2C pos NKp30 low phenotype to assess whether the frequency of this subset could be related to the HCMV viremia, as for the single markers analysed. Our data demonstrated that the expansion of this peculiar subset directly correlates with the highest viral load reached during the reactivation (Figure R.16A).

Furthermore, we evaluated the kinetic of CD158b1b2j pos NKG2A neg NKG2C pos NKp30 low over the time after transplantation (not only at 8-12 months post-HSCT). We showed that this subset is expanded only in R patients and it appears soon after HCMV infection/reactivation, starting from the 3rd month after the h-HSCT. Moreover, it persists also at later time points till one year after transplantation. On the contrary, CD158b1b2j pos NKG2A neg NKG2C pos NKp30 low
NK cells are poorly represented in NR patients at all the time points considered (Figure R.16B), suggesting that this NK cell subset is specifically expanded in response to HCMV reactivation.

Looking more in deep to our cohort of patients experiencing viral reactivation, we noticed that 10 patients showed a single HCMV reactivation event, reaching a viremia higher than 4000 IU/ml only once over the time after transplantation, while 8 recipients experienced multiple viral reactivations. Thus, to better investigate the possible correlation between CD158b1b2\textsuperscript{pos}NKG2A\textsuperscript{neg}NKG2\textsuperscript{pos}NKp30\textsuperscript{low} NK cell expansion and HCMV infection/activation, we assessed the expansion of CD158b1b2\textsuperscript{pos}NKG2A\textsuperscript{neg}NKG2\textsuperscript{pos}NKp30\textsuperscript{low} NK cell subset within the cohort of R patients reactivations. Thus, to better investigate the possible correlation between CD158b1b2\textsuperscript{pos}NKG2A\textsuperscript{neg}NKG2\textsuperscript{pos}NKp30\textsuperscript{low} NK cell expansion and HCMV infection/activation, we assessed the expansion of CD158b1b2\textsuperscript{pos}NKG2A\textsuperscript{neg}NKG2\textsuperscript{pos}NKp30\textsuperscript{low} NK cell subset within the cohort of R patients

![Graph showing the correlation between CD158b1b2\textsuperscript{pos}NKG2A\textsuperscript{neg}NKG2\textsuperscript{pos}CNKp30\textsuperscript{low} NK subset frequency (% on total NK cells) and HCMV viremia (IU/ml) in NR patients (\(
\square\); n=8), and in recipients experiencing a single (\(
\square\); n=5) or multiple (\(
\blacksquare\); n=6) reactivation events at 8-12 months after h-HSCT. (B) Graph showing the frequency (% mean ± SEM) of CD158b1b2\textsuperscript{pos}NKG2A\textsuperscript{neg}NKG2\textsuperscript{pos}CNKp30\textsuperscript{low} NK cells on total NK cells at different time points after the h-HSCT. Multiple t-test; * p <0.05 ** p <0.01; *** p <0.001. (C) Graph showing the frequency (% mean ± SEM) of CD158b1b2\textsuperscript{pos}NKG2A\textsuperscript{neg}NKG2\textsuperscript{pos}CNKp30\textsuperscript{low} NK cells on total NK cells in R1 experiencing a single (\(
\square\); n=10) or multiple reactivations (\(
\blacksquare\); n=8) and in subclinical reactivated recipients (R2) experiencing a single event of HCMV reactivation (\(
\square\); n=8). Multiple t-test; * p <0.05 ** p <0.01; *** p <0.001 \(
\square\) vs \(
\blacksquare\); * p <0.05 ** p <0.01; *** p <0.001 \(
\square\) vs \(
\blacksquare\).
distinguishing between recipients experiencing single and multiple HCMV reactivation events. Interestingly, we observed that patients experiencing multiple HCMV reactivations exhibit a significantly higher expansion of our NK cell subset compared to single reactivated patients (not shown).

Notably, patients experiencing multiple HCMV reactivations are all R1 treated with anti-viral drugs, while single reactivated patients comprise both R1 treated and R2 not treated recipients. Our newly defined subset shows higher expansion in R1 patients experiencing multiple viral reactivations compared to both R1 and R2 patients experiencing only a single event of viral reactivation (Figure R.16C).

These results prompted us to speculate that CD158b1b2j^{pos}NKG2A^{neg}NKG2C^{pos}NKp3^{low} NK cells could proliferate and expand every time HCMV reactivation occurs, “remembering” the virus challenge regardless of the anti-viral treatment they received. To better confirm this hypothesis we selected patients experiencing multiple viral reactivations over the time after h-HSCT and we evaluated the expression of those markers associated with ml-NK cells (CD158b1b2j, NKG2A, NKG2C, and NKp30) at every HCMV reactivation event (Figure R.17A-D). We demonstrated that the expression of CD158b1b2j and NKG2C is significantly increased upon the 2^{nd} and 3^{th} event of HCMV reactivation compared to the 1^{st} one (Figure R.17A, C). Contrariwise, NKG2A levels are reduced after a 2^{nd} and 3^{th} virus challenge with respect to the 1^{st} HCMV infection/reactivation; while no significant difference was observed in NKp30 marker expression (Figure R.17B, D). Moreover, we evaluated the frequency of our newly disclosed NK cell subset in the same patients at the same time points and we showed that CD158b1b2j^{pos}NKG2A^{neg}NKG2C^{pos}NKp3^{low} NK cells expand more upon subsequent viral reactivations (Figure R.17E).
Taken together, our data indicate that multiple viral challenges can highly affect NK cell phenotype driving a higher expansion of ml-NK cells every time a peak of viremia (>4000 IU/ml) is reached.

Merging manual gating strategy and computational analyses, our results showed that only in R patients, a peculiar subset of NK cells, characterized by CD158b1b2jNKG2A<K>posNKG2C<pos>NKp30<low> phenotype, is expanded and it persists till one year post-h-HSCT defining a subpopulation of long-lived cells. Moreover, it appears soon after the occurrence of viral reactivation and its expansion is more pronounced in presence of multiple reactivation events, suggesting that CD158b1b2jposNKG2A<pos>NKG2C<pos>NKp30<low> NK cell subset could recognize and remember the viral challenge upon transplantation, reminding features of ml-NK cells.

9.3 HCMV infection/reactivation and NK cells in h-HSCT patients: molecular signature
Novel insights demonstrated that epigenetic reprogramming might have an important role in defining ml-NK cells properties both in HCMV-seropositive healthy individuals and in recipients experiencing HCMV infection/reactivation after allogeneic HSCT [310, 312]. Thus, in order to better clarify if our newly disclosed subset consists of real ml-NK cells and to investigate the molecular mechanisms governing ml-NK cells and their specific fingerprint, we assessed RNA expression profiles by NGS.

9.3.1 FACS-sorted NK cell subsets from h-HSCT patients

Firstly, we FACS-sorted NK cells from recipients undergone h-HSCT and experiencing or not HCMV infection/reactivation (Table 3), selecting, if possible, the same samples used for previous experiments. Samples were also chosen based on: i) the occurrence of HCMV reactivation, but also their HCMV status (D-/R-, D-/R+, D+/R-, D+/R+), ii) their type of viral reactivation and anti-viral treatment (D-/R-, R0, R1, R2), and iii) the number of reactivation.

<table>
<thead>
<tr>
<th>h-HSCT pts</th>
<th>HCMV status (D/R)</th>
<th>HCMV Reactivation (R)</th>
<th>Single (S)/Multiple (M) R</th>
<th>Months post-h-HSCT</th>
<th>Sorted NK subset</th>
</tr>
</thead>
<tbody>
<tr>
<td>#3</td>
<td>+/-</td>
<td>na</td>
<td>na</td>
<td>9</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$</td>
</tr>
<tr>
<td>#50</td>
<td>+/-</td>
<td>na</td>
<td>na</td>
<td>9</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$</td>
</tr>
<tr>
<td>#89</td>
<td>+/-</td>
<td>na</td>
<td>na</td>
<td>9</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$</td>
</tr>
<tr>
<td>#105</td>
<td>+/-</td>
<td>na</td>
<td>na</td>
<td>11</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$</td>
</tr>
<tr>
<td>#18</td>
<td>+/-</td>
<td>R0</td>
<td>na</td>
<td>12</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$</td>
</tr>
<tr>
<td>#48</td>
<td>+/-</td>
<td>R0</td>
<td>na</td>
<td>9</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$/KIR$\text{C}^0$</td>
</tr>
<tr>
<td>#77</td>
<td>+/-</td>
<td>R0</td>
<td>na</td>
<td>12</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$/KIR$\text{C}^0$</td>
</tr>
<tr>
<td>#88</td>
<td>+/-</td>
<td>R0</td>
<td>na</td>
<td>9</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$/KIR$\text{C}^0$</td>
</tr>
<tr>
<td>#27</td>
<td>+/-</td>
<td>R1</td>
<td>M</td>
<td>10</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$/KIR$\text{C}^0$</td>
</tr>
<tr>
<td>#63</td>
<td>+/-</td>
<td>R1</td>
<td>S</td>
<td>9</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$/KIR$\text{C}^0$</td>
</tr>
<tr>
<td>#25</td>
<td>+/-</td>
<td>R1</td>
<td>M</td>
<td>10</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$/KIR$\text{C}^0$</td>
</tr>
<tr>
<td>#65</td>
<td>+/-</td>
<td>R1</td>
<td>M</td>
<td>9</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$/KIR$\text{C}^0$</td>
</tr>
<tr>
<td>#13</td>
<td>+/-</td>
<td>R2</td>
<td>S</td>
<td>11</td>
<td>KIR$\text{A}^0$/KIR$\text{B}^0$/KIR$\text{C}^0$</td>
</tr>
<tr>
<td>#26</td>
<td>+/-</td>
<td>R2</td>
<td>S</td>
<td>11</td>
<td>KIR$\text{B}^0$</td>
</tr>
<tr>
<td>#42</td>
<td>+/-</td>
<td>R2</td>
<td>S</td>
<td>7</td>
<td>KIR$\text{B}^0$</td>
</tr>
<tr>
<td>#76</td>
<td>+/-</td>
<td>R2</td>
<td>M</td>
<td>7</td>
<td>KIR$\text{B}^0$/KIR$\text{C}^0$</td>
</tr>
</tbody>
</table>

Table 3: FACS-sorted samples
Table showing the patient (h-HSCT pts) information (HCMV status, HCMV reactivation events, time post-h-HSCT) and the sorted NK cell subsets. Pts: patients; KIR: CD158b1b2j; 2A: NKG2A; 2C: NKG2C.
events single (S) or multiple (M). Based on our previous results, we selected PBMC samples at 7-12 months after h-HSCT because our newly identified NK cell subset is still present and we have higher disposability of cells compared to earlier time points (Table 3).

We, then, FACS-sorted 4 main populations of NK cells based on different expression of CD158b1b2j and other markers:

- **CD158b1b2j^{pos} NK cells** expanded in R patients, but having a sufficiently high percentage of cells in NR too (Figure R.18);  

- **CD158b1b2j^{neg} NKG2A^{-}\text{neg} NKG2C^{pos} NK cells** carrying the phenotype disclosed in R patients, except for the negative expression of Nkp30. We excluded this latter marker otherwise the frequency of our novel NK cell subset would had been too low in NR patients to allow a following RNA-seq analysis (Figure R.18A);  

- **CD158b1b2j^{neg} NK cells** expanded mainly in NR patients (Figure R.18B);  

- **CD158b1b2j^{neg} NKG2A^{pos} NKG2C^{neg} NK cells** characterized by an opposite phenotype compared to that expressed by putative ml-NK cells and expanded mainy in NR patients compared to R ones (Figure R.18C).
9.3.2 Gene expression pattern between CD158b1b2j\(^{neg}\) and CD158b1b2j\(^{pos}\) NK cells of h-HSCT patients experiencing HCMV infection/reactivation

We first wanted to assess if CD158b1b2j\(^{pos}\) NK cells were the ones characterized by ml- properties compared to CD158b1b2j\(^{neg}\) NK cells in the same R h-HSCT patients. RNA-seq analysis revealed that CD158b1b2j\(^{pos}\) and CD158b1b2j\(^{neg}\) NK cells cluster separately, thus suggesting a different expression profile. Indeed, 1109 genes are significantly deregulated (p-value adjusted <0.05) in the two groups of comparison (Figure R.19).

To examine the molecular features of our putative ml-NK cells, we used GSEA by pairwise comparison between the two NK subpopulations and genes of the signalling pathways regulating NK cell functions were retrieved from KEGG database. Our results underlined that CD158b1b2j\(^{pos}\) NK cells show an enrichment in some pathways associated to cell activation including *NK cell mediated cytotoxicity* and *Oxidative Phosphorilation* (Figure R.20A). We, therefore, deeply analysed NK cell effector-functions through REACTOME database that confirmed an enrichment in pathways correlated to mitochondrial fitness and respiration, thus metabolic activation (Figure R.20B). Based on the fact that ml-NK cells are characterized by higher cytotoxic responses upon re-stimulation, we further probed genes belonging to *NK cell mediated cytotoxicity* pathway and we noticed a downregulation of *KLRC1* (NKG2A), *FCER1G*
(FcεRγ), SYK (SYK) and TYROBP (DAP-12) in CD158b1b2jpos NK cells; while these cells show a significant increased expression of KLRC2 (NKG2C), KLRC3 (NKG2E), and PRF1 (Perforin). Moreover, besides CD158b1b2j, other KIRs are upregulated indicating a more mature phenotype. These features are in line with current knowledge about hyper-responsive ml-NK cells defined by epigenetic reprogramming and our previous flow cytometric analysis [310, 312] (figure R.20C).

The deep focus on **Oxidative Phosphorilation** pathway highlighted that different deregulated genes are involved in the alteration of mitochondrial membrane potential (TCIRG1, ATP6V0E1, and UCP3) and regulation of mitochondrial fatty acid oxidation (PPARD and EFTB). Furthermore, SDHA, NDUFA5, and NDUFAF5, upregulated in CD158b1b2jpos NK cells, contribute to the transfer of electrons, thus providing additional evidence that CD158b1b2jpos NK cells are more activated than CD158b1b2jneg subpopulation (Figure R.20D).

**Figure R.20:** Gene expression profile of CD158b1b2jneg vs CD158b1b2jpos in R h-HSCT patients
(A) Bar plot representing the top highly deregulated KEGG pathways (|NES| ≥1.8) on CD158b1b2jneg and CD158b1b2jpos NK cells isolated from R patients at 7-12 months after h-HSCT (n=7). (B) Bar plot representing the top highly deregulated REACTOME pathways (|NES| ≥1.8) on CD158b1b2jneg and CD158b1b2jpos NK cells isolated from R patients at 7-12 months after h-HSCT (n=7). (C) Heatmap identification of the main deregulated genes belonging to NK cell mediated cytotoxicity pathway gene set in CD158b1b2jneg and CD158b1b2jpos samples of R h-HSCT patients (D) Heatmap identification of the main deregulated genes belonging to Oxidative phosphorilation gene set in CD158b1b2jneg and CD158b1b2jpos samples of R h-HSCT patients.
9.3.3 D-/R- vs R0: primary HCMV exposure during life does not impact on NK cell features after h-HSCT

The differences highlighted in the previous comparison, allowed us to mainly focus our next analyses on educated CD158b1b2jpos NK cells.

Figure R.21: comparison between sorted CD158b1b2jpos NK cells in D-/R- and R0 patients.
(A) Graph showing the frequency (%, mean ± SEM) of CD158b1b2jpos in D-/R-(red dots; n=4) and R0 (blue dots; n=4). (B) Dendrogram showing sample clustering obtained from counts per million (CPM) data. For each patient (#n) the kind of HCMV-serostatus (D-/R- or R0) is defined. (C) Multidimensional scaling plot of the samples. (D) MA-plot(s) of the comparison; red dots represent significantly differentially expressed features; triangles correspond to features having a too low/high log2(FC).

Figure R.22: differentially expressed genes in sorted CD158b1b2jpos NK cells from D-/R- and R0 patients.
(A) Heatmap showing the differential expression of genes between donor/recipient couples that never encountered HCMV (D-/R-) (red line; n=4) and recipient that do not experience viral reactivation (R0) (light blue line; n=4). For each patient (#n) the kind of HCMV-serostatus (D-/R- or R0) is defined. (B) Table listing the differentially upregulated (red) genes in R0 compared to D-/R-.

Before looking at the differential expression of RNA transcripts between R and NR, we wondered if some differences within D-/R- and R0 patients that do not undergone HCMV
reactivation exist. Between these two cohorts we did not observe any significant changes through flow cytometry experiments in terms of NK cell subset IR (Figure R.9) and phenotype, in particular in CD158b\textsubscript{1b2j} marker expression (Figure R.21A). As expected, the variability exploration between D-/R- and R0 on FACS-sorted CD158b\textsubscript{1b2j}\textsuperscript{pos} NK cells showed that all the samples are clustered together (Figure R.21 B-D). Indeed, RNA-seq analysis showed that only 10 genes are differentially expressed, and in particular upregulated, in R0 compared to D-/R- (p-value adjusted <0.05) (Figure 22).

9.3.4 R1 vs R2: anti-viral treatment can influence the expansion of CD56\textsuperscript{neg} NK cells

Once confirmed that only few genes are deregulated within the group of NR patients, we moved to the cohort of recipients experiencing HCMV reactivation in order to investigate the role of anti-viral treatment in ml-NK cell development. Based on our FACS evidence, in R2 patients experiencing subclinical viral reactivation, the CD56\textsuperscript{neg} NK cell expansion is more pronounced at 3-4 and 10-12 months after h-HSCT compared to R1 clinical reactivated patients (Figure R.23), while any phenotypic variations was observed in NK cells derived from R1 and R2 samples in terms of CD158b\textsubscript{1b2j} expression (Figure R.23A).

![Figure R.23: comparison between sorted CD158b\textsubscript{1b2j}\textsuperscript{pos} NK cells in R1 and R2 patients.](image)

(A) Graph showing the frequency (\%, mean ± SEM) of CD158b\textsubscript{1b2j}\textsuperscript{pos} in R1 (blue rectangles; n=4) and R2 (red rectangles; n=4). (B) Dendrogram showing sample clustering obtained from counts per million (CPM) data. For each patient (#n) the kind of HCMV reactivation/anti-viral treatment (R1 or R2). (C) Multidimensional scaling plot of the samples. (D) MA-plot(s) of the comparison; red dots represent significantly differentially expressed features; triangles correspond to features having a too low/high log\textsubscript{2}FC. (E) Table listing the differentially upregulated (red) and downregulated (blue) genes in R2 vs R1 patients.
The variability exploration between R1 and R2 on FACS-sorted CD158b1b2j<sup>pos</sup> NK cells showed that all the samples are clustered together with no significant changes in their features (Figure R.23B-D). In fact, RNA-seq analysis demonstrated that only 5 genes are differentially deregulated between R1 and R2 (p-value adjusted <0.05) (Figure R.23E). Specifically, *NCAM1*, *SERAC1*, *HBA2*, *ATP8B4* are downregulated, while *PI4KAP1* (pseudogene) is upregulated in R2 subclinical reactivated patients with respect to R1 clinical reactivated recipients (Figure R.23E). Of note, the lower expression of *NCAM1*, encoding for CD56, is in line with flow cytometry data showing a higher expansion of CD56<sup>neg</sup> NK cells in R2 patients.

These results may suggest that anti-viral drugs, actively reducing HCMV viremia in R1 recipients, could limit the expansion of CD56<sup>neg</sup> NK cell subset, which in turn is increased in untreated R2 patients. Concomitantly, it is likely that the CD56<sup>neg</sup> subset, since it is still present at one year after h-HSCT when the viral infection has been cleared, represents a subpopulation of NK cells that has already exerted its cytotoxicity against virus challenge.

Moreover, these first findings on RNA-seq data confirm the results obtained with the flow cytometry platform, indicating that only few changes occur between D-/R- and R0 and between R1 and R2 patients, allowing us to proceed in the analysis of total NR (n=8) versus R patients (n=7), comparison that is strengthened by the number of samples considered.

**9.3.5 Gene expression pattern between not reactivated patients and recipients experiencing HCMV infection/activation**

The RNA-seq analysis revealed that CD158b1b2j<sup>pos</sup> NK cells isolated from NR recipients and R patients cluster separately, thus suggesting that these two groups of patients have a different expression profile (Figure R.24). Indeed, 946 genes are significantly deregulated (p-value adjusted <0.05) in the two groups of comparison. In details, 390 genes are upregulated, while 556 genes are significantly downregulated in R patients with respect to NR recipients (Figure R.24). Of note, h-HSCT R2 patient #26 experiencing subclinical reactivation, is the only recipient characterized by a gene expression intermediated between the two cohorts of samples (R.24E). This results can be explained by the lower HCMV viremia (3000 IU/ml) that characterizes this patient.
9.3.5.1 GSEA

To examine the molecular features of our putative ml-NK cells, we used GSEA by pairwise comparison between not reactivated and HCMV reactivated samples (NR/R). Genes of the signalling pathways regulating NK cell functions were retrieved from KEGG database (Figure R.25A). CD158b1b2pos NK cells derived from R patients show an enrichment in some pathways associated to cell activation including:

**Figure R.24: differentially expressed genes in sorted CD158b1b2pos NK cells NR and R patients**

(A) Dendrogram showing sample clustering obtained from counts per million (CPM) data. For each patient (#n) the kind of HCMV-serostatus, infection/reactivation and anti-viral treatment is defined. R1: clinical reactivation; R2: subclinical reactivation; R0: not reactivation; D-R: donor/recipient that never encountered HCMV. (B) Principal component analysis (PCA) showing the variability of samples in the two groups of comparison, pts: patients. (C) MA-plot(s) of the comparison between NR patients and R recipients representing the log ratio of differential expression as a function of the mean intensity for each feature. (D) Vulcano plot representing the log of the adjusted P value as a function of the log2 ratio of differential expression (log2FC). Red dots represent significantly differentially expressed features; triangles correspond to features having a too low/high log2FC. (E) Heatmap showing the deregulated genes between NR (n=8; blue) and R (n=7; orange) patients. For each patient (#n) the kind of HCMV-serostatus, infection/reactivation and anti-viral treatment is defined.
Figure R.25: Gene expression patterns in sorted CD158b1b2j pos NK cells from pts experiencing or not HCMV reactivation.

(A) Bar plot representing the top highly deregulated KEGG pathways (|NES| ≥1.8) on CD158b1b2j pos NK cells isolated from patients at 7-12 months after h-HSCT experiencing (n=7; orange) or not (n=8; blue) HCMV reactivation. In bold are the pathways better described in the Results section. (B) GSEA was performed to determine the specific enrichment in gene signatures (GeneSet) between the same two cohorts of patients: NK cell mediated cytotoxicity pathway gene set enriched in NR patients is depicted. The top significantly deregulated genes (p-value adjusted <0.05) in the two cohorts of samples are listed: KLRC1 (NKG2A), KIR2DL4 (KIR2DL4), KIR2DS4 (KIR2DS4), FCER1G (FceRγ), SYK (SYK), IFNGR1 (IFN-γ receptor1), GZMB (granzyme-β), IFNAR2 (IFN-α β receptor2), KIR2DL1 (KIR2DL1), NCR3 (NKp30), PRKCA (PKCA), PRF1 (perforin), TYROBP (DAP-12), PTK2B (FAK2), SH2D1B (EAT-2), KRAS (K-Ras), LCK (LCK), IFNG (IFN-γ), PPP3CC (calcineurin A-γ), FAS (FAS). (C) Heatmap showing the differential deregulated genes in NK cell mediated cytotoxicity pathway gene set. FDR: false discovery rate; NES: normalized enrichment score.

- the Janus kinase signal transducer and activator of transcription (JAK/STAT) pathway
plays critical roles in governing immune system, i.e. cellular responses to cytokines and growth factors, proliferation, differentiation, migration, and apoptosis [361];

- the RIG-I-like receptor signaling and Antigen processing and presentation pathways are pattern of recognition receptors responsible for the detection of viral pathogens and for the generation of both innate and adaptive immune responses.

Of note, JAK/STAT pathway and RIG-I-like receptor signalling are involved in downstream IFN-related responses, underlined the HCMV impact in R patients.

Contrariwise, many pathways enriched in NR recipients are correlated to DNA metabolism, indicating an activated state of NK cells in those recipients. The enrichment of NK cell mediated cytotoxicity pathway in NR patients caught our attention (Figure R.25A). Since ml-NK have been described as cells endowed with adaptive features and higher effector-functions upon a secondary exposure to the same antigen [319, 322], the disclosure of this last result pushed us to deeply investigate on the differential expression of those genes included in NK cell mediated cytotoxicity pathway (Figure R.25B-C). Firstly, we focused on the differentially expressed genes comprised in the leading edge subset of this pathway showing a p-value <0.05. KLRC1, encoding for NKG2A, is the top highly downregulated gene in R patients, confirming the significant reduced levels of this iNKR observed by previous flow cytometry experiments (Figure R.15). Other significantly upregulated genes in NR patients are KIR2DS4 and KIR2DL4, encoding for activating and inhibitory receptors, respectively, meaning that KIR molecules (besides CD158b1b2j) can be involved in the viral infection/reactivation control.

According to recent works showing that adaptive NK cells undergo a reprogramming of adaptor molecules [310, 312], we observed that FCER1G (FceRγ), SYK (SYK), and SH2D1B (EAT-2) are downregulated upon viral infection (Figure R.25B-C). The downregulation of these transcription factors is also confirmed by semi-quantitative real-time PCR showing a strong
decrease in PLZF gene expression, likely leading the deregulation of the other above-mentioned genes (Figure R.26).

Moreover, we noticed that RNA expression of genes encoding for cytotoxic granules, granzyme-β and perforin, is reduced in R patients. Finally, in line with the expansion of newly disclosed NK cell subset defined by a CD158b1b2jposNKG2Aneg NKG2CposNKp30low phenotype, we also observed a downregulation of NCR3 encoding for NKp30 receptor. Interestingly, IFNG gene results upregulated, while its receptor IFNGR1 is downregulated in R patients, suggesting that modifications in IFN locus likely occur subsequently to viral infection/reactivation [315] (Figure R.25B-C).

We further analysed NK cell effector-functions in patients undergoing HCMV infection/reactivation through REACTOME database (Figure R.27A). Different signaling pathways related to interferon production and its downstream responses result to be altered in R patients, i.e. IFN-α/β signaling, RIG-I MDA5 mediated induction of IFN-α/β pathways, Interferon signaling, cytokine signaling in immune system, and finally IFN-γ signalling (Figure R.27B).

We deeply focused on the differentially expressed genes comprised in the leading edge subset of this IFN-γ signaling pathway showing a p-value <0.05 (Figure R.27C). Beside different HLA-I genes, GBP1, IFNG, OASL, IRF4, STAT1 genes show enhanced expression upon viral reactivation. GBP1 codes for guanylate binding protein 1, whose expression is induced by IFN, particularly in patients suffering of chronic active Epstein-Barr (EBV) virus infection [362].
Indeed 3 out of 7 recipients experiencing HCMV reactivation, show also EBV infection after h-
HSCT. *IRF4* (interferon regulatory factor 4) and *STAT1* (signal transducer and activator of transcription 1) are involved in the regulation of interferons upon viral infection and in mediating IFN-inducible genes, thus mediating cellular responses to IFNs (Figure R.27C). As previously, these data confirm the hypothesis that HCMV induces changes in pathways related to IFNs. This phenomenon is further corroborated by HALLMARK analysis showing a strong deregulation of genes in response to cytokines, in details IFN-γ, IFN-α, and TNF-α (Figure R.27D).

Finally, both through KEGG and REACTOME analysis, we noticed that some T cell specific pathways are peculiar of NK cells in R patients (i.e. *Antigen processing and presentation*, and *Downstream TCR Signaling*) suggesting shared features with activated T lymphocytes and reinforcing the likely presence of innate cells endowed with adaptive phenotype (Figures R.25 and R.27).

### 9.3.5.2 IPA

The analysis using IPA software was used to interpret data in the context of top upstream regulators and canonical pathways associated with the differentially expressed genes for the comparison between patients experiencing HCMV reactivation and not reactivated ones (R/NR). Filtered dataset (|FC| ≥1.5 and a p-value adjusted ≤0.05) identified 135 genes downregulated and 76 upregulated in R patients. In particular, this analysis revealed that 14 canonical pathways eligible genes are significantly up-regulated, and 10 pathways eligible genes are down-regulated. Three canonical pathways were evaluated in detail: *Th-1, iCOS-iCOSL signaling in T helper cells*, and *PD-1/PD-L1 cancer immunotherapy pathways* (Figure R.28A).

Canonical pathways related to T helper cell function were selected in view of recent published works highlighting similarities between ml-NK cells and Th cell signature, in particular in respect to IFN locus epigenetic modification [315].

In *Th-1 pathway* (z-score = 1.897) *CD3D, CD3E, CD3G, HLA-DOA, HLA-DQA1, HLA-DQB1, HLA-DRA, HLA-DRB5* genes are upregulated while *IL12RB2, KLRC1*, and *DLIL1* genes are downregulated in accordance to IPA expected results based on recent literature update. Of note, IL-12 receptor (*IL12RB2*) is significantly downregulated in R patients, suggesting a reduced cytotoxicity upon stimulation with this cytokine, as reported [208]. Concomitantly, we observed also a reduction of *IL18R1* and *IL18RAP* gene expression in viral reactivated patients (log₂FC = -1.14 and -2.372, respectively).
iCOS-iCOSL signaling in T helper cells pathway (z-score = 2,449) shares some genes with the Th-1 pathway: CD3D, CD3E, CD3G, HLA-DOA, HLA-DQA1, HLA-DQB1, HLA-DRA, HLA-DRB5, and also CAMK4 and FCER1G genes. Excepted for FCER1G, all genes are upregulated upon viral reactivation.

Finally, we focused on PD-1/PD-L1 cancer immunotherapy pathways (z-score = -1,633), suggesting an inhibitory function in these cells. Genes comprised in this canonical pathway are upregulated in R patients: CD274 (PD-L1), HLA-DOA, HLA-DQA1, HLA-DQB1, HLA-DRA, HLA-DRB5, IL12A, and PDCD1 (PD-1) (Figure R.28B). The discovery of PD-1/PD-L1 upregulation in patients experiencing viral reactivation, urged us to deeply investigate on other known inhibitory immune checkpoints. We noticed that also lymphocyte activation gene 3
LAG3, CD223) gene expression is enhanced (log\(_2\)FC= 3.48) following HCMV reactivation in patients, while TIGIT and TIM3, usually associated to exhausted status, are downregulated. Moreover, the increased expression of HLA-DOA, HLA-DQA1, HLA-DQB1, HLA-DRB1, and HLA-DRB5 in R, common to the three above-mentioned pathways, is correlated to the enhanced increased expression of their upstream regulator CIITA (Class II Major Histocompatibility Complex Transactivator) (log\(_2\)FC = 0.78), even higher in multiple reactivated recipients than in patients experiencing a single event of HCMV reactivation (not shown).

9.3.6 PD-1 and LAG3 as potential immune-checkpoints to rouse ml-NK cells

The main results obtained through systems biology approach underlined that expanded CD158b1b2j+ NK cells in R HCMV reactivated patients could be impaired in their effector-functions, thus not able to solve the viral challenge occurring after h-HSCT. In particular, we observed a reduced expression of genes correlated to the pathway of NK cell mediated cytotoxicity (Figure R.25A) concomitantly with an increased expression of checkpoint inhibitors (PD-1 and LAG3) (Figure R.28). Notably, both PD-1 and LAG3 resulted to be upregulated primarily in T cells upon chronic stimulation and their blockade enables the restoration of T cell cytotoxicity [338]. Moreover, recent work showed that chronic stimulation through NKG2C in combination with IL-15 induces also adaptive NK cells proliferation and enhances LAG3 and PD-1 expression [363].

For these reasons we performed masking experiments by blocking either PD-1 or LAG3 or both in the presence or in the absence of IFN-\(\gamma\) stimulated HUVEC expressing their putative ligands (PD-L1 and HLA-II respectively) (FigureR.29A) PBMCs from R (n = 3) and NR (n = 3) recipients at 8-12 months post-h-HSCT were stimulated over night with rhIL-2 and rhIL-12 (100 U/mL each) to induce the expression of PD-1 and LAG-3. The blockade of either PD-1 or LAG-3 resulted in an increased cytotoxicity, assessed through IFN-\(\gamma\) production, in R recipients compared to the co-culture with target cells alone. However, R h-HSCT patients, although retaining a higher IFNG expression, still show reduced IFN-\(\gamma\) release than NR recipients (Figure R.29B), indicating that the upregulation of PD-1 and LAG-3 is only partially involved in NK cell dysfunctional status following persistent viral infection.
**Figure R.29: PD-1 and LAG3 checkpoint inhibitors.**

(A) Representative flow cytometry contour plots showing the surface expression of PD-L1 and HLA-II on HUVEC stimulated (100 ng/mL) or not (unstimulated) three days with IFN-γ. Fluorescence minus one (FMO) was used to set the gate on positive population. (B) Summary statistical graph showing the IFN-γ release on total NK cells from NR (○; n=3) and R (■; n=3) recipients at 8-12 months post-h-HSCT co-cultured with HUVEC either in the absence (-) or in the presence (+) of blocking antibodies (α-LAG3 and α-PD-1). Unpaired t-test: * p<0.05 NR vs R recipients or Paired t-test; # p<0.05 vs NK cells cultured with HUVEC alone in R recipients.

- **Legend:**
  - HUVEC
  - αLAG-3
  - αPD-1
  - ⋆
  - 0.052
  - #
10 DISCUSSION

10.1 The early expansion of anergic unCD56\textsuperscript{dim}/NKG2A\textsuperscript{pos} NK cells represents a therapeutic target in h-HSCT

The development of h-HSCT platforms using RIC and PT-Cy has represented a great step forward for the treatment of patients affected by high-risk hematologic malignancies [42, 364]. Nevertheless the GvL effect, primarily induced by HLA mismatch between donor and recipient cells, has not been fully clarified and efficiently exploited. In this context, it has been demonstrated that faster and qualitatively high is the IR better is patient clinical outcome [40, 64, 78-83].

In this work, we demonstrated that NK cells are the first lymphocytes to recover after h-HSCT, prior to adaptive B and T cells, hence suggesting that NK cells could play a major role in the IR and engraftment. In h-HSCT patients we detected four different NK cell subsets (cCD56\textsuperscript{br}, cCD56\textsuperscript{dim}, unCD56\textsuperscript{br}, unCD56\textsuperscript{dim}), whose distribution is significantly different from healthy counterparts. Interestingly, cCD56\textsuperscript{dim} subset, representing the 90% of circulating NK cells under physiologic conditions, is infrequent till 8 weeks post-h-HSCT when its percentage return similar to HCs. The low number of cCD56\textsuperscript{dim} is counterbalanced by the expansion of cCD56\textsuperscript{br} and, more in particular, of unCD56\textsuperscript{dim} NK cells. This latter donor-derived NK cell subset is detectable from the 2\textsuperscript{nd} week after transplantation and its very early expansion and phenotypic features urged us to hypothesize that it may represent an additional stage of maturation before or between cCD56\textsuperscript{br} and cCD56\textsuperscript{dim} NK cells. Indeed, unCD56\textsuperscript{dim} NK cells do not express NK cell precursor features, while they are characterized by typical markers of differentiated NK cells as NKG2D and NKp30. Of note, these cells lack or express at very low level NKp46 enabling a further discrimination from cCD56\textsuperscript{br} NK cells, which are almost totally NKp46\textsuperscript{pos}. Furthermore, unCD56\textsuperscript{dim} NK cells express lytic granules indicative of a mature cytotoxic phenotype. These flow cytometric data are confirmed by transcriptional profile suggesting that unCD56\textsuperscript{dim} NK cells could represent an intermediate subset between cCD56\textsuperscript{br} and cCD56\textsuperscript{dim} NK cells, properly armed to perform alloreactive responses [238].

We also demonstrated that several cytokines, characterizing the systemic lymphopenic environment early after allogeneic HSCT [360], play an important role in cell development
shaping NK cell subset features. In particular, purified unCD56\textsuperscript{dim}/NKp46\textsuperscript{neg-low} NK cells proliferate, are able to generate themselves and, in a certain degree, to generate also cCD56\textsuperscript{br}/NKp46\textsuperscript{pos} NK cells upon \textit{in vitro} incubation with lonely IL-15+IL-18. At similar extent, a small fraction of FACS-sorted cCD56\textsuperscript{br}/NKp46\textsuperscript{pos} gives origin to unCD56\textsuperscript{dim}/NKp46\textsuperscript{neg-low} NK cells, thus losing the phenotype of the derived NK cell subset. According to literature, neither unCD56\textsuperscript{dim} nor cCD56\textsuperscript{br} NK cell subpopulations can generate cCD56\textsuperscript{dim} NK cells, a subset of already differentiated quiescent cells [365, 366].

This experimental setting underlined a characteristic plasticity of both cCD56\textsuperscript{br} and unCD56\textsuperscript{dim} NK cells, both able to proliferate and to differentiate into another NK cell subset upon a proper stimulation, while it did not answer to the question in regard to the origin of unCD56\textsuperscript{dim} NK cells.

Our results in h-HSCT patients are in line with previous studies showing that unCD56\textsuperscript{dim} NK cells are highly cytotoxic, although poorly represented in healthy donors [194, 202, 205].

Moreover, under physiologic conditions, the origin of unCD56\textsuperscript{dim} NK cells can be explained by the metalloproteinase-17 (ADAM-17)-mediated cleavage of CD16 upon cCD56\textsuperscript{dim} degranulation. The ADAM-17-dependent cleavage is restricted to activated cCD56\textsuperscript{dim} NK cells, meaning that this metalloprotease acts mainly in \textit{cis}. Moreover, the loss of CD16 is correlated to increased effector-functions [367, 368]. However, considering that mature NK cells infused with the graft do not survive to PT-Cy [234], it is unlikely that early expanded unCD56\textsuperscript{dim} in patients can derive from the action of ADAM-17 on activated cCD56\textsuperscript{dim} NK cell subset arising later after h-HSCT. Furthermore, not even cytokine stimulation is able to induce cCD56\textsuperscript{dim} proliferation and generation of unCD56\textsuperscript{dim} NK cells.

Additional experiments are certainly needed in order to better recapitulate the lymphopoiesis processes and to disclose the mechanisms that finely tune the transition to unCD56\textsuperscript{dim} to CD56\textsuperscript{br} NK cells and/or \textit{vice-versa} both under physiologic conditions and in lymphopenic environment of HSCT [369, 370]. In adults, the BM is the main site of NK cell generation and maturation, where the BM niche supports NK cell development through the interaction with stromal cells, cytokines, growth factors and other soluble molecules [369, 371].

Thus, the early stages of NK cell differentiation certainly require cell-to-cell contact and soluble molecules mainly produced by stromal cells in the BM as SCF (ligand of c-KIT/CD117), FLT3L, and IL-7 [372]. On the other hand, the late stages of differentiation need IL-2 and/or IL-
as mentioned in the Introduction section. Indeed, the use of these cytokines as clinical therapeutic strategy to enhance NK cell IR, thus providing a faster development and a better reactivity against target cells, is under investigation [92].

In this regard, we are isolating mesenchymal stromal cells from bone fragments (obtained from hip surgery) with the aim of recreating the complex interactions between BM niche and NK cell precursors in an in vitro setting of maturation (see Materials and methods section).

Our study also pointed out the phenotypic differences between immune-reconstituting NK cells in h-HSCT patients and their healthy counterparts. Specifically, we observed that almost all unCD56dim NK cells highly expanded from the 2nd week post-h-HSCT are characterized by NKG2Apos/NKp46neg-low phenotype soon after h-HSCT.

We showed that the transient increase of CD94/NKG2A inhibitory receptor on unCD56dim NK cells of h-HSCT is responsible, at least in part, for their reduced effector-functions. Indeed, although properly armed to be cytotoxic, patients-derived unCD56dim/NKG2Apos NK cells show impaired degranulation capability, thus affecting their potential alloreactive effect necessary for clearing tumor cell targets. We demonstrated that this temporary functional exhaustion can be restored by the blocking of CD94/NKG2A with a masking mAb, thus including this iNKR among the targetable immune checkpoints also for hematologic malignancy treatment. Indeed, the in vivo administration of anti-CD94/NKG2A (i.e. monalizumab) is now a recent therapeutic approach developed to improve both T lymphocyte and NK cell cytotoxicity both against solid tumors and leukemic cells [92, 373].

10.2 Adaptive ml-NK cells in h-HSCT patients

In the context of h-HSCT, the delay of cell IR leads to an increased risk in developing opportunistic infections, thus worsening patients clinical outcomes [40, 47]. Notably, the 50% of h-HSCT patients experiencing HCMV reactivation in the first months after transplantation [42, 123-126]. Emerging evidence suggest that HCMV infection/reactivation can influence NK cell features after HSCT, promoting the expansion of a dysfunctional CD56neg NK cell subset [319, 330] and driving the development of adaptive NK cells, named ml-NK cells [322].

In this work, we confirmed that HCMV greatly affects NK cell subset development by accelerating NK cell IR after h-HSCT, leading to a faster maturation of cCD56dim NK cells with a concomitant expansion of unusual CD56neg NK cell subset, barely detectable in patients that do
not experience a HCMV infection/reactivation (namely NR) [319, 330]. Interestingly, this latter NK cell subpopulation arises soon after the occurrence of HCMV infection/reactivation and it persists at one year after h-HSCT that is our last time point disposable for sample collection, indicating stable HCMV-driven changes in NK cell subset distribution. In accordance, in different transplantation platforms, recent works showed that this CD56^{neg} NK cell subset is detectable also till two years from the day of graft infusion [374]. Moreover, our data demonstrated that patients able to control the viral reactivation without anti-viral treatment administration (namely R2 patients) show a significantly higher expansion of these cells with respect to R1 treated patients. Based on these results we hypothesize that CD56^{neg} NK cells could represent a subset of exhausted cells that might derived from activated cCD56^{dim} NK cell subset after target engagement. Indeed, it is conceivable that the anti-viral treatment, while controlling the HCMV viremia, can limit the need of expansion of mature NK cells efficiently disposed to solve the viral infection/reactivation and the likely differentiation in CD56^{neg} NK cells after having performed their cytotoxicity against viral-infected cells [278].

Taking advantages of novel computational approaches, we analysed flow cytometry high-dimensional single-cell data using a graph-based clustering algorithm that enabled us to identify 28 phenotypically distinct clusters of identical NK cells among all samples (derived from different patients at different time points after h-HSCT) [347]. This newly method confirmed that cCD56^{dim} and CD56^{neg} share similar phenotype and features, and both of them are more mature than unCD56^{dim}, cCD56^{br}, and unCD56^{br} NK cells, which arise early after h-HSCT [238].

The deep characterization of cluster and marker expression demonstrated that HCMV induces also the shaping of NK cell receptor repertoire based on the time of viral infection/reactivation occurrence. Indeed, five clusters resulted to be enriched only in R patients, both soon after the HCMV reactivation and at later time points, suggesting a persistent modification of NK cell features over the time after transplantation. Thanks to this approach we disclosed a peculiar NK cell subset defined by a CD158b1b2^{pos}NKG2A^{neg}NKG2C^{pos}NKp30^{low} phenotype and mainly enclosing cCD56^{dim} NK cells.

CD158b1b2^{pos}NKG2A^{neg}NKG2C^{pos}NKp30^{low} NK cells expand every time HCMV reactivation occurs and, once again, belong mainly to cCD56^{dim} and CD56^{neg} subsets. Of note, their frequency positively correlates with HCMV viremia, thus “remembering” the virus challenge. In accordance to this hypothesis, we observed that, compared to recipients undergoing
a single HCMV reactivation event, patients experiencing multiple HCMV reactivation events are characterized by higher expansion of $\text{CD158b1b2j}^{\text{pos}}\text{NKG2A}^{\text{neg}}\text{NKG2C}^{\text{pos}}\text{NKp30}^{\text{low}}$ NK cells every time a viral load higher than 4000 IU/ml is reached. Finally, this newly disclosed NK cell subpopulation is barely present in NR patients, where its absence is counterbalanced by presence of a NK cell subset endowed with an opposite phenotype: $\text{CD158b1b2j}^{\text{neg}}\text{NKG2A}^{\text{pos}}\text{NKG2C}^{\text{neg}}$. These $\text{CD158b1b2j}^{\text{neg}}\text{NKG2A}^{\text{pos}}\text{NKG2C}^{\text{neg}}$ NK cells are also less differentiated than $\text{CD158b1b2j}^{\text{pos}}\text{NKG2A}^{\text{neg}}\text{NKG2C}^{\text{pos}}\text{NKp30}^{\text{low}}$ NK cells, being mainly $\text{unCD56}^{\text{dim}}$, $\text{cCD56}^{\text{br}}$ NK cells. These findings further confirm the HCMV impact on NK cell development and features. The persistence of $\text{CD158b1b2j}^{\text{pos}}\text{NKG2A}^{\text{neg}}\text{NKG2C}^{\text{pos}}\text{NKp30}^{\text{low}}$ NK cells after the resolution of infection, together with the markers expressed, strongly remind the adaptive ml-NK cells described in literature.

On the contrary, our results highlighted that our ml-NK phenotype is not marked by increased CD57 surface levels [317, 318]. Indeed, we observed a significant higher expression of this marker of terminal differentiation and senescence only at 8-12 months after h-HSCT following HCMV infection/reactivation. This could mean that the acquisition of CD57 expression occurs only when NK cell IR after h-HSCT is completed and the majority of NK cells are characterized by a fully mature phenotype.

Given the not yet specified ml-NK cell phenotype, we wondered if our newly disclosed subset consists of real ml-NK cells. Thus, we assessed RNA expression profiles of our putative $\text{CD158b1b2j}^{\text{pos}}$ ml-NK cells, expanded in R patients but having a sufficiently high percentage of cells also in NR ones, by NGS in order to investigate ml-NK cell fingerprint, exceeding flow cytometry limitations. To focus on long-living ml-NK cells, we selected samples at 7-12 months post-h-HSCT when NK cell IR is already completed and it would not interfere with the HCMV-driven changes in NK cell phenotype and functions.

The differences highlighted within R h-HSCT patients confirmed that $\text{CD158b1b2j}^{\text{pos}}$ NK cells are more mature/educated than $\text{CD158b1b2j}^{\text{neg}}$ and they show also an increased activation status regulated by the enrichment of metabolic pathways, as OxPhos and Cell-cycle. Finally, $\text{CD158b1b2j}^{\text{pos}}$ NK cells resemble ml-NK cells highly expressing NKG2C, while downregulating NKG2A and cytokine receptors. This first comparison pushed us to deeply focus on similarly educatated $\text{CD158b1b2j}^{\text{pos}}$ NK cells for further analyses.
The absence of significant changes in gene expression between CD158b1b2j\(^{\text{pos}}\) NK cells derived from patients that never encounter the virus (D-/R-) and from patients that do not reactivate (R0) suggests that NK cell shaping in receptor repertoire is mainly driven by viral reactivation after h-HSCT instead of patients HCMV-serostatus. Moreover, also NK cell subset distribution post-h-HSCT is not affected by previous exposure to HCMV during life. In other words, this finding suggests that the CD158b1b2j\(^{\text{pos}}\) NK cells are not transferred from the donor to the recipient with the graft but could eventually be \textit{de novo} generated or expanded after HCMV exposure. However, only the analysis of patient relative donors will answer to this important point.

Looking more in detail within the cohort of R patients the little differential expression of RNA transcripts between R1 and R2 recipients indicates that, despite the anti-viral treatment could affect the frequency of our newly identified subset, it may not have a significant effect on NK cell gene expression at almost one year from transplantation.

Given these results, we proceeded in the analysis among CD158b1b2j\(^{\text{pos}}\) NK cells in NR and the R recipients regardless of the HCMV-serostatus and the anti-viral treatment, respectively. According to literature, we observed a reduced expression of adaptor molecules as PLZF, SYK, EAT-2, and FC\(\varepsilon\)R\(\gamma\) upon HCMV reactivation compared to NR recipients at transcriptional levels, further supporting the hypothesis that our HCMV expanded NK cell subset is composed by ml-NK cells [310, 312]. Moreover, these molecules, together with NKG2A, NKp30, perforin, and granzyme-\(\beta\), are involved in the KEGG pathway of \textit{NK cell mediated cytotoxicity} significantly reduced in R patients. The downregulation of \textit{KLRC1} and \textit{NCR3} genes, codifying for NKG2A and NKp30, respectively, in R recipients confirmed our previous flow cytometric results and strengthened our FACS-sorting gating strategy based on CD158b1b2j\(^{\text{pos}}\)/CD158b1b2j\(^{\text{neg}}\) NK cells. On the contrary, the expression of \textit{KLRC2} gene, codifying for NKG2C, appears to be upregulated in R patients only applying less stringent parameters (p-value= 0.02 vs p-value adjusted= 0.17). Moreover, a general higher NK cell cytotoxicity in NR patients than in R ones could be explained by the fact that ml-NK cells are more functionally efficient against specific targets (i.e. cells expressing HLA-E molecules and HCMV infected cells), no more present at several months from the last pick of viremia detected. However, ml-NK cells should retain the intracellular cytotoxic granules in order to be prone to exert their effector-functions upon a challenge with the same antigen [319, 322]. We hypothesize that the not expected
reduced expression of perforin, and granzyme-β could be due to the presence of low percentages of CD56<sup>neg</sup> cells within the FACS-sorted CD158b1b2j<sup>pos</sup> NK cell subset in R patients (not shown).

Interestingly, CD158b1b2j<sup>pos</sup> NK cells from R recipients express more IFN-γ than in NR patients, suggesting that modifications in IFNG locus occurs upon HCMV infection in accordance with recent works in literature [315]. However, our functional preliminary experiments failed to demonstrate a higher IFN-γ release from CD158b1b2j<sup>pos</sup> than CD158b1b2j<sup>neg</sup> NK cells, thus suggesting that these cells could be somehow impaired in IFN-γ release. In this scenario, our results also highlighted increased levels of PD-1, PD-L1, and LAG3 inhibitory checkpoints in R recipients. Evidence in literature underlines that both PD-1 and LAG3 resulted to be upregulated primarily in T cells upon chronic stimulation [375, 376] and their blockade enables the restoration of T cell cytotoxicity [338]. Moreover, recent work showed that chronic stimulation through NKG2C in combination with IL-15 induces also adaptive NK cells proliferation and enhances both LAG3 and PD-1 expression [363].

For these reasons we performed masking experiments by blocking either PD-1 or LAG3 in the presence or in the absence of tumour target cells expressing their putative ligands. The blockade of these inhibitory checkpoint only partially roused NK cell effector-functions in R h-HSCT patients analysed at 8-12 months post-transplantation and indeed NK cell response in R samples never exceeded the one observed in NR h-HSCT patients. These results highlight that a general impairment of NK cell functions, not only driven by checkpoint inhibitors, could be responsible of multiple reactivations in HCMV non-controller recipients, thus greatly impact on the clinical outcome of h-HSCT patients.
11 BIBLIOGRAPHY


376. Schnorfeil, F.M., et al., T cells are functionally not impaired in AML: increased PD-1 expression is only seen at time of relapse and correlates with a shift towards the memory T cell compartment. J Hematol Oncol, 2015. 8: p. 93.