Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

M.Sc. Anna Mathioudaki

born in: Amarousio, Attiki, Greece

Oral examination: 30$^{\text{th}}$ June 2023

Deciphering the Interplay of

Acute Myeloid Leukemia

with the Bone Marrow Microenvironment

Referees:

Prof. Dr. Stefan Wiemann

Prof. Dr. Oliver Stegle

This thesis was conducted at the European Molecular Biology Laboratory (EMBL), Heidelberg, Germany in collaboration with the Heidelberg University Hospital, starting from October 2018 to April 2023 and was supervised by Dr. Judith Zaugg and Dr. Med. Caroline Pabst.

SUMMARY

Acute myeloid leukemia (AML) is a hematologic malignancy that leads to the accumulation of immature blasts in the bone marrow (BM). The BM is a complex organ, consisting of several cell types, including immune cells as well as non-hematopoietic stromal cells. In AML, chemotherapy may lead to long-term remission, although allogeneic stem cell transplantation (alloSCT) often remains the only therapeutic strategy. However, not every patient responds to alloSCT and often suffer from relapse due to chemo-resistant leukemic stem cells (LSCs).

One of the hypotheses associated with therapy failure is the incapability of donor T cells to recognize and eliminate LSCs, thus escaping graft-versus-leukemia (GVL) effect. The first objective of this thesis was to investigate the role of T cell in alloSCT therapy outcome. Using single-cell RNA-sequencing (scRNA-seq) on BM T lymphocytes and $CD34^+$ hematopoietic stem and progenitor cells (HSPCs) of six AML patients 100 days after alloSCT, I identified T cell signatures associated with either relapse (REL) or complete remission (CR). Among these signatures, a higher frequency of cytotoxic $CD8^+$ effector and gamma delta T cells was observed in CR versus REL samples. Further analyses revealed that in CR, $CD8^+$ T cells were more mature and characterized by higher cytotoxicity, while in REL $CD8^+$ T cells were characterized by inflammatory TNF/NF-κB signaling as well as an immunosuppressive signature. In addition, this analysis identified *ADGRG1*/GPR56 as a surface marker enriched in CR $CD8^+$ T cells. Additional flow cytometry analyses in independent patient cohorts suggested GPR56 as a marker of cytotoxicity as well as a marker of antigen encounter post alloSCT. Together, these data provide a single-cell reference map of BM-derived T cells post alloSCT and propose GPR56 expression dynamics as a surrogate for monitoring alloSCT.

One of the key drivers of AML progression is its interaction with the BM stromal microenvironment. In addition, AML is hypothesized to remodel the BM, creating a protective environment for LSCs. Thus, the second objective of this thesis was to study the impact of AML on the microenvironment and the specific contribution of LSCs in this process. For that, I combined scRNA-seq of AML xenograft models and *in vitro* co-cultures of patient-derived BM mesenchymal stromal cells (MSCs) with AML. These data indicated that AML presence impacts the BM composition, leading to the expansion of Cxcl12-abundant-reticular adipocyte progenitors (Adipo-CAR), decline of osteoblasts as well as disruption of the vasculature. When comparing high LSC-frequency (LSC_{high}) with low LSC-frequency (LSC_{low}) AML, changes in the abundance of several stromal subsets were detected, suggesting the importance of these populations in LSC expansion.

Specifically, LSC_{high} AML was associated with a decrease in the osteo-lineage and an expansion of multiple fibroblast subsets marked by altered extracellular matrix signatures. Furthermore, *in vitro* co-cultures uncovered similar patterns: upon LSC_{high} co-culture, human MSCs suppress the expression of osteoblast lineage genes, while over-expressing fibrosis-related genes.

Collectively, the findings outlined in this thesis provide novel insights into the interaction of AML with the BM microenvironment, which has implications in disease progression and therapy outcome. These insights offer new opportunities for identifying intervention targets which may improve AML patient outcome.

ZUSAMMENFASSUNG

Die akute myeloische Leukämie (AML) ist eine hämatologische Malignität, die zu einer Anhäufung von unreifen Blasten im Knochenmark (KM) führt. Das KM ist ein komplexes Organ, das aus verschiedenen Zelltypen besteht, darunter Immunzellen und nicht-hämatopoetische Stromazellen. Bei der AML kann eine Chemotherapie zu einer langfristigen Remission führen, obwohl die allogene Stammzelltransplantation (alloSCT) oft die einzige kurative Therapiestrategie bleibt. Allerdings sprechen nicht alle Patienten auf eine alloSCT an und erleiden häufig einen Rückfall aufgrund von chemoresistenten leukämischen Stammzellen (LSC).

Eine der Hypothesen, die mit dem Scheitern der Therapie in Verbindung gebracht werden, ist die Unfähigkeit der Spender-T-Zellen, LSCs zu erkennen und zu eliminieren und so dem Graft-versus-Leukämie-Effekt (GVL) zu entgehen. Das erste Ziel dieser Arbeit war es, die Rolle der T-Zellen für den Erfolg der alloSCT-Therapie zu untersuchen. Mithilfe der Einzelzell-RNA-Sequenzierung (scRNA-seq) von KM-T-Lymphozyten und $CD34^+$-Zellen von sechs AML-Patienten 100 Tage nach alloSCT identifizierte ich T-Zell-Signaturen, die entweder mit einem Rückfall (REL) oder einer kompletten Remission (CR) assoziiert sind. Unter diesen Signaturen wurde eine höhere Häufigkeit von zytotoxischen $CD8^+$ Effektor- und Gamma-Delta-T-Zellen in CR- gegenüber REL-Proben beobachtet. Weitere Analysen ergaben, dass die $CD8^+$ T-Zellen in CR reifer waren und sich durch eine höhere Zytotoxizität auszeichneten, während die $CD8^+$ T-Zellen in REL durch eine entzündliche TNF/NF-κB-Signalisierung sowie eine immunsuppressive Signatur gekennzeichnet waren. Darüber hinaus identifizierte diese Analyse *ADGRG1*/GPR56 als Oberflächenmarker, der in CR $CD8^+$ T-Zellen angereichert ist. Zusätzliche durchflusszytometrische Analysen in unabhängigen Patientenkohorten legten nahe, dass GPR56 ein Marker für Zytotoxizität sowie ein Marker für Antigenbegegnungen nach alloSCT ist. Zusammengenommen liefern diese Daten eine Einzelzell-Referenzkarte der aus dem KM stammenden T-Zellen nach alloSCT und schlagen die Dynamik der GPR56-Expression als Surrogat für die Überwachung der alloSCT vor.

Einer der Hauptfaktoren für das Fortschreiten der AML ist die Interaktion mit der stromalen Mikroumgebung des KM. Darüber hinaus wird angenommen, dass die AML das stromale Mikromilieu umgestaltet und so ein schützendes Umfeld für die LSCs schafft. Daher bestand das zweite Ziel dieser Arbeit darin, die Auswirkungen der AML auf das Mikromilieu und den spezifischen Beitrag der LSCs in diesem Prozess zu untersuchen. Zu diesem Zweck kombinierte ich scRNA-seq von AML-

Xenotransplantationsmodellen und In-vitro-Kokulturen von aus Patienten stammenden mesenchymalen Stromazellen (MSCs) mit AML. Diese Daten zeigten, dass das Vorhandensein von AML die KM-Zusammensetzung beeinflusst, was zu einer Expansion von Cxcl12-überschüssigen retikulären Adipozytenprogentiros (Adipo-CAR), einem Rückgang der Osteoblasten sowie einer Störung des Gefäßsystems führt. Beim Vergleich von AML mit hoher LSC-Häufigkeit (LSC_{high}) mit AML mit niedriger LSC-Häufigkeit (LSC_{low}) wurden Veränderungen in der Häufigkeit verschiedener stromaler Untergruppen festgestellt, was auf die Bedeutung dieser Populationen für die LSC-Expansion hindeutet. Insbesondere war LSC_{high} AML mit einer Abnahme der Osteo-Linie und einer Zunahme mehrerer Fibroblasten-Untergruppen verbunden, die durch veränderte extrazelluläre Matrixsignaturen gekennzeichnet waren. Darüber hinaus ergaben In-vitro-Kokulturen ähnliche Muster: Bei LSC_{high}-Kokulturen unterdrücken menschliche MSZ die Expression von Genen der Osteoblasten-Linie, während sie Fibrose-bezogene Gene überexprimieren.

Insgesamt bieten die in dieser Arbeit dargelegten Ergebnisse neue Einblicke in die Interaktion von AML mit der Mikroumgebung des KM, was Auswirkungen auf den Krankheitsverlauf und die Therapieergebnisse hat. Diese Erkenntnisse bieten neue Möglichkeiten zur Identifizierung von Interventionszielen, die das Ergebnis bei AML-Patienten verbessern können.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my supervisors Judith Zaugg and Caroline Pabst for giving me the chance to be here and for baring with me while learning what a hematopoietic stem cell is... Judith, you've been a great mentor, always there for me, patiently teaching me how to be a scientist, well and patience too. You can consider yourself "responsible" for ambition to stay in science, you've facilitated an environment of great creativity. And Caroline I want to thank you for giving me a glimpse of the clinical reality and teaching me how to do research with the ultimate goal of achieving clinical impact. I would like to extend my thanks to my committee members Oliver Stegle, Peer Bork, Stefan Wiemann and Benendikt Brors for generously giving their time to support me throughout this process.

I couldn't have done any of this, without the Zaugg group, current and old members. I would like to thank the 113 office-Christian, Max, Kristy, Bogdan, Miguel, Evi, Frosina, Ivan, Aryan and Guido for all the loud laughs, creative procrastinations and the care-beers when working late. Christian, keep being Christian! All the Zaugg girls, Karin, Annique, Nila, Sara, Olga, Gwen and Rim for having an open ear at all times. And the rest -Mikael, Charles, Brian, Dirk, Yaren, Gerard and Victor for all the feedback and the memories.

I would also like to acknowledge the invaluable contributions from Lukas, Xizhe and Richard throughout my PhD. They've always been kind and willing to share their experiences. And thank my MMPU buddies Daniel, Josephine, Jonas, Michi, Cornelius and Max for all the great food we've shared, and for an eventful conference in Vienna last year.

My PhD experience has been highly collaborative, and I am very grateful for all the inspiring people I got to interact with. Special thanks to Sascha Dietrich for his guidance, as well as the rest of the "Lymphoma team" - Simon Haas, Daniel Huebschmann, Lea, Dominik and Raphael for making Friday evening meetings something to look forward to. And most importantly to Felix, for introducing me to this "world" and for his endless encouragement through these years.

My time at EMBL has been great, not only because of the fun science but also because of the people I had the luck to share this journey with. So, I would like to thank… These girls-Agata, Karolina and Lucia-my "partner in crime", for always speaking the truth. The therapy couch - Wolfy[G], Gilberto and Fergus for Mario Kart-ing away our PhD struggles. A dear fellow, Alex, for his unique ability to rationalize stressful situations and for transforming thesis writing into a fun competition. My writing

buddies, Luca, Aline and Max for spending Easter holidays together, definitely not the way we wanted to... And the rest of the 2018 batch, the party people, the defence parties' organizers -Jesus, Alberto, Matteo, Elisa, Charlie, Javier, Maxime, Andrea, Ana, Martina, Wolfy^{P}. I consider myself extremely lucky for meeting all of you, for standing next to each other all these years, through pandemics, submissions, or when life just happened…

Special thanks to my home away from home, Elena, Danae and Fenia, with whom I hope to build a little settlement, so if you read this, be prepared… As well as to my actual home-my lovely flatmates Diana, Shubham and Alex for being themselves.

Even from a distance, some people's support can still reach you… I would like to thank my Greek friends, that we met at university 10 years ago, and since then, even when we live countries or continents apart, our bond remains strong and it always feels like no day went by.

Last, and certainly not least, I would like to express my gratitude to my parents, Antonios and Kalliopi, for teaching me determination, for believing in me, for being there to listen and advise and for always finding ways to support me! *Ευχαριστώ…*

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AML	Acute myeloid leukemia
alloSCT	Allogeneic stem cell transplantation
BM	Bone marrow
CAR cells	CXCL12 abundant reticular cells
CAR-T cells	Chimeric antigen receptor T cells
CB	Cord blood
cDC	Conventional dendritic cell
CMV	Cytomegalovirus
CITE-seq	Cellular Indexing of Transcriptomes and Epitopes by sequencing
CR	Complete remission
CRISPR/Cas9	Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated protein 9
DE	Differential expression
DEGs	Differentially expressed genes
EC	Endothelial cell
FACS	Fluorescence activated cell sorting
GO	Gene ontology
GPCR	G protein coupled receptor
GRN	Gene regulatory network
GSVA	Gene set variation analysis
GVHD	Graft versus host disease
GvL	Graft versus leukemia
HLA	Human leukocyte antigen
HSPC	Hematopoietic stem and progenitor cell
IFN	Interferon
KO	Knock-out
LMPP	Lymphoid-primed multipotent progenitors
LSC	Leukemic stem cell
MAIT	Mucosal associated invariant T cells
MDP	Monocyte-dendritic progenitors
MEP	Megakaryocyte-erythrocyte progenitors
MLP	Myeloid/lymphoid progenitors
MP	Monocytic precursors
MSC	Mesenchymal stromal cells

NK cells	Natural killer cells
NP	Neutrophil progenitors
PBMCs	Peripheral blood mononuclear cells
pDC	Plasmacytoid dendritic cell
QC	Quality control
REL	Relapse
sgRNA	Single guide RNA
scRNA-seq	Single cell RNA sequencing
TCM	Central memory T cells
TEM	Effector memory T cells
TEMRA	$CD45RA^+$ effector memory T cells
TF	Transcription factor
Treg	Regulatory T cell
UMAP	Uniform manifold approximation and projection

LIST OF FIGURES

LIST OF TABLES

PRELUDE

But the story of leukemia--the story of cancer--isn't
the story of doctors who struggle and survive,
moving from one institution to another. It is the
story of patients who struggle and survive, moving
from one embankment of illness to another.
***Siddhartha Mukherjee**, The Emperor of All Maladies*

The human body consists of trillions of cells, organized in tissues, which give rise to 78 organs. These organs occupy specified locations in the human body. Additionally, blood circulates around the body, where its components interact with different cell types from different environments. The process of the generation of the blood subsets is termed hematopoiesis, which takes place in the bone marrow, producing billions of cells every day.

During the extraordinary biological process of hematopoiesis, errors may occur in the form of genetic mutations. Even though the majority of such errors are corrected through the DNA repair machinery, some remain uncorrected, grow and prevail undetected by the immune system. Such mutations can equip the cell with a proliferative advantage that allows cells to divide uncontrollably and evade neighbouring tissues. These cells are generally termed as **"cancer cells"**, while in the case of blood they are specifically known as leukemia or lymphoma cells, depending on the type of blood cancer[1].

Thus, through a series of fortunate events for this abnormal cell, but unfortunate for the individual, one cell can transform into a cancer cell. However, the survival of a cancer cell is dependent on extrinsic factors as well, since encountering an unfavourable environment could end up in its elimination.

In 1978 Raymond Schofield introduced the term **"stem cell niche"** as the microenvironment for hematopoietic stem cell maintenance in the bone marrow[2]. The niche is defined by its cellular components, their properties as well as their interplay[2]. In the case of malignancy, the neighboring cells which comprise the "niche" exhibit cancer associated gene signatures like stress and interferon response. Though how this niche remodelling is induced, whether it is cancer permissive or simply a consequence, remains to be clarified.

CHAPTER 1: INTRODUCTION & BACKGROUND

1.1 The Bone Marrow Niche in healthy hematopoiesis

Hematopoietic stem cells (HSCs) are capable of self-renewal and they are responsible for the generation of blood and immune cells. HSCs reside in the bone marrow (BM), often termed as the HSC niche. The BM is a large tissue which consists of >99% of hematopoietic cells and <0.1% non-hematopoietic cells[3].

The HSCs of an adult human BM produces 500 billion blood cells daily[4], of either myeloid or lymphoid lineage as well as a tiny fraction of hematopoietic stem and progenitor cells (HSPCs, ~1%)[3]. The HSPCs give rise to all blood cell types broadly classified into lymphoid lineage that gives rise the adaptive immune system, such as B and T cells, and the myeloid lineage that gives rise to erythrocytes and the innate immune system, such as macrophages and granulocytes (**Figure 1**).

Specifically, long-term-HSCs (LT-HSCs) differentiate to short-term-HSCs (ST-HSCa) and subsequently multipotent progenitors (MPPs)[5]. MPPs do not self-renew and give rise to two distinct branches of hematopoiesis, the myeloid and the lymphoid branches thus being considered as the first divergence towards the 2 lineages[5]. Common myeloid progenitors (CMP) generate granulocyte-macrophage progenitors (GMP) and Megakaryocyte-erythrocyte progenitors (MEP) which further differentiate into distinct cell types. Similarly, common lymphoid progenitors (CLP) produce the two different lymphoid branches of the innate adaptive system (T and B cells) as well as natural killer cells (NK). Notably, due to the advances in single cell technologies, the classical model of hematopoiesis has been challenged. These studies support a rather continuous differentiation model where individual HSCs gradually acquire lineage biases rather than committing to strictly defined progenitor populations[6].

Upon commitment to a certain lineage, progenitors migrate to distinct niches either within the BM or in other tissues like the thymus[5]. While the majority of HSPCs are located in close proximity to microvessels[7], there have been reports of differential distribution in the different BM compartments. These compartments are structurally defined by non-hematopoietic cells, thus highlighting the role of the stroma in hematopoiesis.

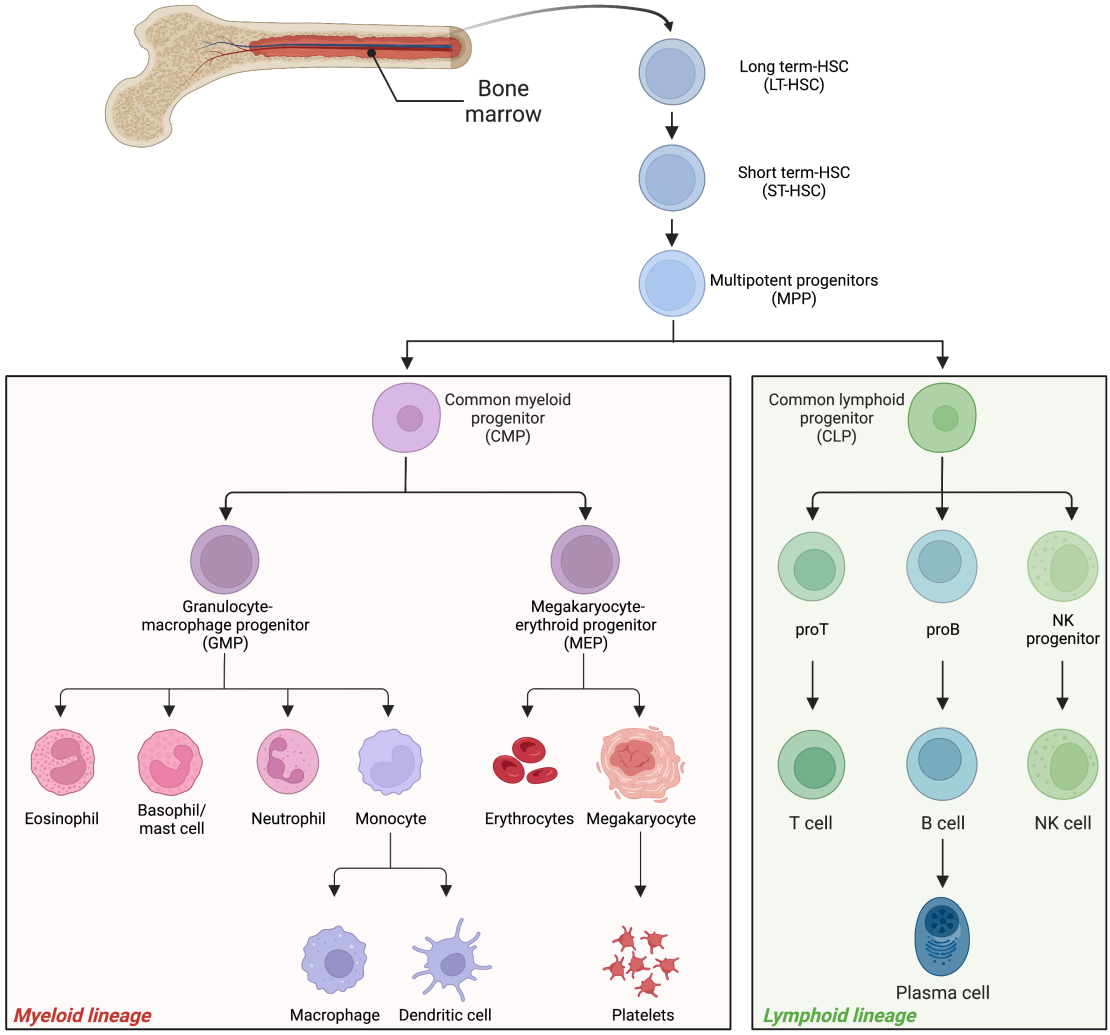

Figure 1 Schematic explaining the classical model of hematopoiesis.
Hematopoietic stem cells (HSCs) reside in the bone marrow (BM) and comprise only a tiny fraction of all the hematopoietic cells. Hematopoiesis follows the developmental sequence from long-term-HSCs to short-term-HSCs and subsequently multipotent progenitors (MPPs). MPPs do not self-renew and give rise to two distinct branches of hematopoiesis, the myeloid and the lymphoid branches. Common myeloid progenitors (CMP) generate granulocyte-macrophage progenitors (GMP) and Megakaryocyte-erythrocyte progenitors (MEP) which further differentiate into distinct cell types. Similarly, common lymphoid progenitors (CLP) produce the different lymphoid branches. Schematic created using Biorender.com.

1.1.2 Stromal cells: Key Regulators of the Bone marrow niche

Several BM resident cell types have been proposed to regulate hematopoiesis, including non-hematopoietic subsets of mesenchymal and endothelial lineage, which are estimated to make up for 0.1% of the BM[3]. These components define BM niches, a term initially used to describe the regulatory unit which preserves and directs the renewal and differentiation of HSPCs[2]. The functionally and spatially distinct BM niches have been

termed as endosteal, sinusoidal, and arteriolar depending on their location and composition[3,8].

Multiple studies have proposed specific BM niches to serve as specialized 'homes' for distinct hematopoietic subpopulations. For example, quiescent HSCs are preferentially found in endosteal BM and are associated with small arterioles[9]. Platelet and myeloid biased HSCs enriched in megakaryocyte niches while the $NG2^+$ arteriolar niche is associated with lymphoid biased HSCs[10].

In the following paragraphs, a brief introduction to each non-hematopoietic cellular component of the BM is given and an illustration of this environment can be seen in **Figure 2A**.

Mesenchymal stromal cells (MSCs) are multipotent cells that can differentiate into various cell types like osteoblasts, chondrocytes, adipocytes and fibroblasts[11] **(Figure 2B)**. BM derived MSCs are critical for the regulation of hematopoiesis and are considered to be quite heterogeneous since colonies derived from a single MSC can behave differently in terms of differentiation and proliferation[11]. It has been shown that MSCs are essential for proliferation, differentiation and quiescence of HSCs *in vivo* and *in vitro*[12,13].

Moreover, the different MSC derived subsets modulate hematopoiesis. Osteoblasts regulate the function[14] and homing of HSPCs after transplantation[15] while they also facilitate the maturation of HSPCs towards the B lymphocyte lineage[16]. On the other hand, adipocytes reduce hematopoietic activity[17] and inhibit lymphoid differentiation[18]. Fibroblasts are responsible for extracellular matrix (ECM) deposition as well as the secretion of cytokines and growth factors, which regulate HSC self-renewal[19].

BM endothelial cells (ECs) form the network of vasculature which is instrumental for HSPCs function, trafficking and homeostasis. ECs directly affect HSC proliferation by expressing E-selectin adhesion molecule[20], while the lack of EC-specific factors like vascular endothelial growth factor (VEGF) results in BM failure[21]. BM vessels can either be sinusoidal or arteriolar. Stem cells antigen-1 positive ($Sca\text{-}1^+$) arterioles, which expand along the bones, are less permeable and maintain HSCs at a low reactive oxygen species state[22]. Sca-1- sinusoids, which span perpendicularly to the long axis and are more permeable, promote activation of HSPCs and are the site of leukocyte exchange to and from the BM[22].

Blood vessels are surrounded by a variety of pericytes[23]. Pericytes can be smooth muscle cells, contributing to vessel movement, or other cells of mesenchymal origin like C-X-C Motif Chemokine Ligand 12 (CXCL12) abundant reticular (CAR) cells. Evidence mainly from murine models support that CAR cells surround sinusoidal endothelial cells and are critical for maintenance of the quiescent HSC pool through CXCL12-CXCR4

(C-X-C chemokine receptor type 4) signalling[24]. Individual CAR cells have been found to express both osteogenic and adipogenic genes and have been hypothesized to be the adipo-osteogenic progenitors of the BM niche[25]. A more recent study has also identified different subsets of CAR cells, expressing either osteo- or adipo-lineage markers[26].

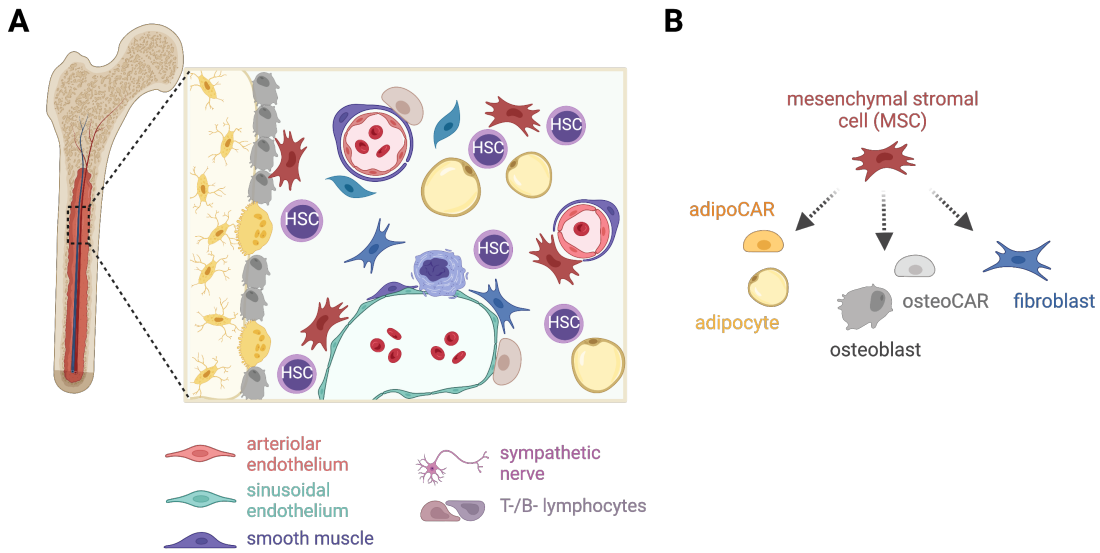

Figure 2 The non-hematopoietic components of the bone marrow niche.
(A) Schematic representation of the hematopoietic stem cell (HSC) niche in the bone marrow (BM) as well as its resident cells. (B) BM mesenchymal stromal cells (MSCs) give rise to cells which form the bone. MSCs give rise to osteoblasts that will subsequently be embedded in the bone matrix to give rise to osteocytes. CXCL12 abundant reticular (CAR) cells, are hypothesized to give rise to either osteo-lineage or adipo-lineage progenitor cells. Modified from schematic from template provided by Dr. Karin Prummel, created using Biorender.com.

1.1.3 Bone marrow resident T cells

Another component of the BM niche are the immune cells. Mature immune cells consist approximately ~20% of the mononuclear cells in the adult human BM, with a ratio of 5:1 T cells/B cells[3]. The majority of the BM resident T cells are antigen-experienced expressing CD44 and CD122, in comparison to the lymph node and the spleen, where the T cells are mainly naive[27].

The 3 main branches of lymphocytes are $CD4^+$ T, $CD8^+$ T and NK cells. While $CD4^+$ T cells assist in the immune system coordination by stimulating other cell types while $CD8^+$ T and NK cells are known for their cytotoxic effect[27].

$CD4^+$ T cells (~1,5% of BM mononuclear cells) can differentiate into different subsets after antigen recognition[27]. This is the consequence of the interaction of T-cell receptor (TCR) and CD4 (co-receptor) with antigen-MHC II complex, presented by antigen presenting cells (APCs)[28]. A co-stimulatory signal from the CD28 co-receptor

must be received in order to achieve activation[28]. This event accompanied with CD3 activation induces downstream signalling pathways which lead to differentiation into the specific $CD4^+$ T cells subsets[28, 29].

These subsets include T-helper 1 (Th1), T-helper 2 (Th2), IL-17^+ $CD4^+$ T-helper 17 (Th17), $CD4^+$ $CD25^+$ regulatory T (Treg) cells. The specification of the different subsets depends on the cytokines present in the microenvironment[28] (**Figure 4**). For example, IL-6, IL-21, IL-23, and TGF-β are the main cytokines inducing Th17 differentiation[29] (**Figure 4**).

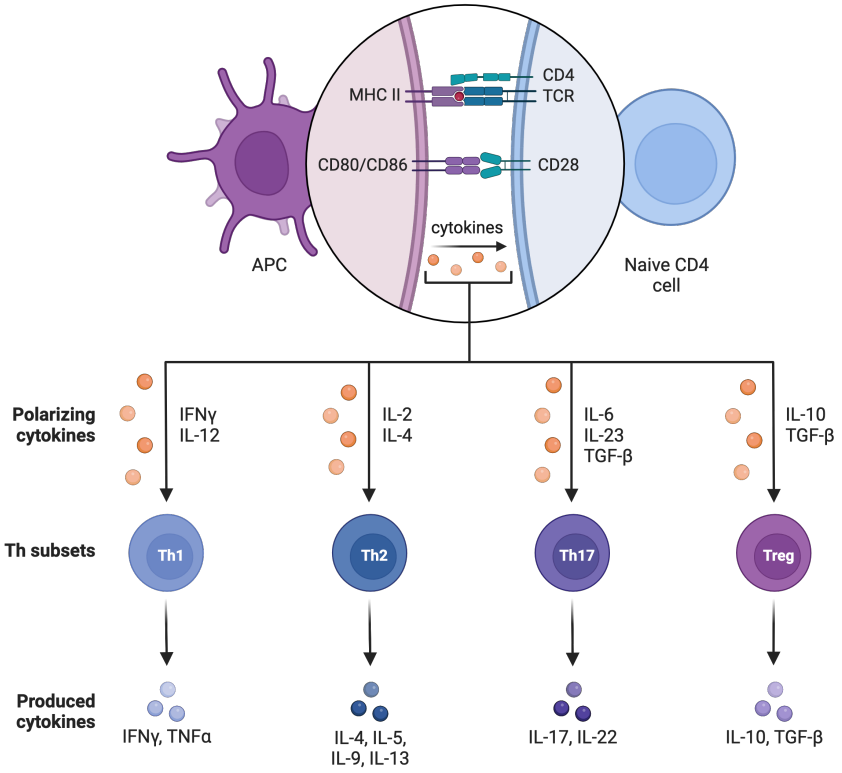

Figure 3 $CD4^+$ naive T cell differentiation into different functional subsets.
$CD4^+$ naive T cells differentiate towards different subsets, after antigen stimulation. This event involves the interaction of CD4 receptor and T-cell receptor (TCR) with antigen-MHC II complex, as well as costimulatory signals coming from CD28 surface protein. After activation, the differentiation to the different T cell subsets occurs and it is driven by different combinations of polarising chemokines. Template provided by Biorender.com.

Similar to the $CD4^+$ T cell differentiation, $CD8^+$ T cells (2-2,5% of BM mononuclear cells) can differentiate from naive to effector cells after antigen stimulation[27]. $CD8^+$ cells are activated after the recognition and interaction of their TCR with the antigen-MHC I complex, which is expressed on almost every cell of our body[30]. To activate the cytotoxic machinery, the co-stimulation of the CD28 co-receptor is

necessary[30]. CD8$^+$ cytotoxic cells recognize antigens derived from intracellular pathogens or tumour-derived neoantigens, leading to cytotoxic killing of infected or cancer cells through the secretion of toxic granzyme proteases like Granzyme B (GZMB) and the pore forming protein Perforin 1 (PRF1))[31] (**Figure 4**).

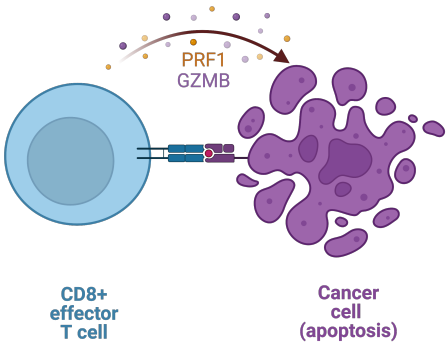

Figure 4 Immune interaction between CD8$^+$ and cancer cells.
In anticancer immunity, CD8$^+$ effector cells main role is to induce cell death through the secretion of molecules like granzyme B (GZMB) and pore forming protein perforin 1 (PRF1). Image created using Biorender.com.

1.2 Acute Myeloid Leukemia (AML)

Acute myeloid leukemia (AML) is a hematologic malignant disorder, in which HSCs grow and differentiate abnormally, leading to the accumulation of immature myeloid precursors in the BM and the peripheral blood[32]. While immature myeloid cells expand, mature myeloid lineages like red blood cells, granulocytes and platelets decrease, leading to peripheral cytopenia[33].

Abnormal hematopoiesis can lead to the production of pre-leukemic HSCs and subsequently leukemic stem cells (LSCs) and AML (**Figure 5**). HSCs can acquire pre-leukemic mutations like DNA (cytosine-5)-methyltransferase 3A (*DNMT3A*) and Tet methylcytosine dioxygenase 2 (*TET2*) which provides them with a proliferative advantage and clonal outgrowth, resulting in a condition called clonal hematopoiesis of indeterminate potential (CHIP) (**Figure 5**)[33]. Additional mutations (e.g., nucleophosmin 1, *NPM1*; fms-like tyrosine kinase 3 gene with internal tandem duplication, FLT3-ITD) in a pre-leukemic clone leads to the transformation of pre-leukemic HSCs to LSCs (**Figure 5**)[33]. A subset of LSCs are therapy resistant thus resulting in cancer reappearance (relapse)[33]. This malignant transformation is a gradual process, and in 30% of the cases AML arises from myelodysplastic syndrome (MDS). MDS is a pre-malignant BM disorder marked by cytopenia, BM dysplasia and abnormal hematopoiesis[34].

Leukemic stem cells (LSCs), often termed as leukemia initiating cells, are defined by their ability to self-renew, engraft into recipient immunocompromised mice and subsequently initiate leukemia[35]. Despite the extensive research in LSCs, their phenotype remains not well understood, thus making it hard to achieve sufficient LSC elimination in therapy. Pabst and colleagues identified G protein–coupled receptor 56 (GPR56) as a novel LSC marker which is independent of the $CD34^{+}$ $CD38^{-}$ phenotype[36]. Moreover, the discovery of the 17-gene LSC score (LSC17), which includes GPR56, allowed clinicians to classify patients into different risk groups prior to therapy since a higher score reflected higher LSC burden and thus resistance to standard therapy[37]. Even though such classifications may be informative, AML heterogeneity defined by genetic subgroups still needs to be addressed. For instance, in *FLT3, DNMT3A*, and *NPM1* triple-mutated high-risk AML, hepatic leukemia factor (HLF) is a critical regulator of LSC fate[38].

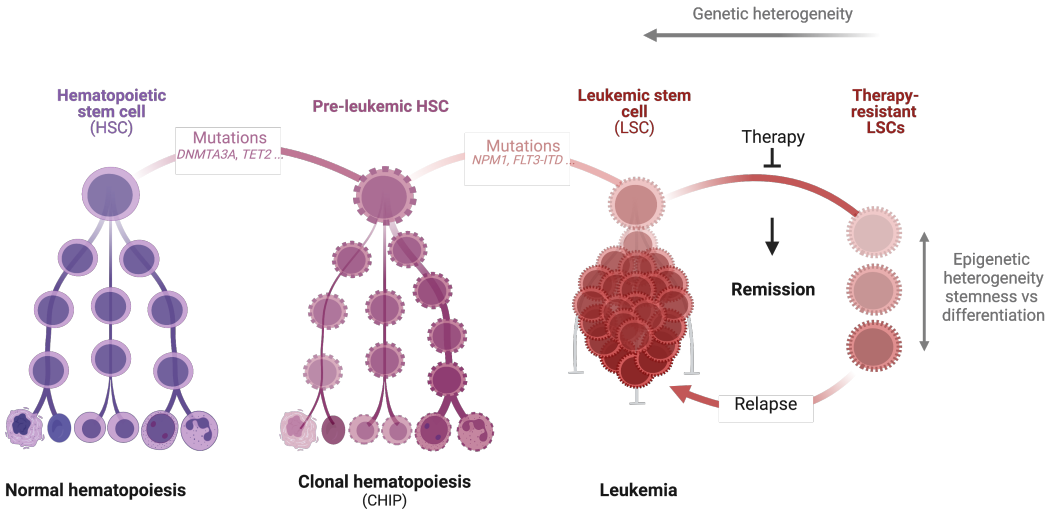

Figure 5 The transformation of healthy HSCs to LSCs.
In healthy situations, hematopoiesis occurs and leads to the generation of the different myeloid and lymphoid branches. HSCs which acquire pre-leukemic mutations like DNMT3A and TET2 gain a proliferative advantage and clonal outgrowth, resulting in a condition called clonal hematopoiesis of indeterminate potential (CHIP). Additional mutations (e.g., NPM1, FLT3-ITD) in this clone transform the pre-leukemic HSC to leukemic stem cell (LSC). A subset of LSCs is therapy resistant thus resulting in cancer reappearance (relapse). Heterogeneity of the subclones contribute to LSC diversity and subsequently therapy resistance. Modified from Trumpp and Haas, 2022[33] using Biorender.com.

1.2.1 AML Epidemiology

In adults, AML is the most common type of leukemia. Although, in comparison to other cancer types its occurrence is less frequent. AML is slightly more frequently diagnosed in males and is considered a disease of the elderly, with a median age of

diagnosis of 70 years[39]. According to United Kingdom (UK) data, the incidence rate rises gradually starting at 40 years of age and follows a steeper increase at 60 years (**Figure 6**)[32].

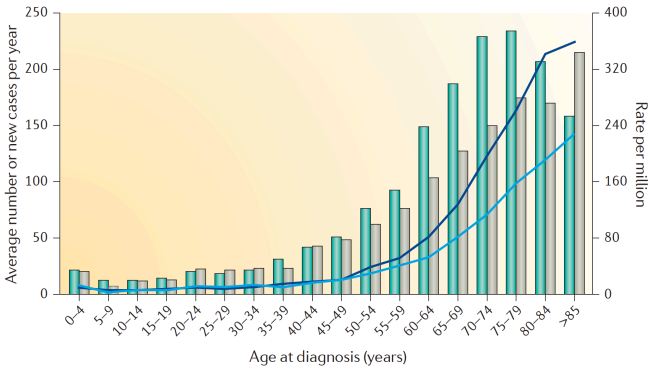

Figure 6 AML incidence trend by sex and age.
Figure from Khwaja et al. 2016[32] generated with data provided from Cancer Research UK.

While acute leukemias are the most common pediatric cancers, AML is less frequent than ALL, as it only accounts for approximately 20% of childhood leukemias[40]. There is no strong evidence regarding family predisposition for AML or MDS, however recent studies report germline genetic predisposition candidate genes in familiar AML/MDS, like DHX34 RNA helicase genetic variants[41,42]. In addition, relatives of patients with AML display an increased risk of hematological malignancies and solid tumours, potentially linked to exposure to environmental factors or shared genetic background which affects genes associated with malignancy[43].

Even though environmental predisposing factors in AML have not been defined, AML and MDS incidence rate is slightly higher in cases of subjection to DNA-damaging agents, including ionisation radiation and cytotoxic chemotherapy, a case of leukemia termed as therapy-related AML[40].

1.2.2 AML subtypes

AML diagnosis initially involves the accumulation of myeloblasts in the BM and blood. To distinguish the different AML subtypes, molecular characterisation and cytogenetics on the myeloblasts are used. In recent years, major advances in AML classification have been made, including genomic diagnostics and molecular markers. Due to that, the European LeukemiaNet (ELN) released an update of these criteria in 2022[44] (**Table 1**).

Table 1: European LeukemiaNet (ELN) of acute myeloid leukemia subtypes.
AML: Acute myeloid leukemia, APL: Acute promyelocytic leukemia, MDS: Myelodysplastic syndrome. Table provided from Döhner et al. 2022(Döhner et al. 2022).

AML with recurrent genetic abnormalities (requiring ≥10% blasts in BM or PB)
• APL with t(15;17)(q24.1;q21.2)/*PML*::*RAR*A
• AML with t(8;21)(q22;q22.1)/*RUNX1*::*RUNX1T1*
• AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22)/*CBFB*::*MYH11*
• AML with t(9;11)(p21.3;q23.3)/*MLLT3*::*KMT2A*
• AML with t(6;9)(p22.3;q34.1)/*DEK*::*NUP214*
• AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2)/*GATA2*, *MECOM(EVI1)*
• AML with other rare recurring translocations
• AML with mutated *NPM1*
• AML with in-frame bZIP mutated *CEBPA*
• AML with t(9;22)(q34.1;q11.2)/*BCR*::*ABL1*
Categories designated AML (if ≥20% blasts in BM or PB) or MDS/AML (if 10-19% blasts in BM or PB)
• AML with mutated *TP53*
• AML with myelodysplasia-related gene mutations (mutations in *ASXL1*, *BCOR*, *EZH2*, *RUNX1*, *SF3B1*, *SRSF2*, *STAG2*, *U2AF1*, and/or *ZRSR2)*
• AML with myelodysplasia-related cytogenetic abnormalities
• AML not otherwise specified (AML-NOS)
Myeloid sarcomas
Acute leukemia of ambiguous lineage
Myeloid proliferations related to Down syndrome
Blastic plasmacytoid dendritic cell neoplasm

1.3 Bone marrow remodeling in AML

AML has been found to interact with the BM in a dual manner. Alterations in the BM niche can affect leukemogenesis but also leukemic cells can alter the BM microenvironment, potentially leading to the accelerated expansion of leukemia[45].

AML is marked by the reduction of the vascular niche, through the induced degradation of the endosteal vasculature as well as the increased permeability of the vascular niche[46, 47]. AML disrupts the adipocytic lineage of the BM which further affects the erythroid differentiation of HSPCs[48]. Moreover, blasts secrete exosomes which remodel the niche into a leukemia-permissive ecosystem while suppressing normal hematopoiesis[49].

Phenotypic changes to T cells, which facilitate anticancer immunity, are also impacted by AML and contribute towards niche remodelling. AML-induced expansion of Tregs and Th17 cells has been connected to immune suppression, thus supporting immune evasion over cytotoxic anti-tumour activity of the immune system[50,51]. In line,

memory T cells of AML patients relapsing after alloSCT were shown to display increased exhaustion markers[52].

1.4 AML therapy strategies

Due to the heterogeneity of AML disorders, different types of treatment strategies can be applied, including chemotherapy, targeted therapies, chimeric antigen receptor T (CAR-T) cell treatment and stem cell transplantations (SCT)[53]. SCT can be further subdivided into autologous or allogeneic stem cell transplantation (alloSCT), in which the HSCs are derived from the patient or a human leukocyte antigen (HLA) matching donor, respectively[53].

In younger patients (60-65 years), the standard care of the initial intensive chemotherapy (induction therapy) is called "3 + 7 regimen"[54]. Briefly, the patient is administered intravenously three days of anthracyclines (daunorubicin, idarubicin) and seven days of continuously infused cytarabine. Follow up consolidation therapy aims to eliminate remnant cancer cells and can be multiple courses of chemotherapy with cytarabine and anthracyclines. Ongoing research introduced additional induction-consolidation AML agents, including epigenetic therapies using hypomethylating agents (azacitidine, decitabine), venetoclax in older patients, targeted therapies such as FLT3 and IDH inhibitors and addition of CD33-targeted monoclonal antibodies[54].

AlloSCT is an accepted standard therapy after first complete remission, and continues to be the main curative strategy. Following alloSCT, donor HSPCs repopulate the BM to regenerate the hematopoietic system, while donor T cells already present in the graft are capable of recognizing and eliminating the patients´ healthy cells, often leading to graft-versus-host disease (GVHD) as well as residual leukemic cells known as graft-versus-leukemia (GVL) effect (**Figure 7**)[55]. These two phenomena are partially connected, since mild chronic GVHD not requiring restart of immunosuppression is associated with better outcome compared to absence of chronic GVHD [56].

The selection of a suitable donor for SCT is critical for the outcome of alloSCT. Multiple factors are critical for the outcome of alloSCT, including donor-recipient histocompatibility, stem cell source (peripheral blood, bone marrow), donor-recipient cytomegalovirus (CMV) and blood type compatibility[32].

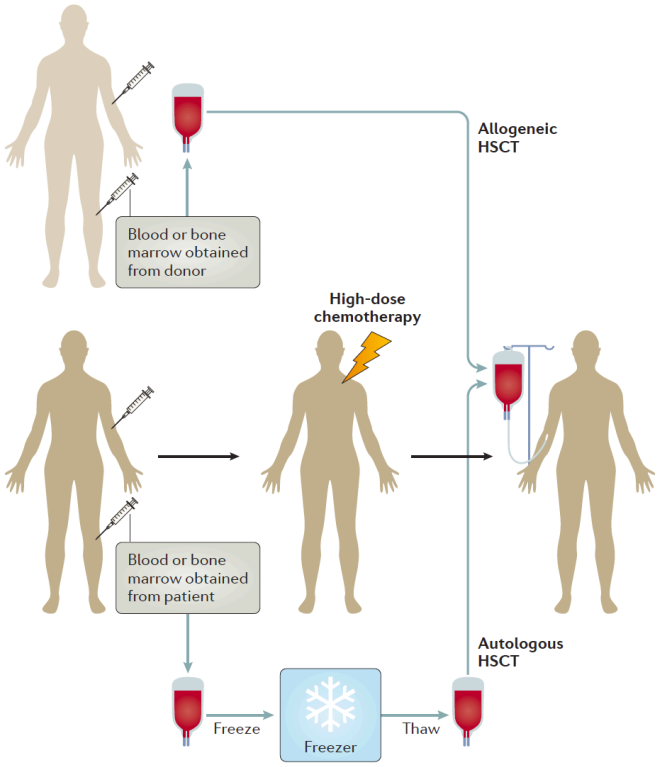

Figure 7 Hematopoietic stem cell transplantation approaches.
Allogeneic hematopoietic stem cell transplantation (alloSCT) is a therapeutic strategy in which an AML patient undergoes after conditioning with cytotoxic and immunosuppressive therapy. Specifically, donor stem cells and leukocytes either from peripheral blood or bone marrow are infused into the recipient. The donor is an antigen matched sibling or unrelated adult donor. In autologous stem cell transplantation, the patient's stem cells are harvested and frozen to be thawed and after high-dose cytotoxic therapy are re-infused to the patient in order to enable hematopoietic recovery. Source: Khwaja et al. 2016 [32].

After SCT, the recovery of the different immune subsets varies, with the innate immunity cells recovering early while T and B cells may take up to 2 years to fully reconstitute (**Figure 8**)[57]. Even after complete reconstitution, T cells often have an abnormal TCR repertoire, thus remaining dysfunctional[58]. Additional post-transplantation factors like administration of immunosuppressive therapies contribute to this delayed recovery of the immune system[57]. This delay is directly linked to the increased risk of relapse[57].

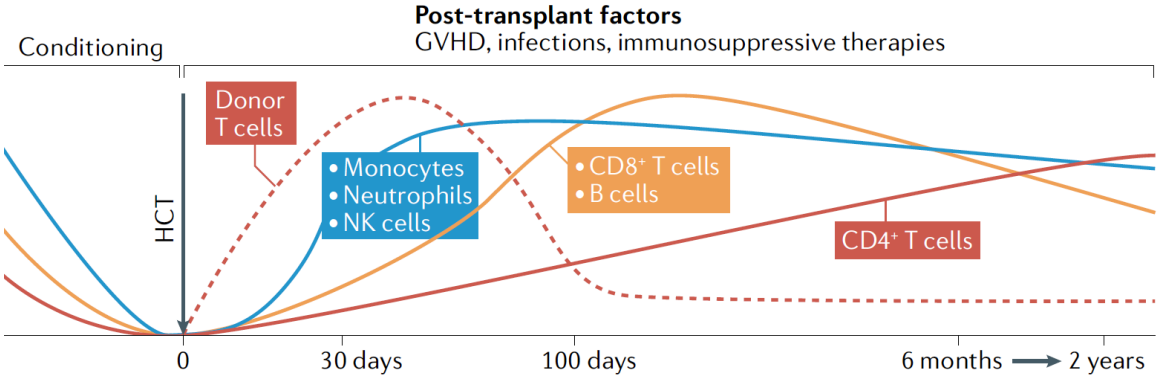

Figure 8 Overview of the T cell reconstitution after hematopoietic stem cell transplantation (HSCT).
Following HSCT, innate immune cells follow early recovery, while T and B cells may take up to 2 years to fully recover. Source: Verardi et al. 2021[57].

1.5 Model systems to study human hematopoiesis and AML

Despite the advances in methods for studying rare populations at a single cell resolution, the study of human HSPCs and their malignant transformation still remains challenging. This is a consequence of lack of appropriate *in vivo* and *ex vivo* models of the BM microenvironment. As previously outlined in this thesis, the BM microenvironment is characterized by increased complexity, and a vast variety of hematopoietic and non-hematopoietic subsets.

Dick and colleagues pioneered the first humanized mouse models, which still allow researchers the study of human HSPCs function *in vivo*[59]. Humanized mice have advanced our understanding of the mechanisms underlying human hematopoiesis. These models, also known as xenograft models, carry mutations on genes which facilitate the increased engraftment of human HSCs and AML cells through inhibiting the organism's immunity. They are often homozygous for the loss of function mutation in protein kinase DNA-activated catalytic subunit (PRKDC) gene, referred to as severe combined immunodeficiency (*scid*) or for targeted mutations at the recombination-activating gene 1 (*Rag1*) or *Rag2*, together with a targeted mutation at the interleukin-2 receptor γ-chain (*Il2rg*) locus[60].

A widely used strain for studying hematopoiesis is NSG mice of NOD genetic background, which carries also a mutation of Il2rg (NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ)[61]. NSG mice lack mature lymphocytes and allow engraftment of human HSPCs[61]. While NSG mice require prior irradiation in order to achieve sufficient engraftment, NSGW41 mice support engraftment of HSPCs without any prior conditioning. NSGW41 are NSG mice with the additional *Kit W-41J* allele[62]. Kit encodes for c-Kit, which is the stem cell factor (SCF) receptor[63]. c-Kit is expressed on hematopoietic cells, and when it binds with

SCF it causes signal transduction necessary for normal hematopoiesis[64]. Consequently, this genetic background is beneficial for the engraftment of donor's HSPCs over the host, without the necessity for prior irradiation[62].

Although xenograft models have significantly advanced our understanding towards the concept of hematopoiesis, these models still impose the limitation of a non-human microenvironment. T. M. Dexter initially developed stromal cocultures as a way of maintaining HSPCs[65]. Since then, *in vitro* co-cultures of human derived cells, often provide mechanistic insights into the multicellular interplay. Such studies, in which primary AML cells were co-cultured with human BM derived stromal cells, revealed the importance of the stromal microenvironment in promoting AML survival[66].

While co-culture models are a well-defined setup to study HSPCs and the interactions with the microenvironment, they still consist of a simplified snapshot of a complex tissue like the BM. A recent study from the Psaila group established 3D human BM organoids, generated from induced pluripotent stem cells (iPSCs) committed to hematopoietic, mesenchymal and endothelial lineages[67]. These structures represent key features of the bone marrow, like lumen-forming sinusoids, sufficient differentiation of myeloid cells as well as BM stroma[67].

1.6 Single cell technologies and the study of rare cell types

Single-cell technologies, such as single-cell RNA sequencing, have revolutionized the study of rare cell types in the BM and other tissues, since they allow measurement of thousands of individual cells. These methods are being used to answer a range of clinically relevant questions in the context of hematological malignancies, like the identification of molecular pathways associated with therapy resistance in multiple myeloma, a neoplastic plasma cell disorder[68].

Well established technologies like flow cytometry and fluorescence activated cell sorting (FACS) have been facilitating the precise isolation and analysis of selected single cells, with the benefit of simultaneous profiling of surface proteins[69]. Recent advances in profiling more cellular modalities have allowed scientists to discover novel cell types, states as well as providing insights into the transitions from one cell to another, during biological processes like development, differentiation and malignant transformation[70]. In 2009, the first single cell transcriptomics (single cell RNA sequencing, scRNA-seq) study was published[71]. Since then, there has been a growing research interest for the improvement of those methods and the incorporation of additional modalities (**Figure 9**). Such methods profile genome sequences[72], chromatin accessibility[73,74], DNA

methylation[75], cell surface proteins[76,77] (**Figure 9**). In the case of cell surface proteins, Peterson and colleagues simultaneously measured gene and protein expression (Cellular Indexing of Transcriptomes and Epitopes by sequencing, CITE-seq)[76]. Another example of simultaneous measurements of multiple modalities is single cell genotyping of transcriptomes (GoT) which integrates genotyping alongside high-throughput scRNA-seq[78]. This method is especially useful in the case of malignant cells, due to the absence of surface markers to distinguish cancer clones and genotype is the only discriminant[78].

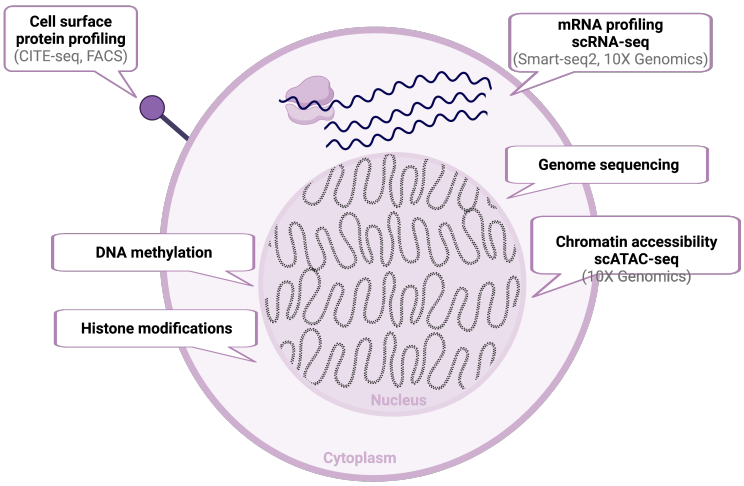

Figure 9 Overview of the current readouts for multimodal analysis of single cells.
CITE-seq: Cellular Indexing of Transcriptomes and Epitopes by sequencing, FACS: Fluorescence-Activated Cell Sorting, mRNA: messenger RNA, scRNA-seq: Single-Cell RNA sequencing, scATAC-seq: Single-Cell Assay for Transposase-Accessible Chromatin using sequencing. Created using Biorender.com.

Despite the developments in single cell methods, issues related to batch effects, detection of cell multiplets as well as the reduction of experiment costs persist. These challenges are often addressed by pooling multiple samples prior to data collection, that their origin can be determined computationally. This can be accomplished by the use of cell hashing barcoded antibodies[79], or DNA oligonucleotide tagging of cellular proteins[80]. In the case of human samples, natural genetic variation can be utilised in order to distinguish cells from multiple individuals[81,82].

1.7 Aims and scopes of this thesis

PhD thesis significance

The previous sections have highlighted the importance of studying the complex interactions between malignant transformation of hematopoietic cells and their tumor microenvironment in the development of acute myeloid leukemia (AML). AML affects thousands of individuals yearly and is characterized by the abnormal growth of immature myeloid cells in the bone marrow (BM).

The BM is a complex organ which consists of two main components, the immune and the non-hematopoietic component. After initiation, one of the key drivers in AML progression is its interaction with both of these components. In malignancy, the BM microenvironment is highly remodelled. The question of whether the microenvironment is remodeled towards a favorable environment for AML maintenance prior to or following malignant transformation of cells remains a "*causality dilemma"* for cancer research.

The high abundance of malignant cells presents a challenge in profiling their microenvironment. However, recent advancements in single-cell technologies have facilitated the high-resolution mapping of rare cell populations.

PhD thesis aims

The overarching goal of my PhD thesis has been to address the long-standing question of the role of the tumour microenvironment in AML, focusing on these two components of the BM-the immune and stroma.

To reach this objective, I pursued two primary goals:

1. First, I sought to determine **the role of the BM immune landscape in therapy outcome post allogeneic stem cell transplantation (alloSCT)**.

   After alloSCT, donor T cells target and eliminate AML blasts. Yet, 50 % of patients still relapse[83]. Poor therapy outcome is hypothesised to be driven by impaired anticancer immunity of the donor T cells.

   To address this hypothesis, I first performed scRNA-seq analysis of patient derived T cells after alloSCT. My analysis identified several T cell signatures associated with favourable outcomes following alloSCT, including enrichment of G protein–coupled receptor 56 (GPR56) expression in the $CD8^+$ T cells of patients in remission. Then, in collaboration with the Schmitt group at Universitätsklinikum Heidelberg it was demonstrated that GPR56 becomes specifically upregulated on T cells upon antigen encounter with AML cells. Lastly, these findings were further expanded in an independent cohort of 139 AML patients profiled using flow cytometry. The results of this first project indicated

that surface expression of GPR56 may be used as an easily detectable biomarker for alloreactivity and therapy response for patients receiving alloSCT.

2. In light of the effect of AML in the stromal microenvironment and the complex interaction of various cell types in the BM, I then pursued to uncover **the impact of leukemic stem cells (LSCs) on the BM stromal microenvironment.**
   The motivation for this second project was that LSCs often escape classic therapy strategies, resulting in relapse. Therefore, mapping their impact on their microenvironment is instrumental for understanding the mechanisms behind therapy evasion.
   Initially, I investigated the impact of LSC_{high} and LSC_{low} burden AML to the BM stromal microenvironment using scRNA-seq of NSGW41 xenograft mouse models. Since the BM ecosystem of these murine models was never extensively profiled, my first goal was to do a comparative study comparing the BM niche of the NSGW41 strain to other well-studied strains. After achieving this, I performed differential expression, transcription factor activity, cell composition and cell-cell interaction analyses of the stromal cells in LSC_{high} versus low burden xenograft models. These multifaceted analyses revealed that LSC burden is associated with an imbalance of the adipo-lineage over the osteo-lineage as well as fibrotic signatures. Since the major caveat of the aforementioned setup had been to study the impact of human AML on mouse stroma, I sought for further evidence, by profiling *in vitro* co-cultures of human BM derived MSCs with LSC_{high} and LSC_{low} burden AML.

CHAPTER 2: RESULTS

2.1 Project 1: The remission status of AML patients post alloSCT is associated with a distinct single cell signature of bone marrow T cells

The results presented in Section 2.1 are part of a manuscript currently in preparation with title "The remission status of AML patients post alloSCT is associated with a distinct single-cell bone marrow T cell signature".

List of authors: Anna Mathioudaki, Xizhe Wang, David Sedloev, Richard Huth, Aryan Kamal, Michael Hundemer, Yi Liu, Spyridoula Vasileiou, Premal Lulla, Carsten Müller-Tidow, Peter Dreger, Thomas Luft, Tim Sauer, Michael Schmitt, Judith B. Zaugg, Caroline Pabst

With the exceptions detailed below, the experiments and analysis presented in this Chapter were designed and performed by me, in discussion and guidance from **Dr. Judith Zaugg** and **Dr. med. Caroline Pabst**.

In detail, I contributed to the experimental design, conducted experiments (scRNA-seq experiments, library preparations, flow cytometry), performed scRNA-seq data analysis as well as clinical data analysis, interpreted the data and generated figures.

scRNA-seq experiments were performed jointly with **Xizhe Wang.** In addition, Xizhe Wang performed flow cytometry experiments and assisted with cell type annotation. **David Sedloev**, performed the CAR-T cell experiments. HL60 CD33 KO cell lines and 3G.CD33.CAR construct was provided by **Dr. Yi Liu**. **Richard Huth** assisted with the CAR-T cell experiments and the flow cytometry experiments. **Clarissa Holitsch, Rebeca Kruhmann** and **Sophie Leonhardt** assisted with sample collection.

2.1.1 Summary

The focus of this Chapter is to investigate the role of the bone marrow (BM) T cell landscape in the therapy outcome of acute myeloid leukemia patients (AML) after allogeneic stem cell transplantation (alloSCT). AlloSCT is the most established immunotherapy in AML, where donor T cells present in the graft can recognise and target residual leukemic cells. A holistic view of the bone marrow T cell repertoire is presented to distinguish T cells from relapse patients (REL) in whom early graft versus leukemia (GvL) failed, versus T cells in patients with long-term complete remission (CR). This Chapter pinpoints several T cell signatures associated with favourable outcomes, including enrichment of specific $CD8^+$ T cell subsets and increased T cell cytotoxicity. On the other hand, relapse patients were associated with certain $CD4^+$ T cell subsets, such as regulatory T cells as well as TNF/NF-κB signalling. Amongst these signatures, the adhesion G-protein coupled receptor 56 (GPR56) is proposed as a dynamic biomarker in T cell reactivity after transplantation and validated using a chimeric antigen receptor T (CAR-T)/HL-60 coculture system and multi-colour flow cytometry on a cohort of more than 100 alloSCT bone marrow samples. A summary of these findings is illustrated in **Figure 10**.

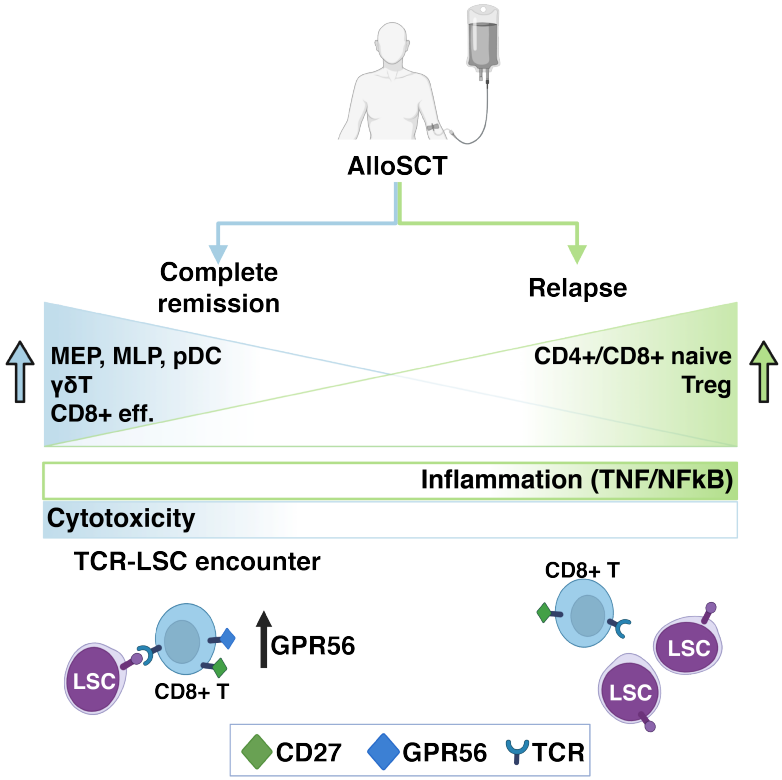

Figure 10 Graphical summary of the findings presented in this Section.
MEP: megakaryocyte-erythroid progenitors; MLP: myeloid/lymphoid progenitors; pDC: plasmatocytoid dendritic cells; Treg: Regulatory T cells; TCR: T cell receptor; ; LSC: Leukemic stem cell.

2.1.2 Experimental overview and quality control of scRNA-seq data

Primary BM samples of 3 Complete Remission (CR) and 3 Relapsed (REL) patients were thawed (sample characteristics in **Table 4**), stained with CD3 and CD34 (common T cell and hematopoietic stem/progenitor cells (HSPCs) markers, respectively). Then, T cells and HSPCs were sorted, and processed with the 10X platform to produce single cell RNA-sequencing (scRNA-seq) libraries. Overview of the gating scheme is presented in **Figure 52** (Section 4.2). Per group, all 3 patient cells were mixed and pooled together into one 10X reaction (**Figure 11**).

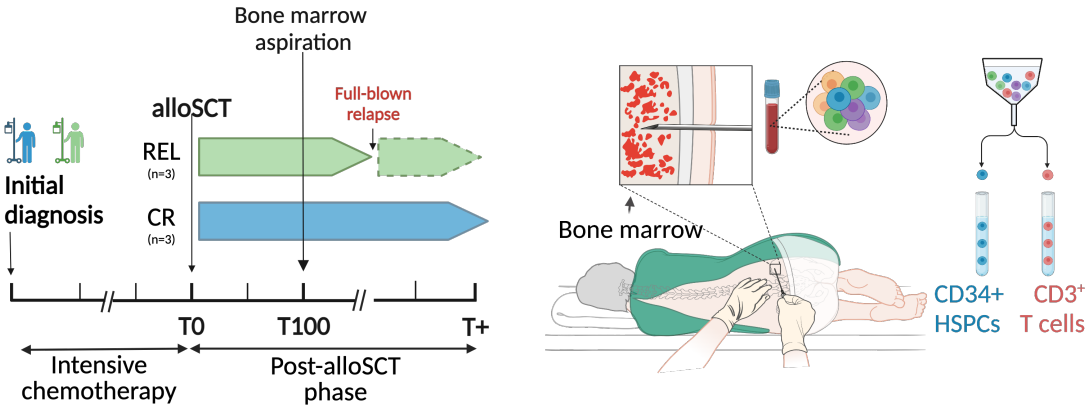

Figure 11 Overview of the experimental design to characterize the bone marrow (BM) T cell and hematopoietic stem and progenitor cells (HSPCs) landscape of AML patients after allogenic stem cell transplantation (alloSCT). CR: Complete Remission; REL: Relapse.

Following quality control (QC; cells with more than 200 genes detected and less than 15% mitochondrial genes) 4,038 from 3 CR & 4,461 cells from 3 REL patients were retained for downstream analysis (average number of cells per patient = 1,414; **Figure 12**). Using souporcell[81], I assigned each single cell to individual samples (see Methods, Section 4.13.4). Briefly, souporcell is a method for clustering single cells of mixed-genotypes, based on common variants from the 1000 genomes project, filtered for variants of minor allele frequency greater than $5x10^{-4}$ [84].

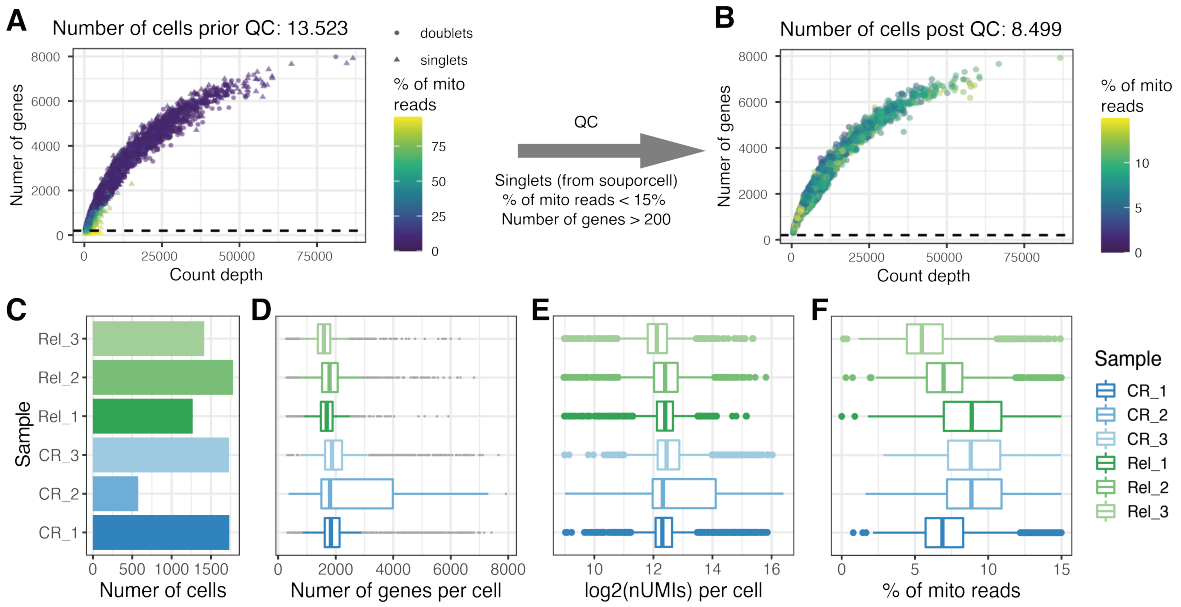

Figure 12 Quality control of scRNA-seq data.
Scatter plot that each dot indicates a single cell, count depth (x-axis) plotted against the number of genes. The colour indicates the fraction of mitochondrial reads prior to quality control (QC) filtering (A) and post QC (B). (C) Number of cells per sample. (D) Number of genes per cell per sample. (E) Number of unique molecular identifiers (UMIs) in a log2 scale. (F) Percentage of mitochondrial reads per single cell. CR: Complete remission, REL: Relapse.

Since no prior genotypic information per sample was acquired, further assignment of samples to individual patients was performed using sex chromosome genes, *XIST* as an X-linked gene and *RPS4Y1* as a Y linked gene. This analysis revealed that souporcell clusters/samples REL_2 and CR_3 were female while the rest are male (**Figure 13**).

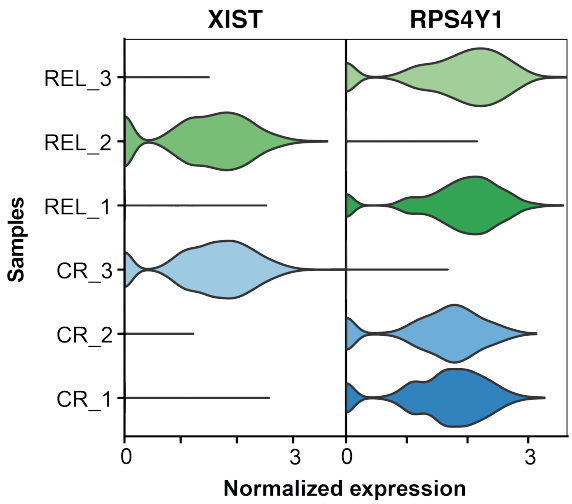

Figure 13 Normalised expression of X-linked gene XIST and Y-linked gene RPS4Y1. CR: Complete remission, REL: Relapse.

2.1.3 The bone marrow landscape of post-transplant AML patients

After Louvain clustering, the single cells were grouped into separate clusters and annotated based on known marker genes into nine HSPCs, eight CD8$^+$, five CD4$^+$ and two unconventional T cell clusters. Cluster annotation was performed by incorporating both gene expression as well as TF activity levels, which was calculated using *SCENIC*[85].

HSPCs annotation

Among the HSPCs, I identified nine clusters including myeloid/lymphoid progenitors (MLP) expressing *SPINK2* and *SELL*; precursors of B cells (preB) expressing *VPREB1/3, DNTT, JCHAIN, CD79A* and *CD24* as well as precursors of T cells (pro/preT) expressing *CD3G, CD3D* as well as *CD4* and *CD8A/B* simultaneously. Moreover, neutrophil progenitors (NP) expressing higher level of *ELANE, AZU1, PRTN3, CTSG* and *RNASE2* and megakaryocyte-erythroid progenitors (MEP) expressing *AHSP, HBB, CA1, KLF1, ITGA2B, GATA2, CNRIP1* and *MYC* expressing were detected. Lastly, monocyte-dendritic progenitors (MDP) expressing *SCT, IRF8, IRF7* and *TCF4* as well as plasmacytoid and conventional dendritic cells (pDCs/cDCs) (*IRF7, CLEC4C*) and monocytic precursors/monocytes (MP/mono) expressing *CD14r* were detected (**Figure 14, 15**). Notably, MDP, MP/mono, DC depicted low or no expression of *CD34* at the level of RNA/gene expression. The discrepancy between RNA and protein levels for *CD34* has been previously observed, for the same HSPCs populations, in recent CITE-seq (Cellular Indexing of Transcriptomes and Epitopes by sequencing) studies, potentially associated with post-transcriptional RNA regulation [86,87]. The annotation of the aforementioned subsets was performed according to the findings from Velten et al. 2017[88] (**Figure 14**).

T cells annotation

The dataset presented here spans across all known T cell populations, several of which are understudied in the context of AML and therapy outcome (**Figure 15**).

In particular, CD8$^+$ T cells were sub categorised into 13 clusters including naïve (CD8$^+$ NV) expressing *CCR7* and an IFN response (CD8$^+$ IFN) cluster with shared features with corresponding CD4$^+$ clusters, which will be further described below. In addition, I identified a memory-like cluster termed CD8$^+$ hobit, characterised by high expression of Hobit (*ZNF683*), which is associated with long-lived effector memory cells[89].

The remaining CD8$^+$ populations comprised various effector clusters expressing cytotoxicity genes *GZMB, FGFBP2, GZMH, CX3CR1, PRF1* and *GNLY* as well as exhaustion genes *TIGIT1, CD160* and *CTLA-4* at different levels[90]; memory related cells expressing *GZMK* and different levels of exhaustion genes, though with low expression of cytotoxicity genes. In particular, I identified 2 effector subsets termed CD8$^+$ effector

1 and 2 (CD8$^+$ eff. 1: *CD160*$^+$, CD8+ eff. 2: *CD160*; expressing *NKG7* and *GZMB*), CD8$^+$ memory 1, 2, and 3, characterised by expression of *GZMK* and lower levels of cytotoxic genes (**Figure 14, 15**).

Additionally, I identified unconventional T cells including Gamma-delta (γδ) T cells with high expression of *TRDC*, *TRGC1* and *KLRC1*[91] as well as mucosal associated invariant T cells (MAIT) cells expressing *KLRB1, SLC4A10, NCR3*. The latter are hypothesised to temporarily reside in the BM due to circulation. Recent evidence highlights the role of MAIT cells in PD1 therapy outcome in AML, however these conclusions were derived from a limited number of patients[92].

The CD4$^+$ cells were further segregated into 5 clusters, including CD4$^+$ T cells (CD4 T NV) expressing *CCR7*, regulatory T cells (Treg) expressing *FOXP3*, CD4$^+$ IFN expressing high levels of IFN response genes, Th17 expressing *LMNA* and *CCR10*[93] and a cluster of CD4$^+$ T memory (CD4 T mem) cells. Notably, the identification of CD4$^+$ effector subsets was challenging and since they did not correspond to any known subsets. This could potentially be explained by the slower CD4$^+$ T cells reconstitution after alloSCT[57], since the donor T cells analysed in this study were acquired only 100 days after alloSCT (**Figure 14, 15**).

In order to further corroborate accurate cluster annotation, TF activity inferred with *SCENIC*[85] was used as an additional level of information. MP/mono showed co-activity of SPI1, EGR1, and IRF8[94], naive CD4 and CD8 T cells had highest co-activity of LEF1 and TCF7[95], while Tregs showed the typical FOXP3 activity[96,97] (**Figure 14**).

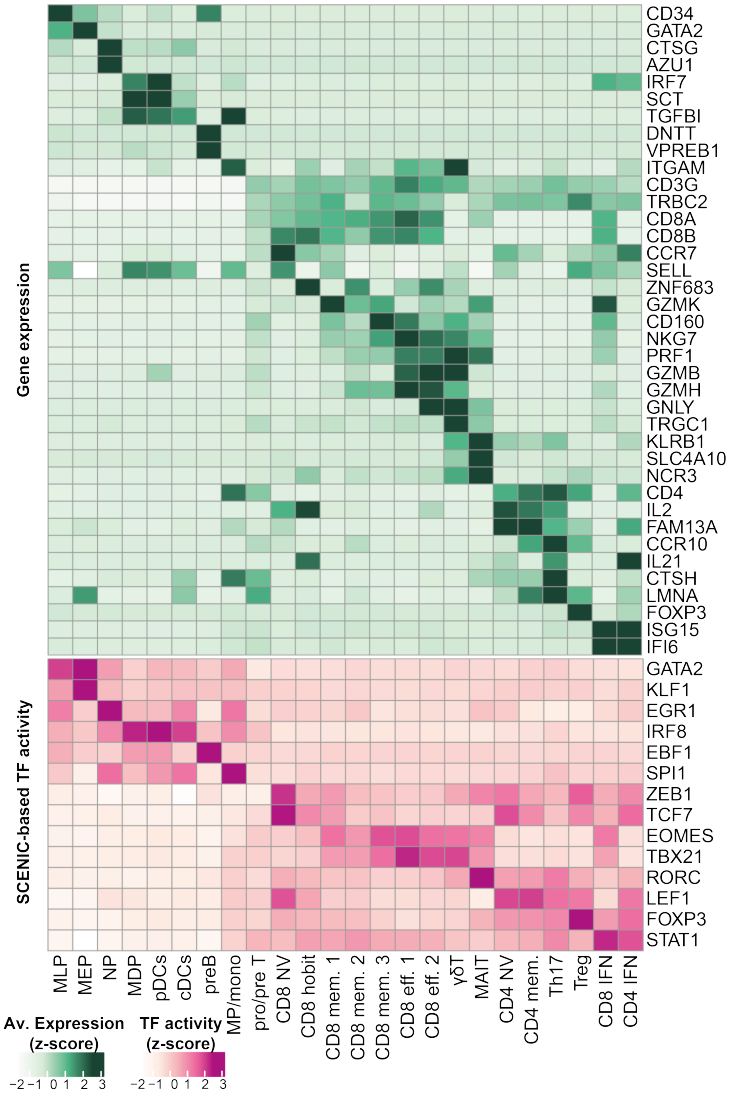

Figure 14 Marker genes and transcription factors (TFs) of BM resident T cells and HSPCs. Heatmap depicting scaled expression of marker genes (green) and scaled TF activity of marker TFs (pink; calculated using SCENIC[85]) across all 24 unsupervised clusters. Values are averaged across all cells in the cluster. The genes and TFs represented contributed to the annotation of the cell types.

Overall, this dataset covers a variety of HSPCs and T cell subsets (**Figure 15**). Here, clusters were annotated based on RNA features. T cells are traditionally studied and defined by surface molecules at a protein level, often using FACS. With the rise of CITE-seq, which allows simultaneous profiling of RNA and protein, scRNA-seq studies began to accommodate similar terminology[86]. This point will be addressed in Section 2.1.5, where I computationally aligned my dataset with previously published CITE-seq datasets.

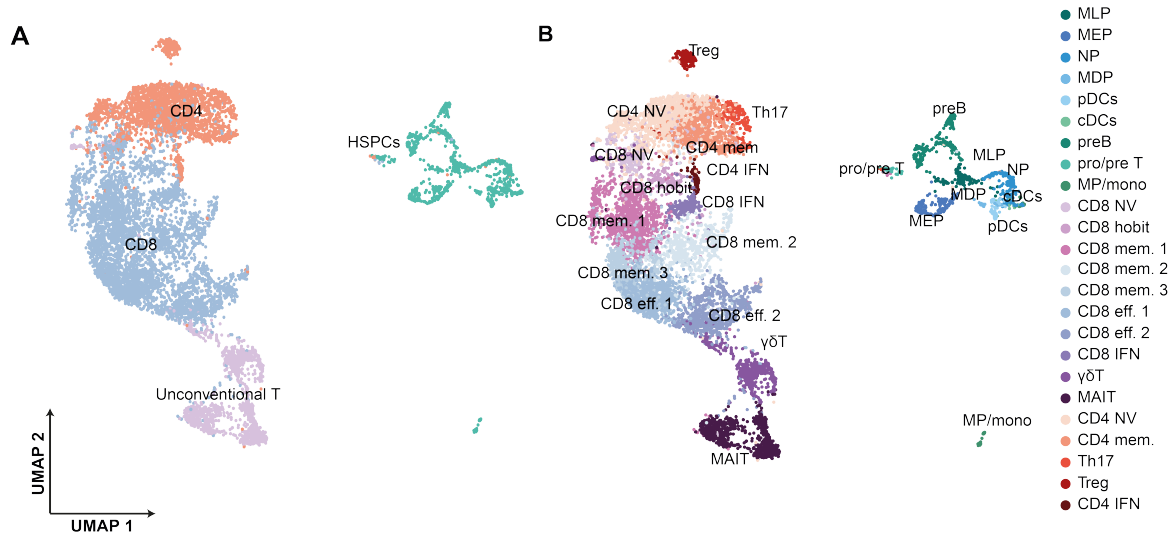

Figure 15 The landscape of BM resident T cells and HSPCs after alloSCT.
Uniform manifold approximation and projection (UMAP) of 8,492 post-QC cells representing BM HSPCs and T cells from 6 AML patients. Cells are coloured according to the four major populations (A) or the twenty-four detailed cell types/states (B).

2.1.4 Bone marrow composition and therapy outcome

Since the main objective of this project was to investigate whether changes in the T cell landscape could impact the therapy outcome after alloSCT, and considering that the T cell reconstitution plays a crucial role in anticancer immunity, I initially investigated whether there was a association between the BM composition and the alloSCT outcome.

While all cell types were present in all patients, their abundances differed between the two groups (**Figure 16A**). Relapse specimens showed depletion of the most immature, multipotent MLPs and of MEPs (**Figure 16B;** Fisher's test, log2OR < 0, p.adj < 0.05 after Bonferroni correction). Depletion of normal erythroblasts in AML is a known feature, due to the accumulation of myeloid blasts stalling normal production of myeloid progenitors and mature cells. Importantly, even though the samples included in this study were collected prior to relapse (<5% blasts in the BM, sample characteristics in **Table 4**), the presence of a few AML blasts cannot be entirely ruled out. In addition, in REL samples proT/preT and pre-B cells were enriched indicating a normal hematopoietic maturation towards lymphoid lineages (**Figure 16B;** Fisher's test, log2OR < 0, p.adj < 0.05 after Bonferroni correction). CR samples were significantly enriched for pDCs, known to stimulate immune responses against AML (**Figure 16B;** Fisher's test, log2OR > 0, p.adj < 0.05 after Bonferroni correction) [98].

Within the T cell compartment, all clusters were present in both conditions (**Figure 16C**). However, $CD4^+$ NV, Tregs, $CD8^+$ NV and $CD8^+$ hobit, and the two interferon

clusters (CD4⁺ IFN, CD8⁺ IFN) were enriched in REL patients (**Figure 16D;** Fisher's test, log2OR < 0, p.adj < 0.05 after Bonferroni correction) while CD8⁺ effector cells (CD8⁺ eff. 1 and 2), and mature CD8⁺ mem. 2 and 3 as well as γδT cells (**Figure 16D;** Fisher's test, log2OR > 0, p.adj < 0.05 after Bonferroni correction).

Overall, this analysis indicates a T cell maturation stall in REL, associated with the enrichment of naive T cells (CD4⁺ and CD8⁺ NV cells) as well as early memory populations (CD8⁺ hobit, CD8⁺ mem. 1). This cannot be associated with HSPCs' deficiency to differentiate towards the lymphoid lineages, since pro/preT cells were enriched in REL. Moreover, REL samples exhibited a depletion of highly cytotoxic CD8⁺ effector clusters, indicating that decreased killing capacity of T cells may be associated with AML re-appearance after alloSCT. The Treg enrichment in REL is in line with previous studies, linking Treg with poor outcome in AML, due to their immunosuppressive capacity[99].

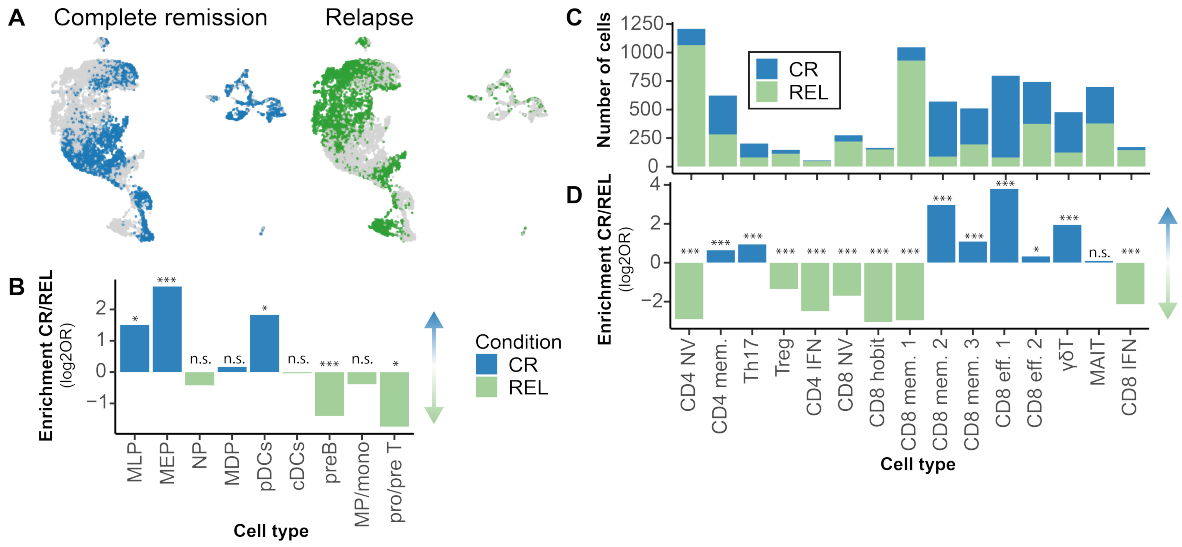

Figure 16 Bone marrow (BM) compositional changes in therapy outcome.
(A) UMAP highlighting complete remission (CR) and relapse (REL) cells (green: REL, blue: CR). (B) Differential abundance analysis per cell type within the hematopoietic stem and progenitor cells (HSPCs) using Fisher's exact test. The bars represent the log2 odds ratios calculated using Fisher's test (Fisher test; p-values were adjusted for multiple comparisons using the Bonferroni correction method), n.s.: not significant, *P < 0.05, **P < 0.001, ***P < 0.0001) (C) Absolute numbers of cells across CD3⁺ T cell types. The different colors indicate REL (green) and CR (blue) samples. (D) Differential abundance per cell type, within the CD3⁺ population using Fisher's exact test. The bars represent the log2 odds ratios (Fisher test, p-values were adjusted for multiple comparisons using the Bonferroni correction method, n.s.: not significant, *P < 0.05, **P < 0.001, ***P < 0.0001).

I then sought to further corroborate the findings related to CD8⁺ compositional alterations in AML and therapy outcome using an alternative approach, independent of clusters of cells. To do so, I performed pseudotime analysis using *Monocle3*[100] (See Methods, Section 4.13.7). Briefly, *Monocle3* is a computational approach which orders a given set of cells based on similarity, which may reflect a biological process. In the case of T cells, this process is T cell maturation. The first step is the definition of a starting population, which in this analysis was CD8⁺ NV. Unconventional T cells were excluded from the analysis.

The reconstructed pseudotime within the CD8⁺ T cells followed the expected sequence. Expression of genes associated with naive T cells (*IL7R, CCR7 LEF1*) was high in the beginning of the reconstructed trajectory (**Figure 17 B**). Next, the expression of early memory genes (*ZNF683*/HOBIT and *GZMK*) increased, followed up by CD8⁺ effector genes (*NKG7*, *KLRG1*, *GZMH*, *GZMB*) known to have a cytotoxic role (**Figure 17 B**). The expression of effector genes then declined when known exhaustion genes like *TIGIT* and *PDCD1* began to be expressed (**Figure 17 B**).

Consistent with the results from the Fisher's test, REL cells were significantly less advanced in pseudotime (**Figure 17 A, C**). The fraction of CD8⁺ REL cells increases again at the latest part of the trajectory, which is characterised by high expression of exhaustion markers. These observations suggest abnormal CD8⁺ T cell development in REL samples relative to CR, as well as rapid T cell exhaustion.

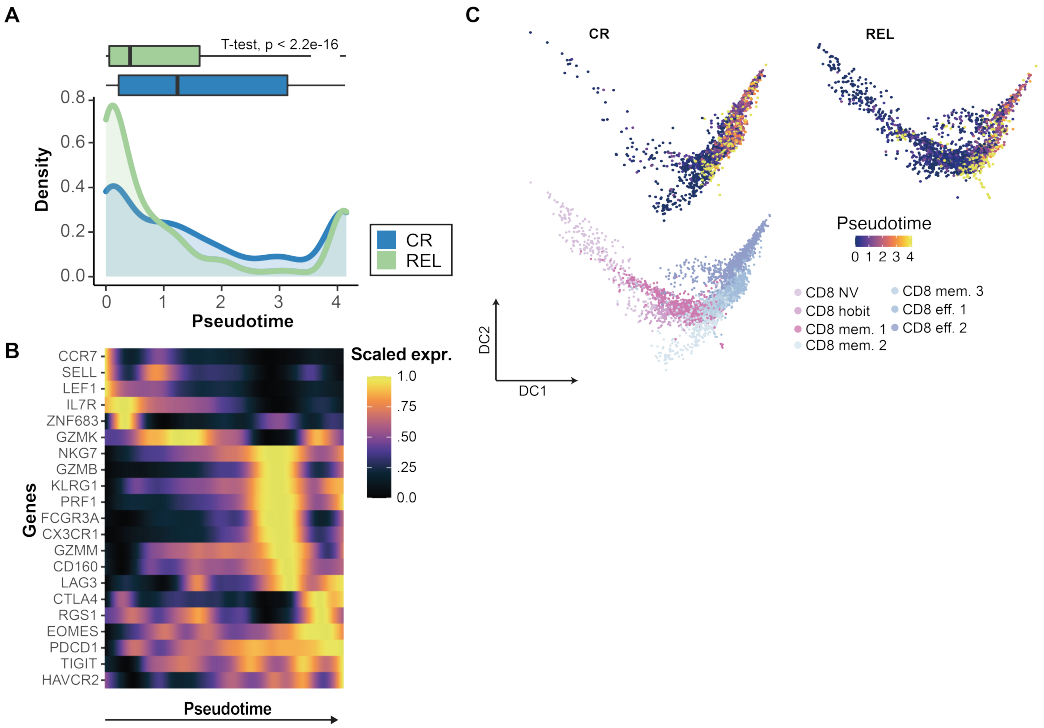

Figure 17 Pseudotime analysis of CD8⁺ EM cells.

(A) Boxplot (top) and density plot (bottom) depicting the pseudotime of CD8$^+$ effector memory (EM) cells in CR (blue) and REL (green). Comparison was performed using Student's t-test. (B) Heatmap depicting the scaled expression across pseudotime of selected effector, memory and exhaustion genes in CD8$^+$ cells. (C) Diffusion maps for the CD8$^+$ cells coloured according to the inferred pseudotime using Monocle3 split per condition (top) and based on the clusters (bottom).

2.1.5 Comparison with publicly available CITE-seq datasets

T cell subsets are classically defined by the presence of surface markers at protein level, while from scRNA-seq data we can derive information regarding gene expression profiles. In brief, T cells can be subdivided in four groups using the expression levels of two surface proteins; CCR7 and CD45-RA at a protein level[101]. These 2 proteins are sufficient to define the four subsets of T cells which mature from naive (CCR7$^+$ CD45RA$^+$) to central memory (TCM, CCR7$^+$ CD45RA$^-$), then to effector memory (TEM, CCR7$^-$ CD45RA$^-$), and lastly CD45RA$^+$ effector memory cells (TEMRA, CCR7$^-$ CD45RA$^+$)[101]. The latter are characterised by elevated expression of effector molecules, which in the case of CD8$^+$ cells, are related to cytotoxicity[101].

In this section, I further characterized the T cell subsets. For that, I integrated the dataset with a recent CITE-seq dataset of peripheral blood mononuclear cells (PBMCs) [102] and then performed *LabelTransfer* analysis[103]. In this study, T cell subsets were defined by both RNA and surface protein expression[103].

My analysis allowed the mapping of CD8$^+$ hobit and CD8$^+$ mem. 1 clusters, which were enriched in REL to CD45RA$^-$ CCR7$^+$ TCM and CD45RA$^-$ CCR7$^-$ TEM, respectively (**Figure 18**). On the other hand, CD8$^+$ clusters depleted in REL (CD8$^+$ eff. 1 and 2, and CD8$^+$ mem. 2 and 3), corresponded to different subsets of CD45RA$^+$ CCR7$^-$ effector memory cells (TEMRA), which differed in their expression levels for CD45RA and the cytotoxic molecules *GZMB*, *GZMK*, and *GNLY* (**Figure 18**). This analysis reinforces the hypothesis of impaired CD8$^+$ T cell maturity and cytotoxicity to be linked with alloSCT outcome.

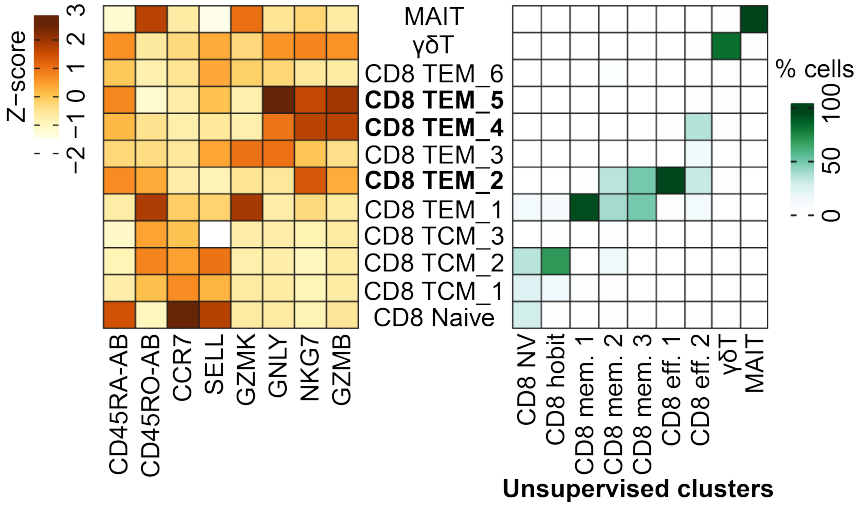

Figure 18 Correspondence of in house scRNA-seq CD8$^+$ T cells clusters with CITE-seq publicly available PBMCs clusters.
Left: Scaled expression (z-score) of publicly available CITE-seq data. Features (x-axis) with AB suffix indicate that the measurement was performed on protein level. Right: Percentage of cells per cluster (x-axis) that map to public reference clusters (y-axis). Bold cluster names indicate CD8$^+$ TEMRA subsets (CD45RA$^+$ CCR7$^-$ effector memory).

2.1.6 Identification of TFs associated with alloSCT outcome

In the previous section (Section 2.1.4), I identified BM T cells compositional changes to be associated with therapy outcome. T cells differentiate from naive to the different effector populations after activation and additional extrinsic signals, like chemokine polarisation. These signals trigger signalling cascades that subsequently induce expression of gene modules after activation of specific transcription factors (TFs). Following up on this, I then investigated whether these compositional changes may be reflected by distinct TFs.

To address this, I performed differential TF activity analysis using an adaptation of the *SCENIC* workflow[85]. *SCENIC*[85] is a computational tool which uses co-expression networks and motif discovery analysis in order to infer gene regulatory networks (GRN; see Methods, Section 4.13.5). In total, this GRN consisted of 280 TFs and 8,206 target genes (average: 100 target genes per TF). To identify differentially active TFs, I first detected the differentially expressed (DE) genes between the two conditions using MAST[104] (**Figure 19 A**). I found 422, 539, and 644 differentially expressed genes (DEG) for CD4$^+$, CD8$^+$, and HPSCs, respectively (log2FC > 1, p.adj < 0.05 after Bonferroni correction). Per TF and population, I then calculated the enrichment for these genes over the rest of its target genes using Fisher's exact test (**Figure 19 C, D, E**).

These results identified 12 TFs more active in REL HSPCs and CD8$^+$ T cells, and 10 TFs in CD4$^+$ REL T cells (**Figure 19 B**; p.adj < 0.05, after Bonferroni correction). In CR cells, 4 TFs were more active in HSPCs and only 1 in CD8$^+$, while in CD4$^+$ T cells no significant enrichment was detected (**Figure 19 B**; p.adj < 0.05, after Bonferroni correction).

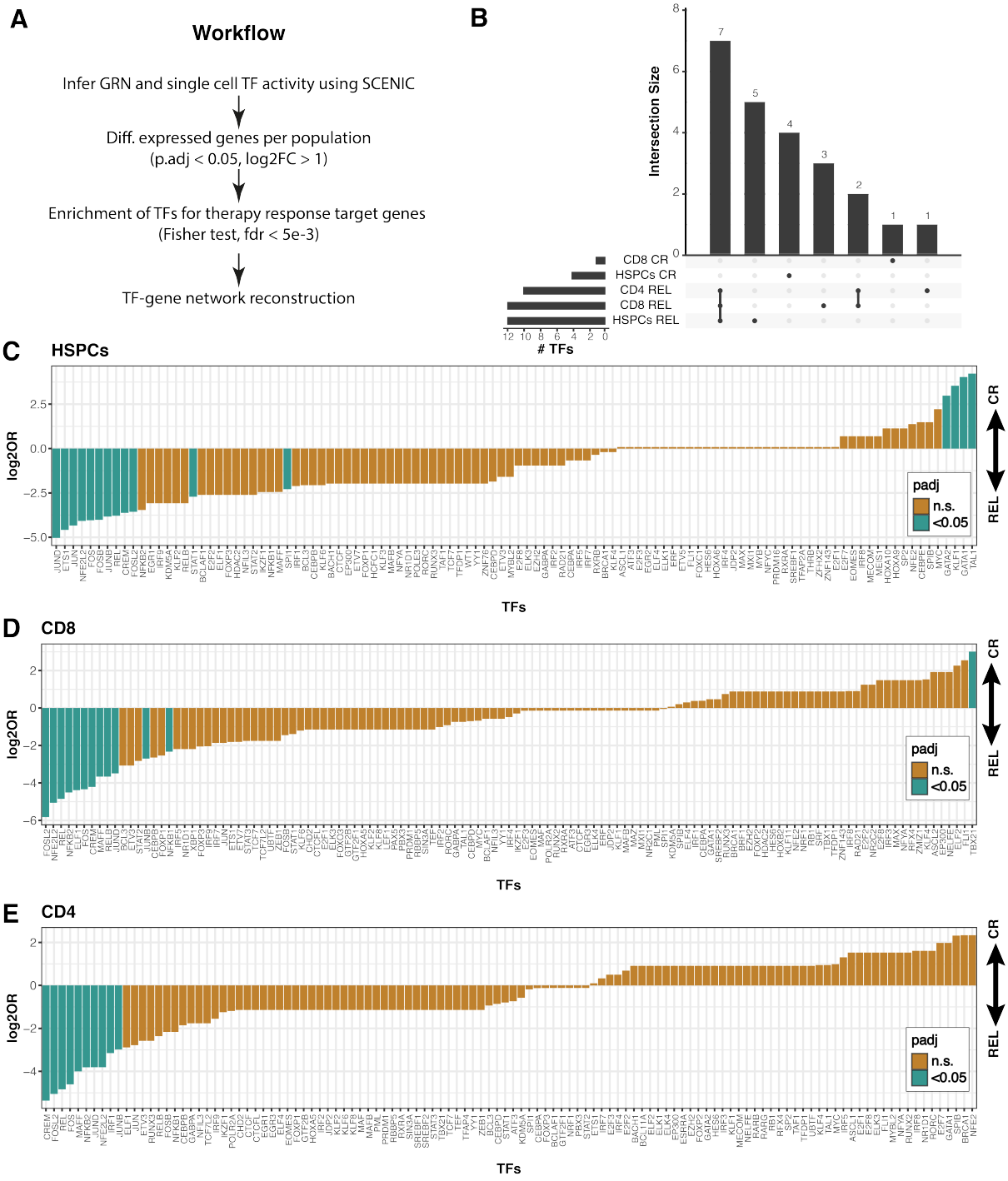

Figure 19 Transcription factor activity analysis using SCENIC.
(A) Overview of SCENIC (3) workflow for identifying differentially active transcription factors (TFs) between two conditions. (B) Upset plot indicating the overlap of the differentially active

TFs between the 3 major populations and the two conditions. (C) Bar graphs depicting the log2OR calculated using Fisher's test. Green colour indicates p.adj < 0.05 after Bonferroni correction.

This analysis revealed in total 23 TFs with regulons enriched among the DEGs (Fisher test; p.adj < 0.05 after Bonferroni correction; **Figure 20 A**). In the HSPCs, *GATA1* and *KLF1* were among the TFs enriched in CR, which are known key regulators of erythroid lineage. This observation is in line with the observed MEP decline in REL versus CR samples (**Figure 16**). In REL samples, most of the TFs enriched were shared amongst $CD4^+$, $CD8^+$ and HSPCs populations. These TFs included *FOS* and *JUN*, the NF-κB family member *REL*, as well as *CREM* and *NFE2L2/NRF2*, which have been linked to an immunosuppressive tumour microenvironment and CD8 exhaustion, respectively[105,106]. Notably, the target genes of the same TFs differed between $CD4^+$ and $CD8^+$ cells, highlighting the importance of studying gene regulation and common TFs at a cell-type specific context[107] (**Figure 20 C**).

In $CD8^+$ T cells, the only TF more active in CR was TBX21. Its regulon mainly comprised of cytotoxic genes like granzymes (*GZMB, GZMH)* as well as killer-like receptor genes (*KLRG1, KLRD1),* was enriched among the genes upregulated in CR CD8 cells (**Figure 20 B**).

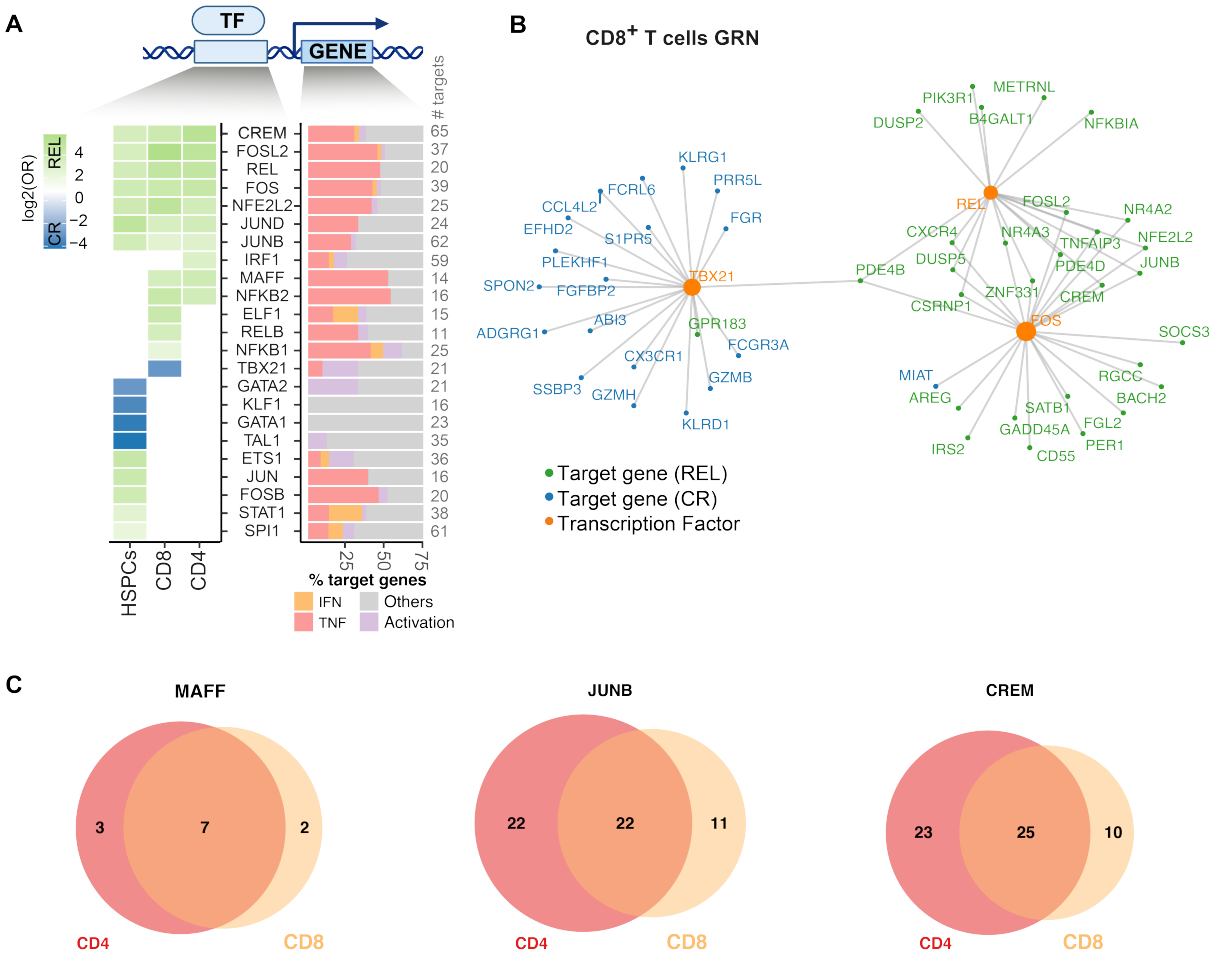

Figure 20 Identification of transcriptional networks associated with therapy outcome post alloSCT.

(A) Differential TF activity analysis using SCENIC. Heatmap indicating log2 odds ratios calculated using Fisher's test (left). Characterization of target genes based on known gene sets (right). Assignment of target genes to known functions was performed using publicly available gene sets (see Suppl. methods). The coloured bars represent the fraction of target genes per TF which belong to the different gene sets (IFN: interferon response, Activation: Immune cell activation, TNF: TNF signalling). (B) CD8$^+$ T gene regulatory network (GRN) of exemplary differentially active TFs (TBX21, REL, FOS) and their target DEGs (C) Overlap of target genes per TF between CD4$^+$ and CD8$^+$ T cells.

2.1.7 Relapse patient CD8 EM T cells have lower cytotoxic potential

The results of the previous section suggest that a key difference between CR and REL in CD8$^+$ cells is related to T cell activation and cytotoxicity. Thus, the core of the rest of the analysis was the CD8$^+$ effector and memory subsets (**Figure 21 C**).

Differential expression analysis in the CD8$^+$ EM cells revealed 144 genes upregulated in CR while 91 genes in REL (**Figure 21 A**; MAST[104]; log2FC $>$ 0.5, p.adj $<$ 0.05 after Bonferroni correction) (**Figure 21 A**). The genes overexpressed in CR CD8$^+$ EM mainly included genes involved in cell killing processes (*KLRG1, KLRD1, GZMB,*

GZMH) as well as *ADGRG1* which encodes for adhesion G-protein-coupled receptor (GPCR) 56 (GPR56). Among the genes overexpressed in REL CD8⁺ EM cells, I identified *CXCR4* previously associated with homing of naive T cells to the BM[108], *GZMK*, as well as *JUNB*, *NFKB*, inflammatory genes like *IFI6* and *CD27*, a member of TNF receptor superfamily, a known regulator of cell activation[109] (**Figure 21 A**). In line, gene ontology (GO) and hallmark enrichment analysis revealed significant enrichment of terms related to T cell activation and cytotoxicity in CR, including the RAS signalling pathway which is essential for the regulation of T cell activation after TCR stimulation (**Figure 21 B**)[110]. On the other hand, terms upregulated in REL comprised NF-κB dependent TNF-alpha signalling, IFNγ response as well as IL-2 production (**Figure 21 B**).

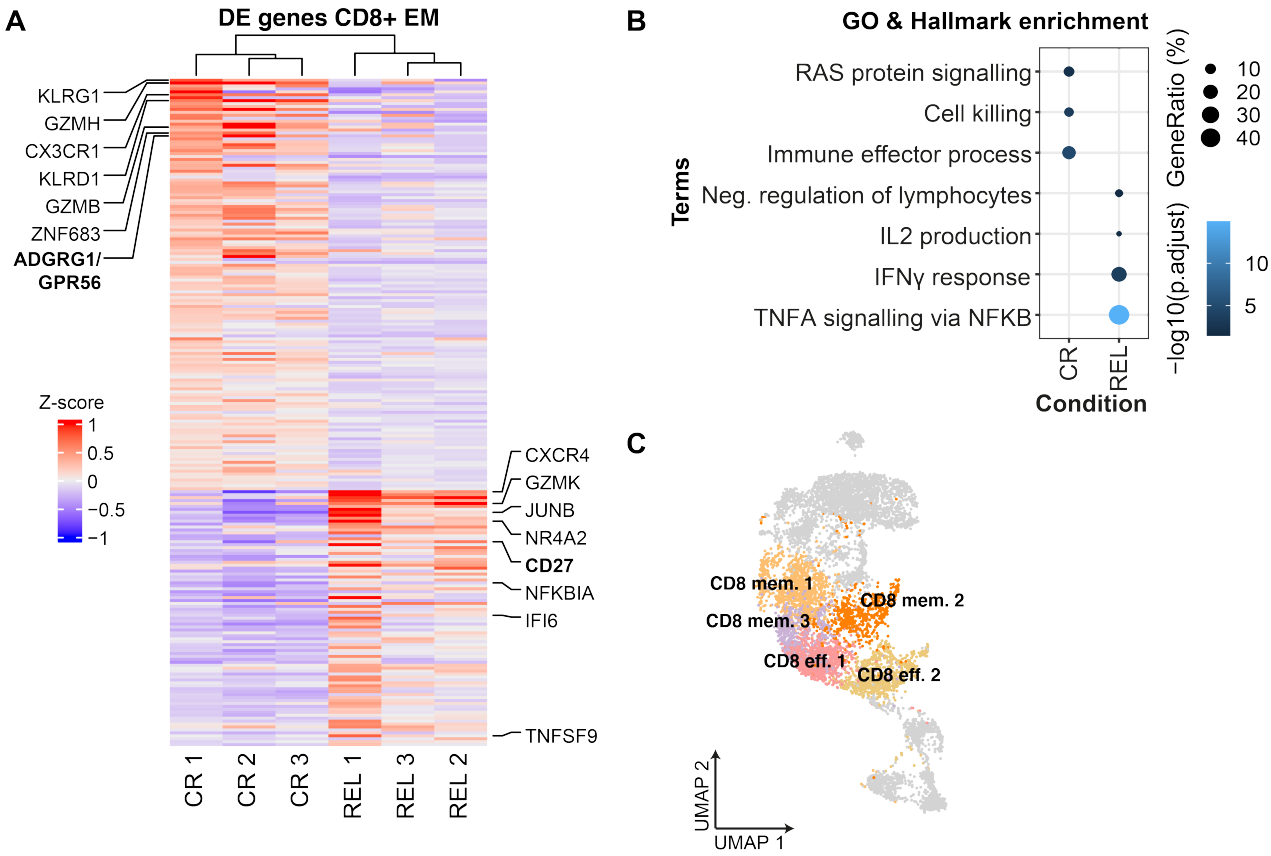

Figure 21 CD8⁺ EM T cells of relapse samples are characterized by lower cytotoxicity. (A) Heatmap depicting scaled expression (z-score) across all 3 samples of 235 differentially expressed genes (DEGs; CR: 144 genes, REL 91 genes) in CD8⁺ effector memory (EM) clusters. Analysis was performed using the MAST algorithm (log2FC > 0.5, p.adj < 0.05 after Bonferroni correction). (B) Gene ontology and hallmark enrichment analysis on the DEGs from A. (C) UMAP highlighting the T cell clusters which belong to the CD8⁺ effector memory (EM) cells and were used for the DE analysis.

2.1.8 GPR56 & CD27 may serve as potential markers for alloSCT outcome

Following up on this analysis, I investigated whether any of the DE genes associated with therapy outcome post alloSCT could serve as a therapy biomarker, with the ultimate goal of being used as a means of monitoring T cell reconstitution after alloSCT and subsequently therapy outcome.

Among the DE genes, I identified two surface molecules which would allow easy detection using flow cytometry without prior extensive processing of the samples; GPR56 was detected to be enriched in CR $CD8^+$ EM cells while *CD27* in REL. Notably, GPR56/*ADGRG1* was detected to be a target gene of TBX21, the only TF more active in CR $CD8^+$ T cells (**Figure 20 B**). GPR56 has previously been reported to be expressed on $CD8^+$ TEMRA and NK cells, and seems to play an inhibitory role on the latter[111,112]. While its functional role in T cells is not well defined yet, Pabst and colleagues previously identified a complex and essential role for GPR56 in driving self-renewal in AML[36]. On the other hand, *CD27*, also known as TNFRSF7, is a TNF receptor instrumental in early activation of T cells[109]. Therefore, I further investigated the dynamics of GPR56 and *CD27* initially in the scRNA-seq data and in the next sections mechanistically as well as in larger independent cohorts.

First, I found a significant increase in the fraction of GPR56 positive cells among the CR samples, while the fraction of *CD27* positive cells decreased (**Figure 22 B, D**). When assessing their expression across pseudotime, which is associated with T cell maturation (**Figure 22 A, C**), a steady increase of GPR56 and a steady decrease in *CD27* in both CR and REL samples was observed. Of note, the baseline levels differed between the 2 conditions, with GPR56 expression being constantly higher at all maturation stages in CR, while CD27 levels were lower (**Figure 22 A, C**).

The aforementioned findings regarding GPR56 from the pseudotime analysis suggested its increases during maturation. I further validated this finding on the same samples using flow cytometry. CCR7 and CD45RA were used to define the different maturation stages of $CD8^+$ T cells and GPR56 levels were compared between the 2 groups (**Figure 22 E**). In line with the RNA levels Figure 13 A, GPR56 was higher in the two effector subsets (TEM, TEMRA) in both conditions, while CR $CD8^+$ cells always showed higher levels of GPR56 on a protein level.

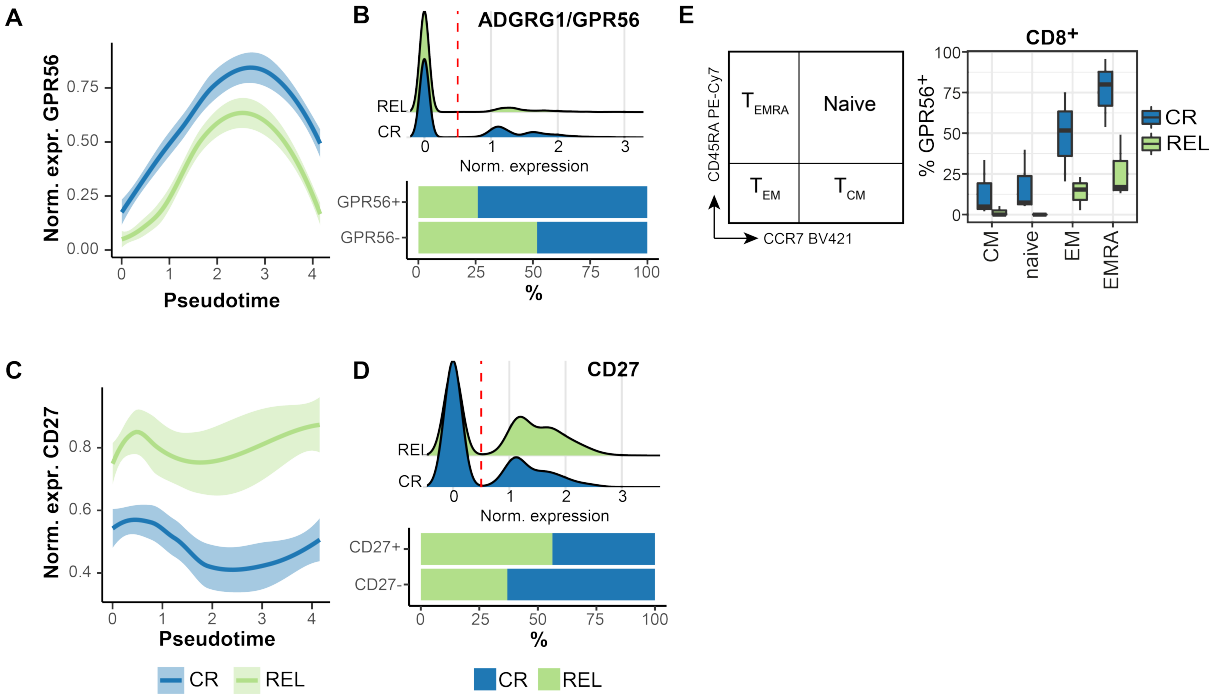

Figure 22 GPR56 and CD27 dynamics across T cells maturation.
(A) GPR56 and CD27 (C) expression across pseudotime of CD8$^+$ cells, split per condition. (B) Density plot indicating the distribution of the normalised expression of ADGRG1/GPR56 and CD27 (D) within the CD8$^+$ EM cells. Vertical red line indicates the threshold (0.5) used for defining a cell as GPR56$^+$ and CD27$^+$ cells. Bottom: Bar plots indicating the percentages of GPR56$^+$ and CD27$^+$ cells in REL and CR samples when considering all CD8$^+$ clusters. (E) Left: Gating strategy used to identify naive, central memory (TCM), effector memory (TEM), and TEMRA cells using CCR7 and CD45RA. Right: Percentage of GPR56 positive cells in the indicated fractions.

2.1.9 GPR56: A potential marker of T cell alloreactivity

GPR56 is co-expressed with cytotoxic molecules

To best associate GPR56 with known T cell phenotypes like T cell activation and allorecognition, I performed DE analysis between GPR56$^+$ and GPR56$^-$ CR CD8$^+$ EM cells. To account for the differences due to the therapy outcome this analysis was performed exclusively in the CR samples, since as presented in Section 2.1.7, the two groups are marked by transcriptional changes. In total, the analysis revealed 24 genes differentially expressed between GPR56$^+$ and GPR56$^-$ CD8$^+$ EM cells (MAST[104]; log2FC > 0.5, p.adj < 0.05 after Bonferroni correction).

At the level of RNA, GPR56$^+$ cells expressed higher levels of *NKG7*, *GZMB*, *GZMH*, *KLRD1*, *GNLY* and *PRF1*, suggesting that GPR56 marks functional, cytotoxic cells. When comparing the GPR56$^+$ fractions between CR and REL (**Figure 23 B**) I identified similar signatures as presented in **Figure 21**, including higher cytotoxic

signatures (*KLRG1*) in CR and increased inflammation (*IFI6*) and *CXCR4* in REL. Moreover, CR GPR56^{+} EM cells were marked with higher expression of *IL32*. In a murine melanoma model, IL32 treatment increased the recruitment of activated tumour-specific CD8^{+} T cells, resulting in systemic induction of anti-tumour immunity[113].

Taken together, CR derived CD8^{+} EM cells are characterised by a higher fraction of GPR56^{+} cells, which based on RNA levels are highly cytotoxic (**Figure 23 A**) relative to GRP56^{-} cells, exerting higher antitumor activity relative to REL (**Figure 23 B**).

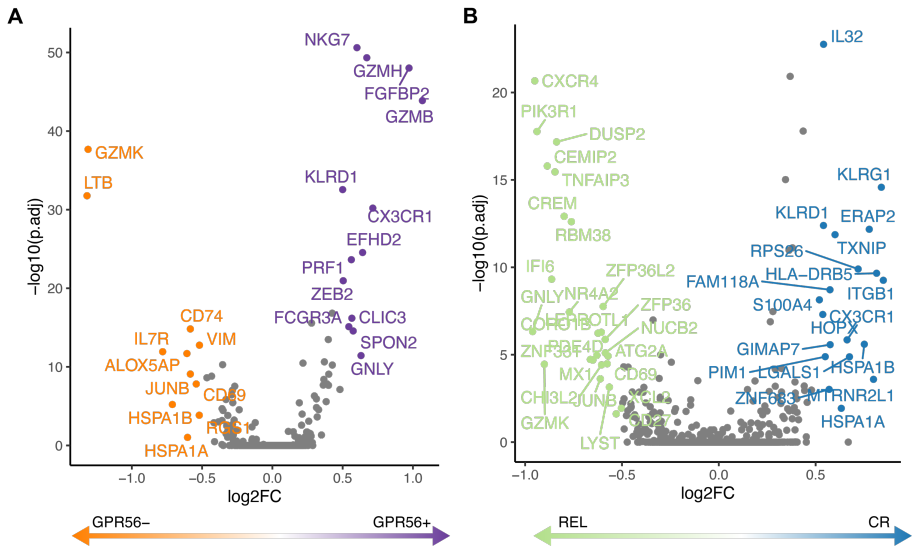

Figure 23 GPR56 co-expression with cytotoxicity molecules at the RNA level.
Volcano plots illustrating the differentially expressed genes between GPR56^{+} (purple) and GPR56^{-} (orange) CD8^{+} effector memory (EM) cells of CR patients (A) and CR versus REL of GRP56^{+} CD8^{+} EM cells (B). Y axis represents p-value after Bonferroni correction (p.adj) and points were coloured according to absolute log2FC > 0.5 and p.adj < 0.05.

These findings were further corroborated using flow cytometry analysis on PBMCs from an independent cohort of 10 AML patients in CR. In detail, the levels of PRF1 and GZMB, central cytotoxic response molecules, were estimated in GPR56^{+} CD8^{+} cells and GPR56^{-} CD8^{+} (**Figure 53**; see Methods Section 4.3). This analysis illustrated that GPR56^{+} cells exhibited higher levels of PRF1 and GZMB in comparison with the GPR56^{-} counterparts (**Figure 24**; paired Student's t-test, p-value < 0.001).

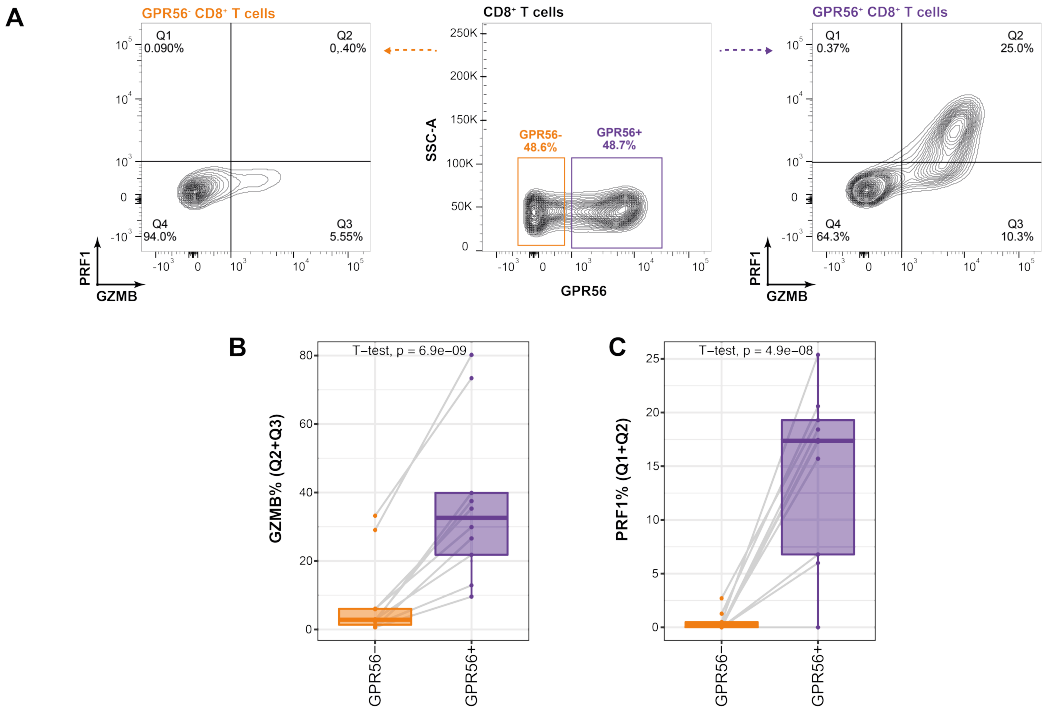

Figure 24 GPR56 co-expression with cytotoxicity molecules at the protein level.
(A) Representative gating scheme for FACS strategy used to separate GPR56$^+$ (purple) from GPR56$^-$ (orange) CD8$^+$ T cells (middle). The 2 fractions were further analysed for PRF1 and GZMB intracellular protein expression. Analysis was performed on peripheral blood mononuclear cells (PBMCs) from a cohort of 10 AML patients in remission. (B) Boxplots illustrate the comparison of the percentage of GZMB$^+$ T cells between the GPR56$^+$ and GPR56$^-$ fractions. Connected points indicate fractions originating from the same sample. P-value was calculated using paired Student's t-test, n = 10. (C) Boxplots illustrate the comparison of the percentage of PRF1$^+$ T cells between the GPR56$^+$ and GPR56$^-$ fractions. Connected points indicate fractions originating from the same sample. P-value was calculated using paired Student's t-test, n = 10.

The findings presented so far suggest that GPR56 may serve as a molecule for monitoring alloreactivity and cytotoxicity in the alloSCT setting, since it was found to be co-expressed with effector/cytotoxic molecules. To further substantiate this observation, the next goal was to investigate its association with T cell activation markers like PD-1 and CD107a using flow cytometry. This analysis showed that GPR56$^+$ cells were characterised by higher levels of CD107a activation molecule, with a median percentage of CD107a$^+$ cells to be 0.993 over 0.441 on the GPR56$^-$ fraction, however CD107a is a very lowly abundant molecule (**Figure 25 A, B**; paired Student's t-test, p-value = 0.0022). No significant association was detected between GPR56 and PD-1 expression (**Figure 25 C, D**; paired Student's t-test, p-value = 0.35).

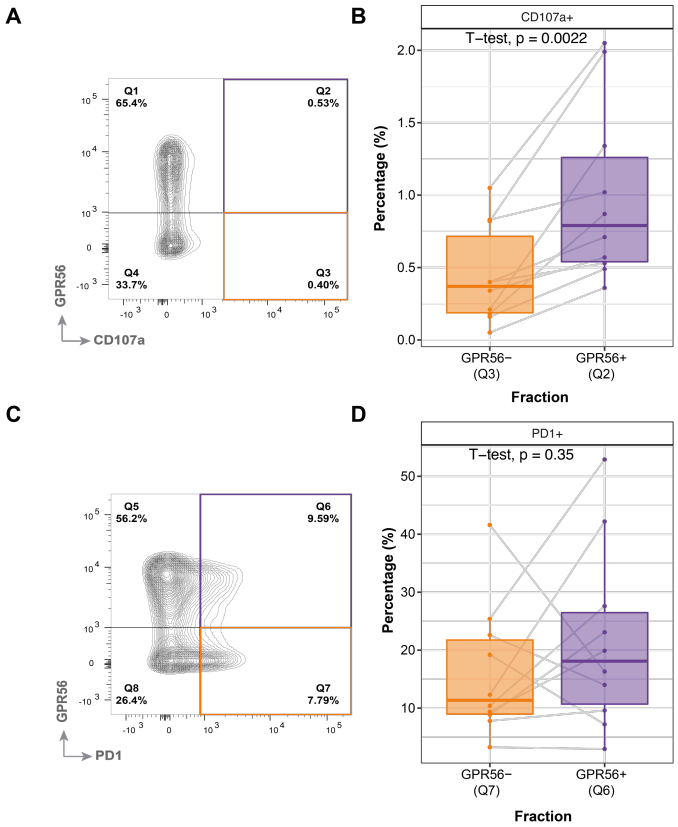

Figure 25 GPR56 is not directly associated with T cell activation molecules.
Representative gating FACS strategy used to separate the different fractions, according to GPR56 and CD107a (A) or PD1 (C) expression. (B) Boxplots illustrating the comparison of the percentage of CD107a$^+$ T cells (B) and PD1$^+$ T cells (D) between the CD8$^+$ GPR56$^+$ and CD8$^+$ GPR56$^-$ fractions. Connected points indicate fractions originating from the same sample. P-value was calculated using paired Student's t-test. Analysis was performed on peripheral blood mononuclear cells (PBMCs) from a cohort of 10 AML patients in remission.

GPR56 is upregulated upon target recognition

Increased GPR56 expression on mature CD8$^+$ T cells suggested that GPR56 expression might be associated with more antigen encounters and therefore a higher GvL potential.

To test this, in collaboration with the Schmidt group at Universitatsklinikum Heidelberg my colleagues and I used Chimeric antigen receptor T (CAR-T) cells as an effector-target cell model. CAR-T cells mimic GvL, as their T cell receptors (TCRs) are genetically engineered to specifically recognise a defined (leukemic) target and eliminate cells expressing this target. CD33-directed CAR-T cells engineered from activated T cells (ATCs) of four healthy donors were co-cultured together with the AML cell line HL-60, in which CD33 was knocked-out[114] (HL60 KO) versus HL-60 with preserved CD33

expression (scrambled CRISPR control, HL60 WT). Non-transduced ATCs were used as negative controls (outlined in **Figure 26 A**).

After the first co-culture, GPR56 upregulation occurred exclusively on CD8$^+$ CD33.CAR-T cells exposed to CD33$^+$ HL-60 WT cells (**Figure 26 B, C**). GPR56 levels remained unchanged on both CAR-T cells exposed to HL-60 KO cells lacking CD33 expression and on non-transduced ATCs exposed to HL60 WT cells (**Figure 26 B, C**). This observation indicated that GPR56 upregulation occurs after TCR-antigen encounter.

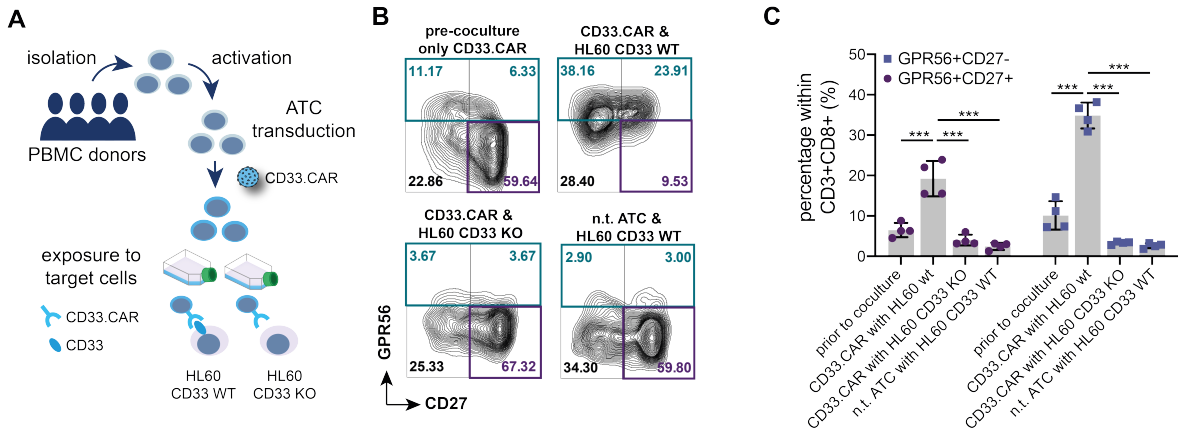

Figure 26 GPR56 upregulation on CAR-T cells after target recognition.
(A) Overview of the experimental setup. Peripheral blood mononucleated cells (PBMCs) from four healthy donors were first activated (activated T cells; ATCs) and then transduced with a retroviral vector comprising the construct of CD33.CAR. On day 15, CAR-T cells were co-cultured with the AML cell line HL60 presenting CD33 on the surface, or with HL60 cells with CD33 knocked-out (KO) using CRISPR/Cas9 (HL60 CD33 KO). (B) FACS plots showing CD27 and GPR56 expression on CAR-T cells after activation and transduction, without contact to leukemia cells (pre-coculture), after a 5-day co-culture with CD33$^+$ HL60 (CD33.CAR & HL60 CD33 WT), after coculture with HL60 CD33 KO cells (CD33.CAR & HL60 CD33 KO), and on non-transduced (n.t.) ATCs after co-culture with HL60 CD33$^+$ WT (n.t. ATC & HL60 CD33 KO). (C) Analysis of the fractions shown in panel B across all 4 samples. Wilcoxon test *** $p < 0.0005$. Panels were generated by Dr. med. Caroline Pabst and experiments were conducted by David Sedloev.

GPR56 dynamics after alloSCT: Insights from independent patient cohort

Given the increase of GPR56 along the pseudotime and its upregulation upon antigen encounter, me and my colleagues hypothesised that GPR56 might serve as a molecule to monitor the dynamics of donor T cell activity in AML patients following alloSCT. The panels and gating schemes are displayed in **Figure 54**.

This hypothesis was investigated together with Caroline Pabst and Xizhe Wang, where we collected and analysed 338 BM aspirates from 139 AML patients. Since several patients sampled at multiple time points and in order to avoid individual over-representation, per patient only the latest available time point was used.

First, the impact of transplantation on GPR56 expression on T cells was investigated. This question was addressed by comparing the BM of patients who never underwent alloSCT, to those prior and post alloSCT. While the percentage of CD8$^+$ TEM in BM did not significantly differ between the three groups, the median fraction of GPR56$^+$ cells in the CD8$^+$ TEM compartment significantly increased after alloSCT (**Figure 27,** median percentage of GPR56$^+$ on CD8$^+$ TEM 12% versus 34% before and after alloSCT, respectively, p<0.005). Similarly, the overall fraction of TEMRA in BM was only slightly higher after alIoSCT, but the fraction of GPR56$^+$ cells within the TEMRA compartment significantly increased after alloSCT (**Figure 27,** median GPR56$^+$ fraction 61% vs. 35% in patients after vs. without alloSCT, p=0.004). Overall, the high GPR56 positivity on CD8$^+$ T cell subsets, which was found to exceed the levels of non-transplanted patients, supported that GPR56 may serve as a hallmark of the alloSCT setting.

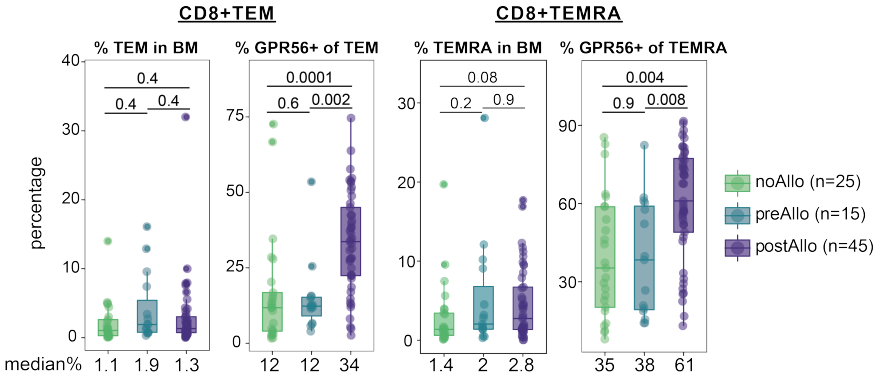

Figure 27 GPR56 is upregulated after alloSCT.
(From left to right) Boxplots depicting medians and quartiles of the fractions of CD8$^+$ TEM, GPR56$^+$ of CD8$^+$ TEM, CD8$^+$ TEMRA and GPR56$^+$ of CD8$^+$ TEMRA in the bone marrow (BM) of patients without (noAllo), before (preAllo) and after (postAllo) allogeneic stem cell transplantation (alloSCT). Numbers on the x axis indicate the median percentage. Numbers between groups of patients indicate the adjusted p-values after unpaired Wilcoxon test. The total fraction of CD8$^+$ TEM and TEMRA in BM does not significantly differ between the three groups.

To further characterise the differences between transplanted and non-transplanted T cells, combined characterisation of GPR56 and CD27 surface staining was performed.

Whereas $CD8^+$ cells in healthy BM predominantly expressed CD27 but not GPR56 (**Figure 28 A**, upper), a broad variability in alloSCT samples ranging from predominant $CD27^+$ $GPR56^-$ and $CD27^+$ $GPR56^+$ patterns to samples, in which the $GPR56^+$ $CD27^-$ fraction represented the largest population was observed (**Figure 28 A**, middle and lower). Given that the samples in this cohort were harvested at variable time points after transplantation, the available CR samples were grouped according to their aspiration time point and were then compared for the expression of the two surface markers in these groups. While the $GPR56^-$ $CD27^+$ fraction continuously decreased over time, the $GPR56^+$ $CD27^+$ double positive (DP) fraction increased in the early months and declined at later time points (**Figure 28 B**). The $GPR56^+$ $CD27^-$ fraction continuously increased one year post alloSCT compared to 10% in non-transplanted patients (**Figure 28 B**). $CD3^-$ $CD56^+$ NK cells did not display dynamic changes of the $GPR56^+$ fractions over time after alloSCT (**Figure 28 C**).

These findings suggest a phenotype switch explicitly on the $CD8^+$ T cells upon antigen encounter post alloSCT from a mostly $CD27^+$ $GPR56^-$ to a $CD27^-$ $GPR56^+$ phenotype, with a temporary increase in the $CD27^+$ $GPR56^+$ fraction in between (**Figure 28 D**).

Multiple clinical parameters are critical for alloSCT outcome, including the cytomegalovirus (CMV) status of both the donor and the recipient [32]. Thus, the next step was to investigate the impact of CMV serostatus on the GPR56 expression on $CD8^+$ effector T cells. Interestingly, the recipient's CMV serostatus (but not the donor's) was identified as a main contributor to high GPR56 expression even in non-transplanted patients (**Figure 28 E**). The median percentage of $GPR56^+$ on $CD8^+$ TEMRA was 40% vs. 18% in CMV IgG positive versus negative non-transplanted patients (**Figure 28 E**). While GPR56 expression on TEMRA increased in both CMV IgG positive and negative patients, this occurred on a higher baseline level in CMV positive patients (**Figure 28 E**, median $GPR56^+$ rising from 50% to 75% from month 1-3 to >24 months in CMV IgG positive and from 27% to 57% in CMV IgG negative patients).

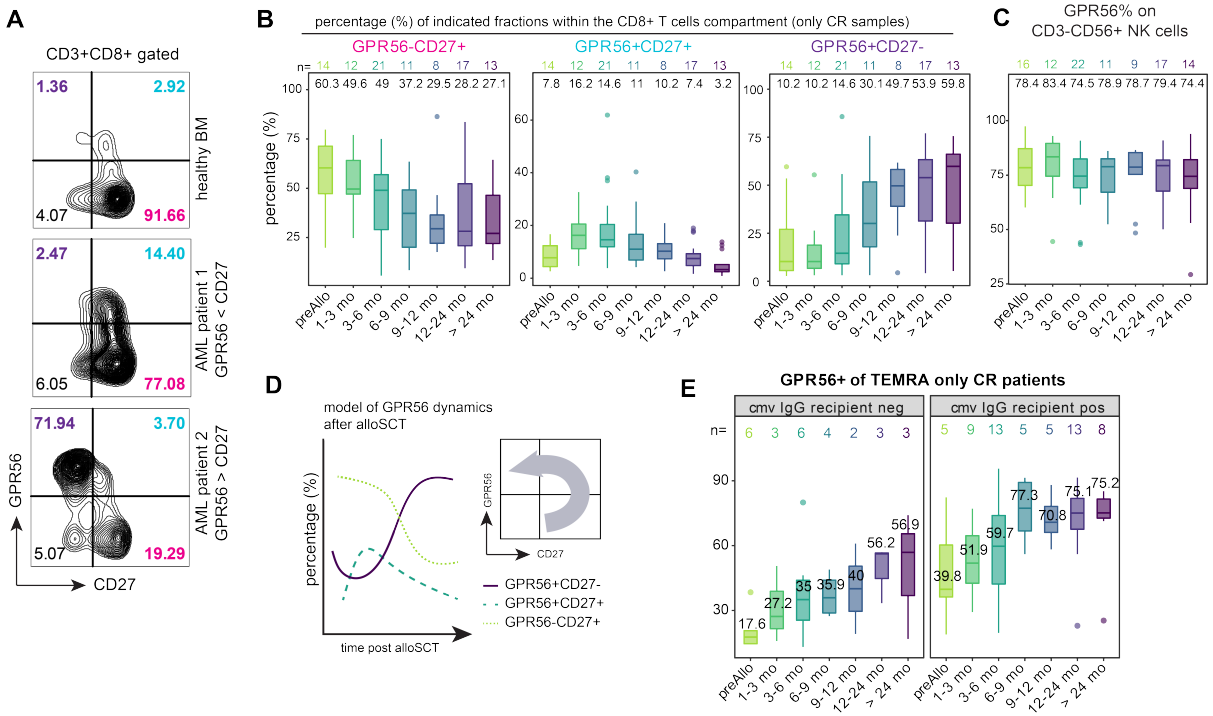

Figure 28 Continuous GPR56 increase on donor T cells after alloSCT.
(A) Representative FACS plots depicting CD27 and GPR56 dynamics on $CD3^+$ $CD8^+$ cells of healthy bone marrow (BM; top) and two AML patients, with either low GPR56 levels (middle) or with a dominant $GPR56^+$ $CD27^-$ fraction (bottom). (B) Boxplots illustrating the medians and quartiles of the percentages of $GPR56^-$ $CD27^+$, $GPR56^+$ $CD27^+$, and $GPR56^+$ $CD27^-$ in the $CD8^+$ compartment in CR patients post alloSCT over time. (C) Boxplots illustrating the medians and quartiles of the percentages of $GPR56^+$ cells on $CD3^-$ $CD56^+$ NK cells. (D) Proposed model derived from (B) of $CD8^+$ T cell phenotype switch post alloSCT. (E) Percentage of $GPR56^+$ on $CD8^+$ TEMRA in CMV IgG negative and positive recipients. Numbers next to box plots indicate the median percentages. Box plots represent medians, quartiles, and outliers. Panels were produced by Dr. med. Caroline Pabst.

2.1.10 Synopsis

All together, the observations presented in this Chapter provide a more comprehensive understanding of T cell biology in the context of alloSCT. This was achieved by analyzing HSPCs and T cells from day +100 BM aspirates of patients who either remained in complete remission or suffered relapse shortly after sampling, using scRNA-seq. By mapping these cells, I was able to identify key differences in their transcriptional patterns associated with therapy outcome, including TNF-α/NF-κB signaling as well as T cell cytotoxicity. Ultimately, these findings suggest GPR56 expression on T cells to be an indicator for a favorable alloSCT outcome. Advancing the understanding of the factors which contribute to GPR56 upregulation, could potentially enhance graft versus

leukemia effects, enabling more effective donor selection and immunomodulation after alloSCT.

Further discussion of the findings presented in this Chapter will be explored in Section 3.1.

2.2 Project 2: Single cell profiling of xenograft mouse models unveil the bone marrow microenvironment remodeling in Acute Myeloid Leukemia

With the following exceptions, the experiments and analysis presented in this Chapter were designed and performed by me with discussion and guidance from **Dr. Judith Zaugg** and **Dr. Med. Caroline Pabst**.

In detail, I contributed to the experimental design, conducted experiments (flow cytometry, scRNA-seq experiments, library preparations, *in vitro* co-cultures), performed scRNA-seq data analysis and data interpretation.

Dr. Swati Garg provided the CRISPR/Cas9 HLF KO AML lines and helped with the *in vitro* co-culture experiments. **Dr. Alicia He** performed the mouse transplantations. The xenotransplantation experiments were designed and performed jointly with **Dr. Karin Prummel**. Mouse harvesting and FACS was performed together with **Dr. Shubakar Soob** and the rest of **Dr. Marieke Essers** team. **Rim Moussa** assisted with the computational analysis under my guidance.

2.2.1 Summary

The focus of this Chapter is to investigate impact of acute myeloid leukemia (AML) and in particular leukemic stem cells (LSCs) to the bone marrow (BM) microenvironment. To achieve this goal, a combinatorial approach was utilized, where scRNA-seq of AML xenograft models as well as in vitro co-cultures of patient-derived bone marrow mesenchymal stromal cells (MSCs) with AML cells were combined. While assessing the impact of AML, the LSC burden was considered. The evidence presented here outline the AML induced compositional as well as phenotypic switch of the non-hematopoietic BM compartment.

2.2.2 Experimental overview

In order to comprehensively study the impact of Acute myeloid leukemia (AML) and leukemic stem cells (LSCs) on the bone marrow (BM) niche, my colleagues and I used AML patient-derived-xenograft (PDX) mouse models. As a model of high LSC burden AML (LSC_{high} AML) a patient derived triple-mutated (FLT3-ITD/NPM1/DNMT3A) AML line was used, while for low LSC burden AML (LSC_{low} AML) the same line with hepatic leukemia factor (HLF) knock out (KO) was used. Pabst, Zaugg and colleagues previously identified HLF as a stemness regulator in AML[38]. In particular, it was demonstrated that a CRISPR/Cas9-mediated HLF knockout in AML cells leads to a significantly lower LSC frequency after injection in recipient mice[38]. Thus, here HLF KO AML cells were used as a surrogate of LSC_{low} AML. As control conditions, non-transplanted mice and mice transplanted with healthy cord blood-derived (CB) $CD34^+$ cells were included. All conditions contained 3 biological replicates (**Figure 29 A, B**). I verified the lower abundance of HLF in AML cells after CRISPR/Cas9-mediated knockout by western blot (**Figure 29 B**).

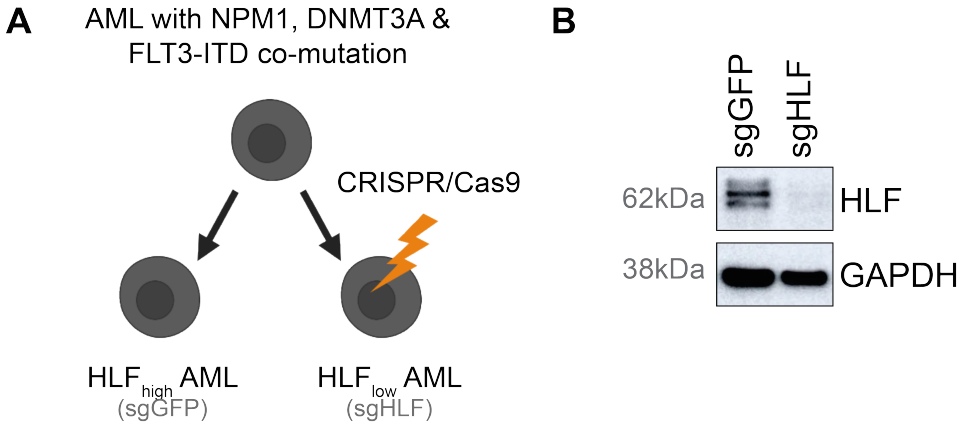

Figure 29 Strategy for acquiring LSC_{high} and LSC_{low} AML.
(A) Schematic overview of the strategy of acquiring AML lines with high leukemic stem cell (LSC) burden (LSC_{high}) and low LSC burden (LSC_{low}). Patient derived triple mutated AML cells (NPM1/DNMT3A/FLT3-ITD) were subjected to CRISPR/Cas9 induced knock outs using a single guide RNA (sgRNA) against hepatic leukemia factor (HLF, sgHLF), hereafter termed as LSC_{low} AML cells or against GFP (sgGFP, LSC_{high} AML). (B) Western Blot validating reduced levels of HLF protein expression in HLF KO AML line. GAPDH was used as a control.

For the human cell xenotransplantation, KIT-deficient NOD/SCID *Il2rg−/−KitW41/W41* (NSGW41) murine strain was used, which allows engraftment of cells without prior irradiation[62] (**Figure 30**). Control mice, as well as mice transplanted with human derived CB-$CD34^+$ control cells, LSC_{high} and LSC_{low} AML cells were sacrificed 7 weeks after transplantation (**Figure 30**, see Methods, Section 4.8). At this point, 70-80% of engraftment was achieved (**Figure 41 A**).

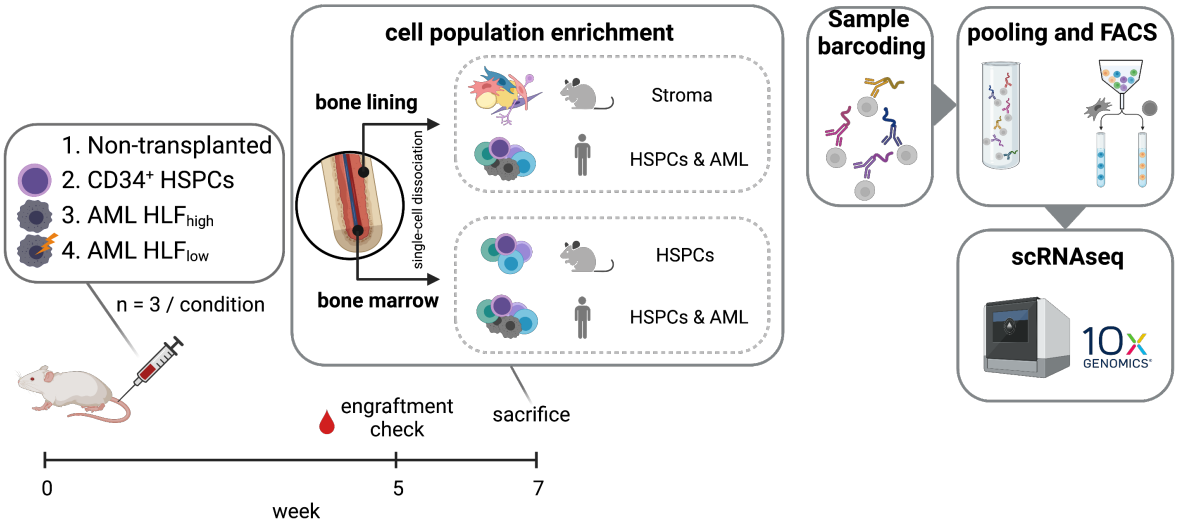

Figure 30 Overview of the experimental design to characterize the BM microenvironment in AML using NSGW41 xenograft mouse model.
Recipient mice were transplanted with human derived AML cells and cord blood (CB) derived CD34$^+$ hematopoietic stem and progenitor cells (HSPCs). After 7 weeks of engraftment, both cells from the bone lining and the bone marrow (BM) compartments were isolated using the optimised tissue digestion protocol outlined in Baccin et al., 2020[26]. Non-hematopoietic mouse niche cells were enriched as well as human and mouse HSPCs, and human AML cells using fluorescent activated cell sorting (FACS). Since samples were barcoded using TotalSeqB oligonucleotide tagged antibodies, they were pooled and subjected to droplet-based single cell RNA-sequencing (scRNA-seq) using the 10x Genomics platform.

In order to capture both highly abundant and rare BM resident cell types, me and my colleagues performed scRNA-seq using the 10x Genomics platform of cells from the total non-digested bone marrow after the progressive depletion of highly abundant cell types or after the enrichment of rare populations of enzymatically digested bones (Figure 31, see Methods Section 4.9). In particular, using fluorescent activated cell sorting (FACS) mouse niche cells (mCd45$^-$ mTer119$^-$ mCD41$^-$ mCD71$^-$ huCD45$^-$ stromal cells and mCD31$^+$ endothelial cells), human and mouse HSPCs (huCD45$^+$ huCD34$^+$ & mCD45$^+$ mCD34$^+$), and human AML cells (huCD45$^+$) were enriched (**Figure 31**). Furthermore, these cells were labelled with sample-specific oligo-conjugated antibodies (TotalSeq-B; see Methods Section 4.10) prior to loading on the 10X, allowing for sample deconvolution during downstream analyses.

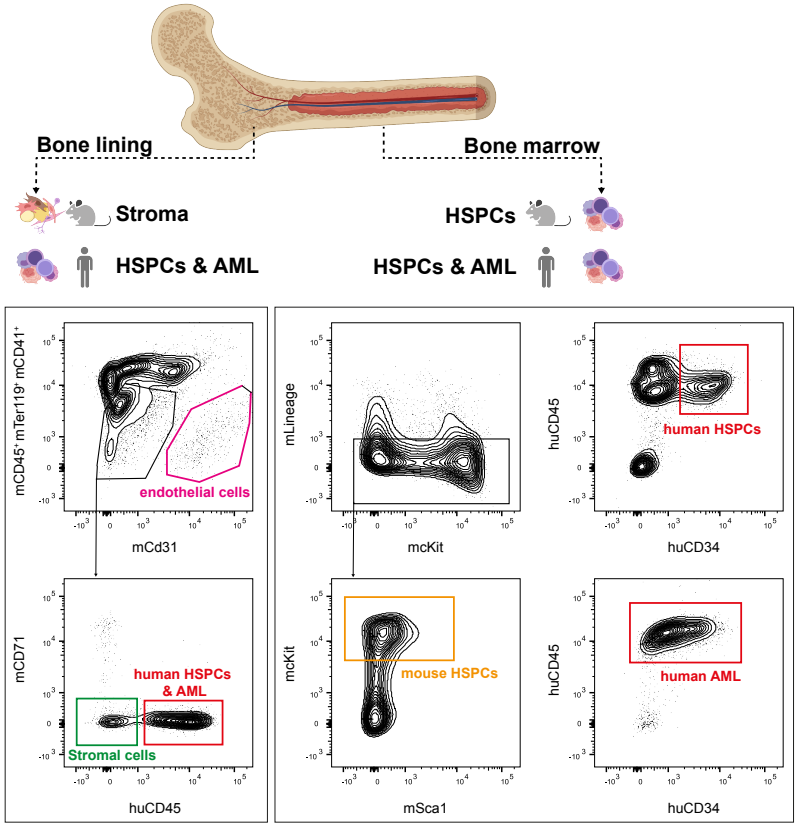

Figure 31 Overview of the sorting strategy.
In short, bone lining cells obtained after bone crushing and digestion as well bone marrow (BM) cells were subjected to mouse hematopoietic lineage depletion. (Left) From the bone, endothelial cells (mCd45$^-$ mTer119$^-$ mCD41$^-$ mCd31$^+$), stromal cells (mCd45$^-$ mTer119$^-$ mCD41$^-$ mCD45$^-$ mCD71$^-$) as well as human HSPCs/AML cells (mCd45$^-$ mTer119$^-$ mCD41$^-$ mCD71$^-$ huCD45$^+$) were sorted. (Right) BM samples were used for the enrichment of mouse HSPCs (mLin$^-$, mKit$^+$), human HSPCs (mLin$^-$, huCD45$^+$ huCD34$^+$) and human AML (mLin$^-$, huCD45$^+$) cells. mLin: mTer119/mCD11b/B220/mCD4/mGr1/mCD8a. Figure was generated using Biorender.com. Flow cytometry panels were provided by Dr. Karin Prummel.

2.2.3 Computational overview of scRNA-seq analysis of xenotransplanted models

Cell loss during experiments

After FACS, cells were subjected to droplet based 10X Genomics scRNA-seq (see Methods Section 4.5). Cells were first loaded on a microfluidics chip, though the number of recovered cells sequenced may decrease due to technical reasons. Extensive efforts have been made in the field of single cell genomics to better understand possible sources of cell loss during such experimental procedures[115]. As illustrated in **Table 2**, the number of cells loaded and finally sequenced in the control non-transplanted mice did not differ, though in the rest of the samples an average loss of approximately 50% cells was observed.

Table 2: Number of cells subjected to scRNA-seq, as estimated upon loading on the chip (# cells loaded) as well as after sequencing, as estimated with CellRanger, the computational pipeline for processing these scRNA-seq data acquired with the 10x genomics platform.

Sample	# cells loaded	# cells - CellRanger
Non-transplanted stroma	9,400	9,874
Non-transplanted HSPCs	11,300	9,846
Inj. $CD34^+$ stroma	10,000	4,305
Inj. $CD34^+$ HSPCs	15,400	7,260
Inj. LSC_{high} stroma	8,200	3,930
Inj. LSC_{high} HSPCs	13,200	6,212
Inj. LSC_{low} stroma	9,400	3,897
Inj. LSC_{low} HSPCs	12,900	5,557

Species and replicate assignment single cell data

In the dataset obtained both human and mouse derived single cells were represented. I first performed the alignment of the data against single-organisms reference genomes (human: GRCh38, mouse: mm10). This approach led to single cell barcodes assigned to both mm10 and GRCh38 genomes. As an example, in the group transplanted with CB-$CD34^+$ HSPCs, 2,245 barcodes were assigned in both genomes (**Table 3**). This issue might have been a result of high genome similarity between the two species. To further address this issue, I realigned the data on a prebuilt reference genome which contains both human and mouse reads provided by 10x Genomics, termed barnyard reference genome (see Methods, Section 4.13.1).

As demonstrated in **Table 3**, alignment against the barnyard reference genome resulted in a smaller number of cells assigned to each species, especially in the case of mm10 (1,069 cells, in comparison to 2,688 cells). Though, the number of cells calculated after single species alignment is likely overestimated, since 2,245 cell barcodes were assigned to both human and mouse species.

Table 3: Number of cell barcodes recovered after performing single species alignment (mm10, GRCh38), as well as by performing alignment using the barnyard reference.

	mm10, GRCh38 separate references *	mm10, GRCh38 separate references	Barnyard reference
mm10	2,246	2,688	1,069
GRCh38	2,578	2,952	2,379
Assigned to both mm10 & GRCh38		2,245	

** low coverage sequencing*

After concluding that the barnyard approach could recover a reliable number of cells, the next step was to assign each single cell to a single species. To do so, the log

transformed number of mm10 relative to GRCh38 reads was calculated (**Figure 32 A**). That way, doublets containing cells of both human and mouse origin were detected (**Figure 32 A**). Following doublets exclusion, in order to accurately assign single cells to single organisms, I used the ratio of log transformed GRCh38 mapped UMIs divided by the UMIs mapped to mm10. Cells with a ratio greater than zero were considered of human origin (**Figure 32 B**). Cells of poor quality were excluded from further downstream analysis (percentage of human and mouse mitochondrial reads > 10% and total number of genes < 250).

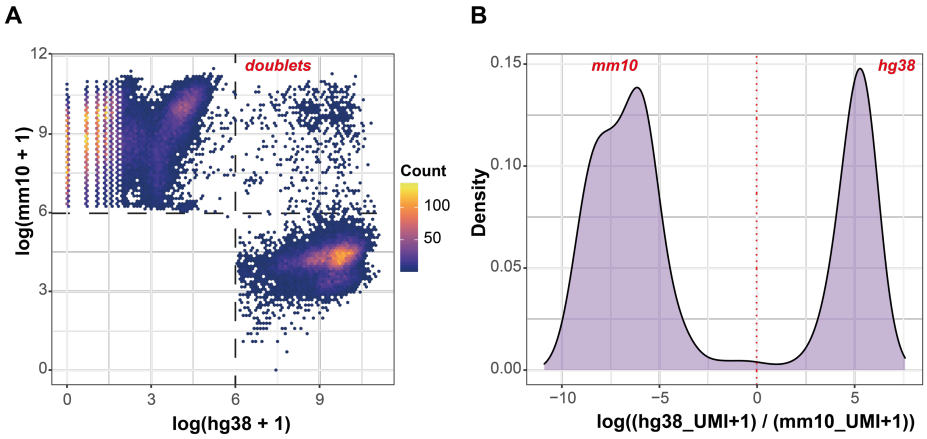

Figure 32 Computational approach for species assignment.
(A) Scatter plot depicting the log normalized GRCh38 (hg38) unique molecular identifiers (UMIs) on the x axis and log normalized mm10 UMIs on the y axis per single cell. Cells considered as doublets based on the number of UMIs are labelled. (B) Density plot of ratio used for mm10 to GRCh38 single cell distinction.

In total, the dataset comprized 22,958 cells of murine origin as well as 14,316 xenotransplanted human derived cells. After extracting species specific barcodes, single reference re-alignment was performed. This resulted in a differentially distributed number of genes (**Figure 33A, B;** Barnyard reference: mean = 2,740.598 and median = 2,585; mm10 reference: mean = 2,479.262 and median = 1,897) and reads (**Figure 33A, B;** Barnyard reference: mean = 12,736.28 and median = 8,477; mm10 reference: mean = 1,138.75 and median = 6,755).

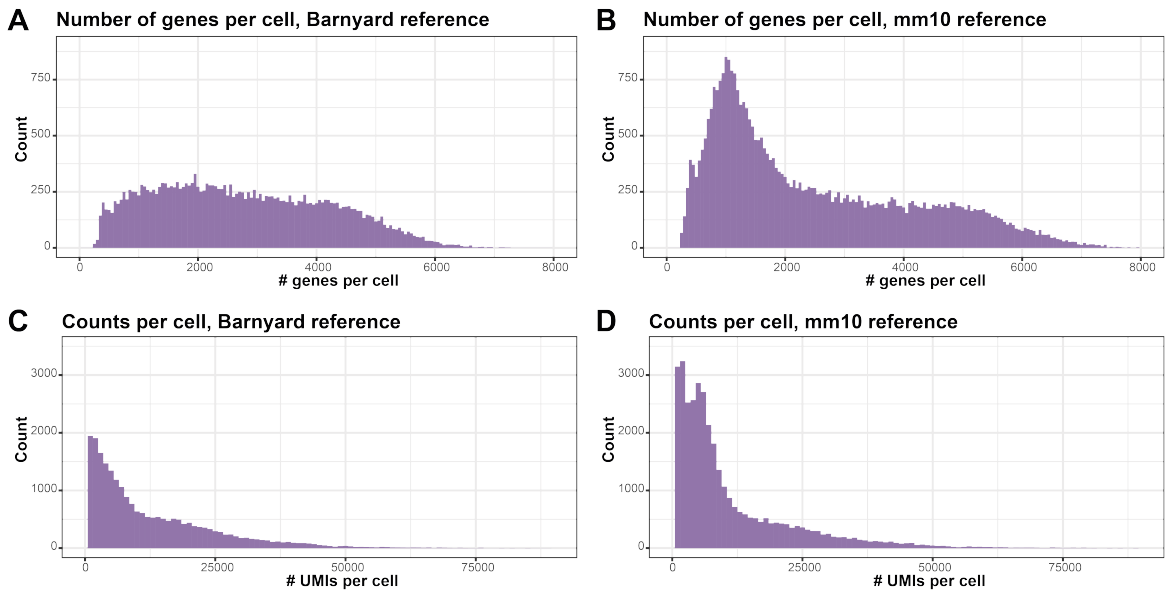

Figure 33 Comparison between barnyard and single species alignment.
Read counts and number of genes distributions after Barnyard (A, C) and single species (B, D), mm10 reference alignment, respectively. Panels were generated by Rim Moussa.

As depicted in **Figure 30**, single cells from all replicates per group were pooled together after labelling with oligo-conjugated antibodies targeting pan-cell surface antigens (TotalSeq-B; See Methods Section 4.10) – an experimental procedure termed cell hashing. Because each replicate per group was assigned a unique cell hashing barcode, it was possible to computationally deconvolute the sample origin of individual single cells. Of note, I was able to confidently map 18,626 out of 50,881 single cells to individual samples (36.6%, **Figure 34 A**). The rest of the cells were either assigned as doublets (i.e., associated with more than one sample oligonucleotide) or unassigned (**Figure 34 A**). Interestingly, the majority of the cells not assigned to single replicates/barcodes originated from the bone lining, such as murine, non-hematopoietic-stromal cells (**Figure 34 B**). The target antigen of the cell hashing antibodies was MHC-I, which is known to be ubiquitously expressed on all mammalian cells, including stroma, this result was unexpected. Since the majority of single cell studies and optimisations have been performed on hematopoietic cells, a possible explanation is that the properties of these cells still need to be better clarified.

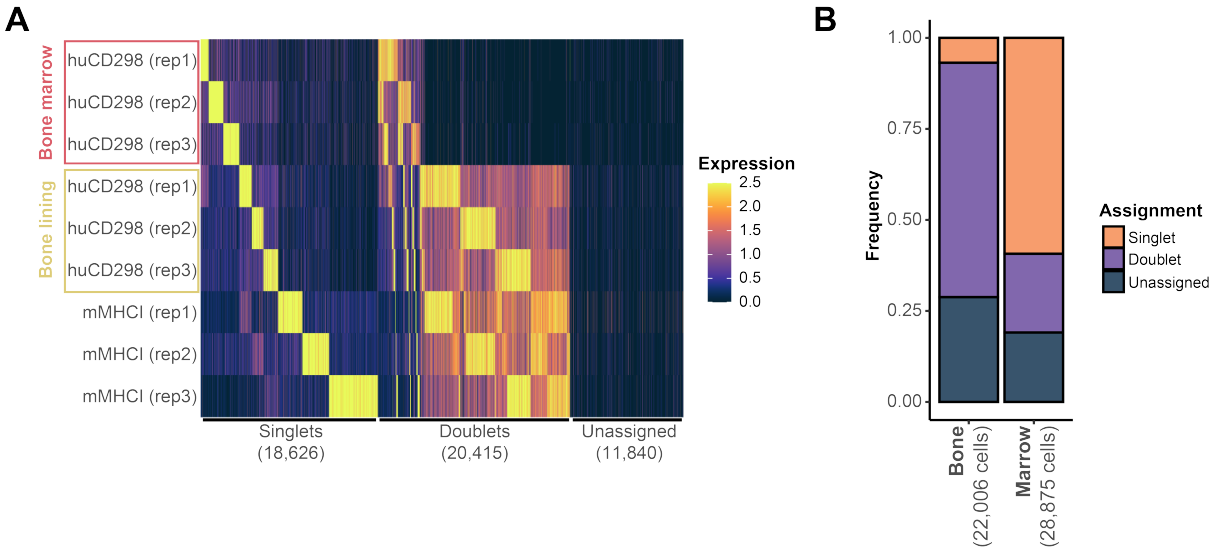

Figure 34 Replicate assignment of scRNA-seq data using TotalSeq-B antibody oligonucleotides. (A) Heatmap of scaled (z-scores) normalized antibody oligonucleotide values based on the calculated assignments. Doublets express more than one, while unassigned/negative cells express non. The data illustrated here are prior quality control. (B) Fractions of assigned cells, depending on the region of origin.

2.2.4 Bone marrow landscape of PDX NSGW41 mouse model

After assigning single cells to species and samples, I performed cell type annotation of the unsupervised yet distinct hematopoietic and non-hematopoietic clusters based on known marker genes. Detailed marker genes used for the annotation can be found in **Figure 35**. This analysis revealed the presence of rare non-hematopoietic populations including smooth muscle cells/pericytes (*Myh11, Mustn1, Tagln, Acta2* positive), schwann cells (*Mal, Mag* positive) as well as a subset of Ng2$^+$ glial cells (Ng2/*Cspg4*, *Kcna1* positive), myofibroblasts (*Acta2, Myf5*), and endothelial clusters. The endothelial clusters (*Cdh5*, Cd31/*Pecam1* positive) were further subdivided to 2 sinusoidal (*Emcn* positive) and 3 artelioral subsets (*Ly6a* positive). Furthermore, I identified *Pdgfra*$^+$ mesenchymal populations. These included 2 osteoblast clusters (*Bglap*, *Col1a1* positive), chondrocytes (*Sox9, Acan* positive), 6 fibroblasts-like populations, 2 of which were CD34$^+$ (Fibroblasts I-IV, CD34$^+$ Fibroblasts I-II) as well as Cxcl12-abundant reticular (CAR) cells (*Cxcl12*, *Kitl, Lepr* positive). The latter population could be subdivided to 2 Adipo-CAR (*Adipoq*) clusters and CAR expressing a combination of adipo- and osteo-lineage genes (*Alpl, Mmp13*). The presence of the aforementioned CAR subsets is in line with previous studies in C57BL/6J mice, which have identified CAR cells characterising by simultaneous expression of both lineages[25] or single lineages gene sets[26]. The analysis also revealed the presence of skeletal muscle cells (*Ckm, Ttn* positive) as well as

lymphatic vessels' *Lyve1*$^{+}$ ECs (*Prox1, Lyve1* positive), which were hypothesized to originate from the outer layer of the bone.

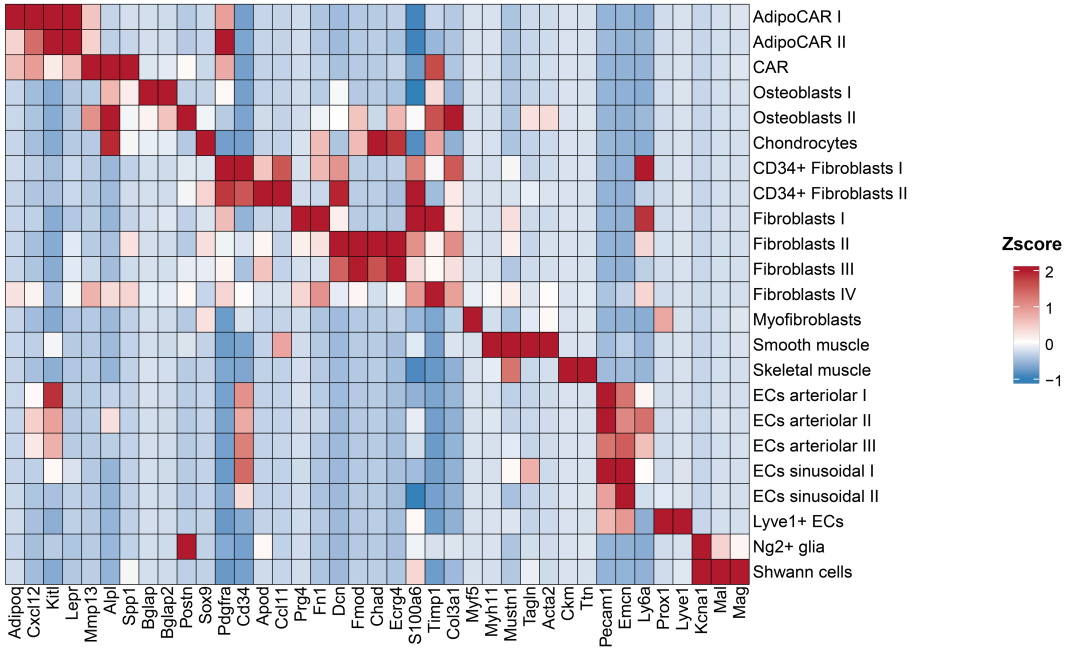

Figure 35 Heatmap depicting scaled expression of marker genes across all 23 unsupervised non-hematopoietic clusters.
Values are averaged across all cells within each cluster and then scaled (z-score) across all clusters.

Similarly, I performed the annotation of Lin^{-} cKit^{+} cells which cover the NSGW41 HSPCs populations according to previously published studies[26,88]. This analysis showed the existence of 12 mouse HSPCs clusters, spanning mainly across the myeloid blanches of hematopoietic progenitors. Specifically, I identified a branch of megakaryocyte–erythrocyte progenitors (Ery/Mk progenitors I & II, Erythroid progenitors, Erythroblasts), a branch of neutrophil/monocyte progenitors and a separate cluster of eosinophil/basophils progenitors (Eo/Baseo. progenitors). In addition, I identified a cluster of lymphoid-primed multipotent progenitors (LMPPs; *Kit*, *Msi2*, *Pim1*) but no further committed lymphoid branches. This observation is in line with the literature since NSGW41 mice genetic background does not allow the maturation towards the lymphoid lineage[62]. An overview of all the genes which contributed to the cluster assignment is presented in **Figure 36**.

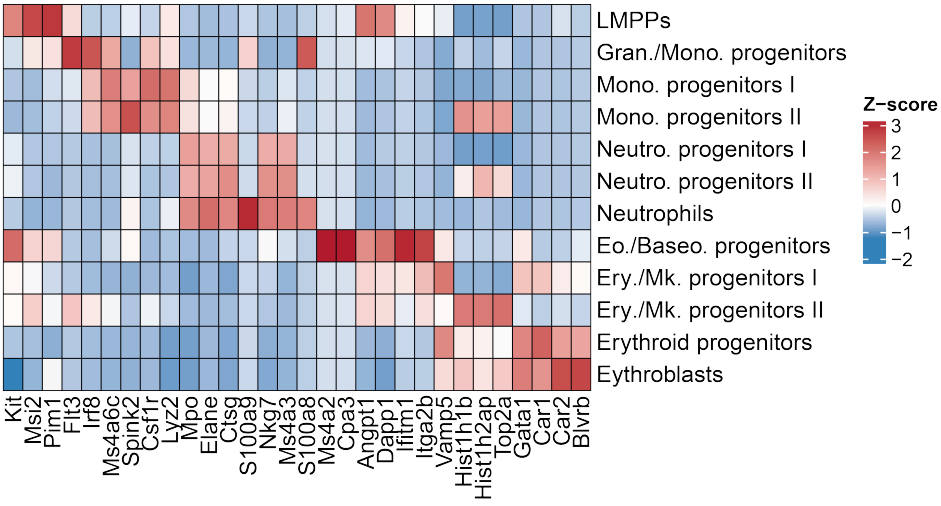

Figure 36 Heatmap depicting scaled expression of marker genes across all 12 unsupervised hematopoietic clusters of NSGW41 mice.
Values are averaged across all cells within each cluster and then scaled (z-score) across all clusters.

Overall, this dataset recapitulates all stromal and HSPCs populations which have been previously described in other mouse models (i.e.: C57BL/6J mice mice) (**Figure 37**).

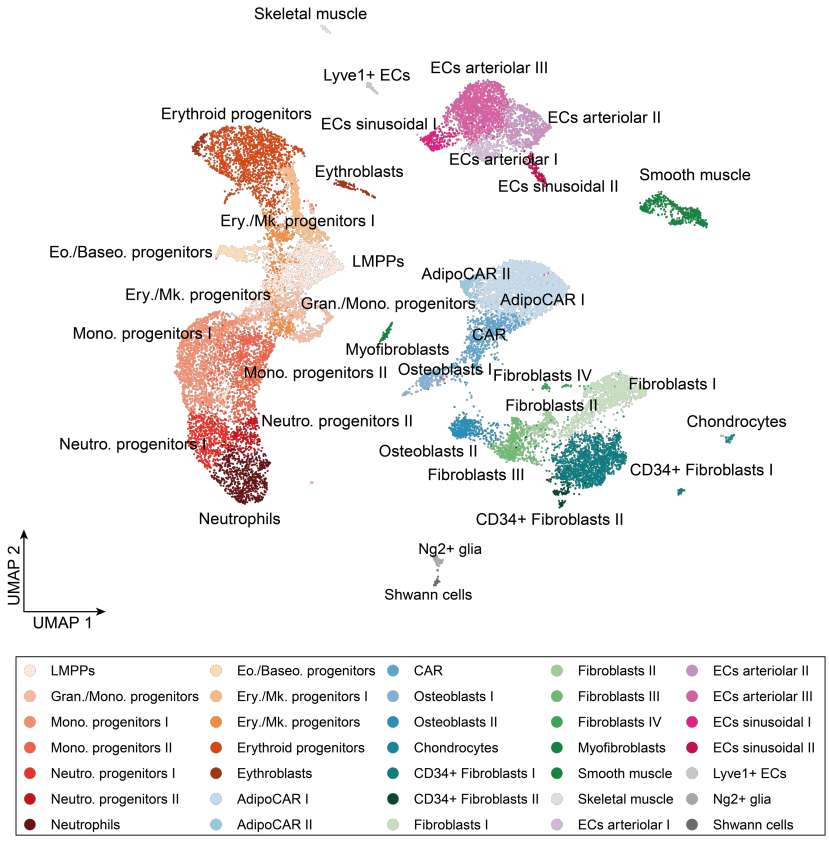

Figure 37 The single cell landscape of the BM niche of NSGW41 mice.

Uniform manifold approximation and projection (UMAP) of 22,981 post-QC single cells of the NSGW41 bone marrow mice. Cells are coloured according to cell type.

To ensure that this model is suitable for studying human hematopoiesis, I then assessed the expression of genes known to support HSPCs *in vivo*; *Kitl* [116], *Il7* [117], *Igf1* [118], *Cxcl12* [119], *Csf1* [120] and *Bmp4* [121]. As illustrated in **Figure 38**, CAR cells as well as ECs express these molecules in various levels.

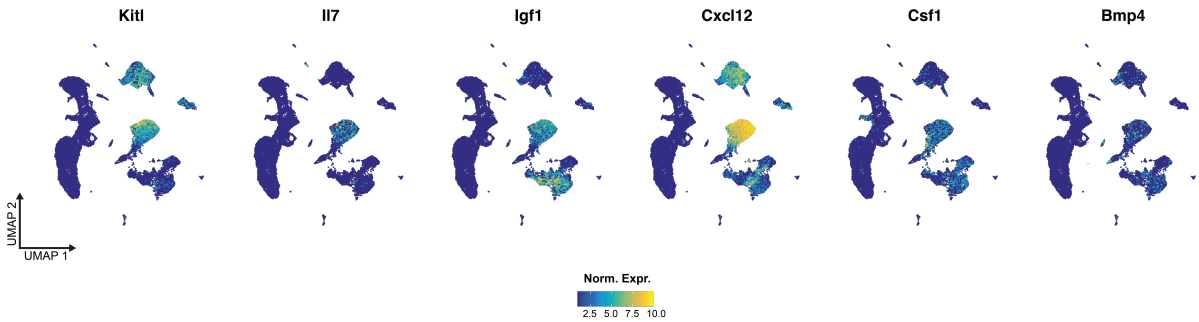

Figure 38 BM stroma cells of NSGW41 mice express genes known to support hematopoiesis. Uniform manifold approximation and projection (UMAP) plots representing the log-normalized expression of selected genes.

2.2.5 BM stroma niche comparison with publicly available C57BL/6J mice datasets

In this study, humanized mouse models were generated by engrafting human AML and HSPCs into NSGW41 recipients. These mice support stable and long-term engraftment of human HSPCs without prior conditioning therapy[62]. In addition, they manifest an increased engraftment of human myeloid cells in the BM and spleen and efficient differentiation of human donor HSPCs into erythroblasts and megakaryocytes in NSGW41 recipients compared to conventional irradiated NSG recipients[62]. NSGW41 mice lack T cell receptor (TCR) on the T cells and immunoglobulin expression in B cells, preventing their efficient maturation. Moreover, the NOD genetic background contains various alleles resulting in the absence of functional macrophages, dendritic cells, and NK cells while myeloid cells are retained. Despite these apparent biological differences, no one has ever evaluated the differences/similarities in the stromal cell composition of NSGW41 mice compared to C57BL/6J mice, which have been previously profiled.

I then conducted a comparative analysis between the BM stromal cell compartment of NSGW41 and C57BL/6J mouse strains, which may be influenced by the absence of functional lymphocytes and other immune populations. To accomplish this, I performed a *LabelTransfer* analysis using Seurat[103], comparing the dataset presented in this Section

with publicly available scRNA-seq BM niche data generated from C57BL/6J mice [26,122,123] (see Methods, Section 4.13.3). A combinatorial analysis of these three datasets was previously performed by Dolgalev and Tikhonova 2021[124]. Briefly, this method integrates two datasets after identifying integration anchors/points between the two[103]. Skeletal muscle cells and Lyve1^{+} ECs were excluded from the analysis since they originate from the outer surface of the bone and were absent from the previously published C57BL/6J of the datasets.

The different stromal cell types were found to be differentially distributed in the 3 datasets, which could be potentially explained by the different dissociation protocols followed in the three studies (**Figure 39 B, C**). This analysis revealed that the unsupervised and manually annotated clusters of NSGW41 mice match the clusters of the same lineages in the C57BL/6J datasets, indicating that genetic background of the strain does not drive large differences in cell state or identity (**Figure 39 D**). Though, the question about abundance differences of the lineages between the two species cannot be addressed using this analysis since dissociation protocols differ between the different studies. Additionally, the age and sex of the mice used between the studies may vary, introducing another possible source of variation.

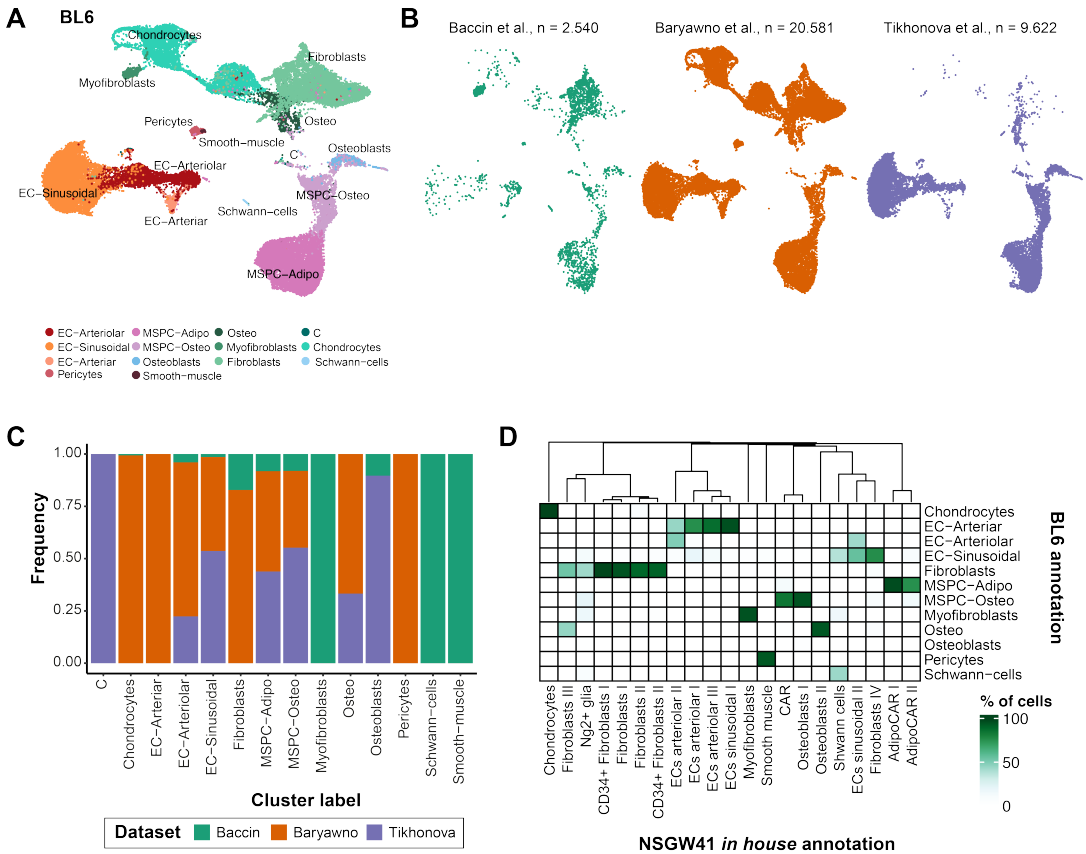

Figure 39 Comparison of the scRNA-seq data of NSGW41 and C57BL/6J strains.

(A) Uniform manifold approximation and projection (UMAP) of 32,743 cells of C57BL/6J mice. Cells are coloured according to cell type. (B) Split UMAPs per dataset. (C) Bar plots indicating the fractions of each dataset in each cell type. (D) Heatmap illustrating the percentage of cells per manually annotated cluster assigned to cluster annotations extracted from the previously published C57BL/6J studies.

2.2.5 The single-cell landscape of engrafted human AML and HSPCs

Ultimately, I annotated the human engrafted CB-CD34$^+$ cells based on known marker genes (**Figure 40 B**)[88]. I identified 9 distinct clusters of myeloid lineages (Neutrophils, Monocytes, pDCs, MDPs), a cluster of LMPPs and several subsets of lymphoid B cell progenitor subsets (pre/proB, small preB cells-sB, Immature B) (**Figure 40 A**). In **Figure 37** it was illustrated that murine HSPCs did not differentiate towards the lymphoid lineages due to the genetic background of the NSGW41 mice used in this study. However, CB-CD34$^+$ HSPCs were able to differentiate towards both myeloid and lymphoid lineages, though lymphoid lineage cells only consisted of B cells (**Figure 40 A**). CB derived HSPCs are able to differentiate towards both myeloid and lymphoid lineages in the presence of appropriate cytokines *in vitro*[125] but they exhibit a slight differentiation bias towards the myeloid lineage, as reported in a previous study[126]. Within the lymphoid lineage, no maturation bias has previously been reported. It should be noted that the differentiation capacity of CB-HSPCs may be influenced by their source, which can pose a challenge to generalising these findings.

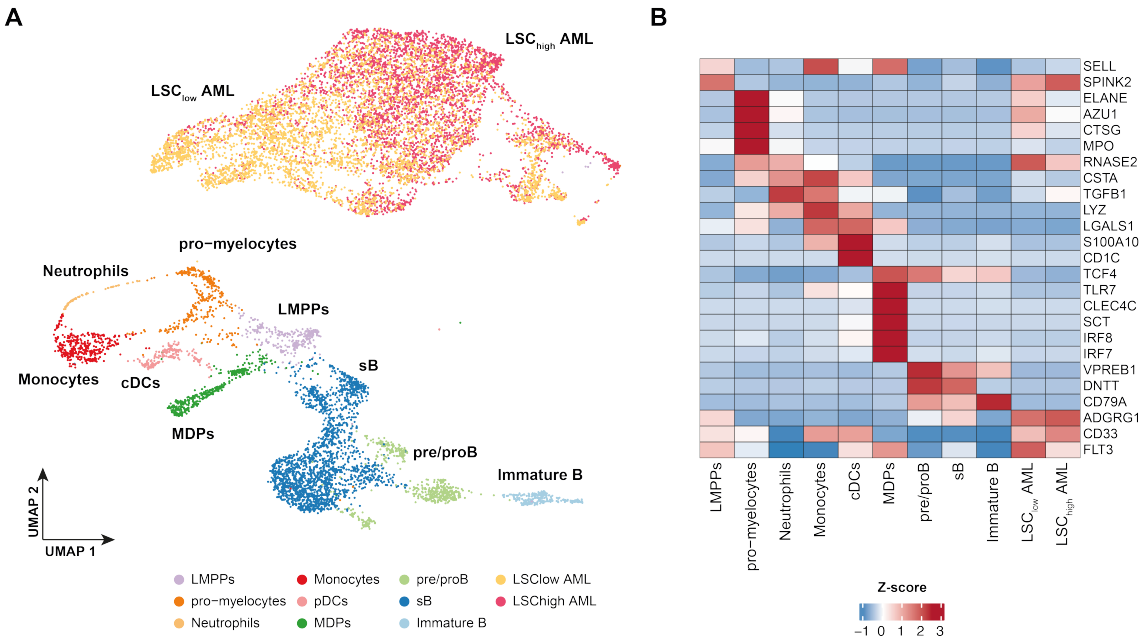

Figure 40 Transcriptional profiling of human CB-CD34$^+$ engrafted into NSGW41 mice.

(A) Uniform manifold approximation and projection (UMAP) of 14,316 post-QC xenotransplanted to NSGW41 mice. Cells are coloured according to cell type. (B) Scaled expression (z-score) of marker genes are shown for each cell type. Expression values were averaged across all cells in the cluster. CB: Cord blood, QC: Quality control.

Altogether, this work enabled the characterization of the single-cell landscape of human engrafted CB-derived HSPCs, as well as AML cells of variable LSC burden, using scRNA-seq. This dataset represents the first of its kind on xenotransplanted cells and it illustrates that this model can be used to sufficiently study differentiation of HSPCs *in vivo*.

2.2.6 Engraftment analysis

To ensure that differences in the BM niche between treatment groups do not originate from variable cell numbers, I initially compared the levels of engraftment between the 3 xenotransplanted groups using flow cytometry. This analysis did not reveal a significant difference in the engraftment of human blood cells between different samples and conditions (**Figure 41 A,** unpaired Student's t-test, p-value > 0.05). Next, the fraction of LSCs was assessed first using flow cytometry analysis of both CD34 and GPR56, two established LSC markers[36, 37] as well as comparing the 17-gene LSC score (LSC17) in the scRNA-seq data[37]. LSC17 is a highly prognostic gene set used for predicting initial therapy resistance[37]. As expected, the percentage of $CD34^+$ $GPR56^+$ double positive cells was higher in LSC_{high} AML (i.e. HLF_{high} AML) relative to LSC_{low} AML (**Figure 41 B**; unpaired Student's t-test, p-value < 0.05). This observation was further corroborated using the LSC17 score, which illustrated that LSC_{high} AML showed higher LSC17 score and subsequently higher LSC burden (**Figure 41 C**; unpaired Student's t-test; p.adj $<$ 2e-16 after Bonferroni correction).

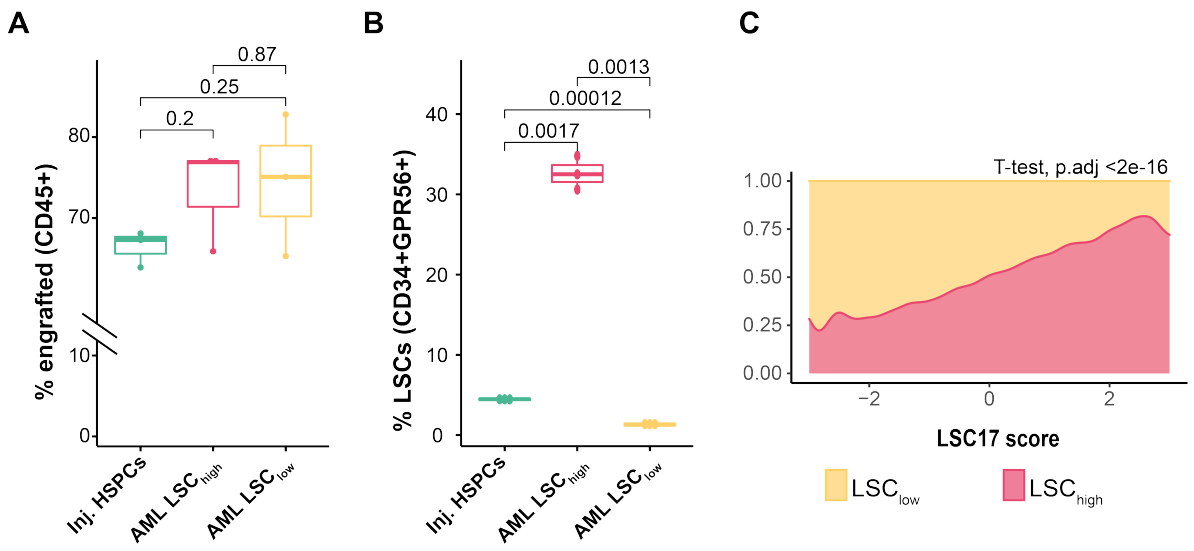

Figure 41 Engraftment analysis of xenotransplanted mice.
(A) Percentage of CD45$^+$ human cells in NSGW41 mice (unpaired Student's t-test, p-value > 0.05). (B) Percentage of LSCs per condition using flow cytometry (GPR56$^+$ CD45$^+$ cells, Middle) and using LSC17 score on scRNA-seq data defined by Ng et al. 2016[37]. The LSC17 score was calculated using Seurat's AddModuleScore function[103]. P-values were calculated using unpaired Student's t-test and in panel (C) p-values were adjusted for multiple testing using Bonferroni correction method.

2.2.7 AML impacts the stromal composition of the bone marrow niche

One way to understand the impact of perturbed cell states is to look into abundance shifts as a response to a biological stimulus, in this case AML and LSC exposure. Building on this concept, I performed pairwise differential abundance analysis between conditions using Fisher's exact test (see Methods, Section 4.13.9).

Multiple studies have reported altered composition of the BM vasculature in AML[46,47]. However, such studies rarely looked into these changes in a global, unsupervised manner. Here, I first examined the impact of transplanting human cells (CB-CD34$^+$ cells) on the EC compartment. These findings showed that the transplantation of human CB derived CD34$^+$ cells have no striking impact on the EC compartment, except for a decrease in the EC arteriolar I cluster (**Figure 42 A**; Fisher's test, log2OR < 0, p.adj < 0.05 after Bonferroni correction). This suggests that arteriolar ECs are more prone to perturbation in presence of xenotransplanted cells. Secondly, I explored the impact of AML and in line with previous reports, sinusoidal ECs were found expanded, while the arteriolar subsets were depleted (**Figure 42 B**; Fisher's test, p.adj < 0.05 after Bonferroni correction)[46,47]. Additionally, LSC burden only impacted the EC

arteriolar I cluster (**Figure 42 C**; Fisher's test, log2OR < 0, p.adj < 0.05 after Bonferroni correction).

Gene set variation analysis (GSVA; see Methods Section 4.13.6) revealed distinct molecular signatures associated with these differentially abundant subsets: ECs sinusoidal I and EC arteriolar I cells enriched in AML, were associated with higher VEGFA signalling, cholesterol transport cell adhesion (**Figure 42 D**). In addition, EC arteriolar II cluster, which was found depleted in AML, showed an increase in several signatures, including hypoxia (**Figure 42 D**). The role of these signatures in AML has been frequently pinpointed. Hypoxia is known to regulate the proliferation of AML cells in the BM[127], while cholesterol modulation results in elimination of AML[128]. Though, no extensive associations between ECs and these signatures have been previously studied, thus highlighting the significance of these findings.

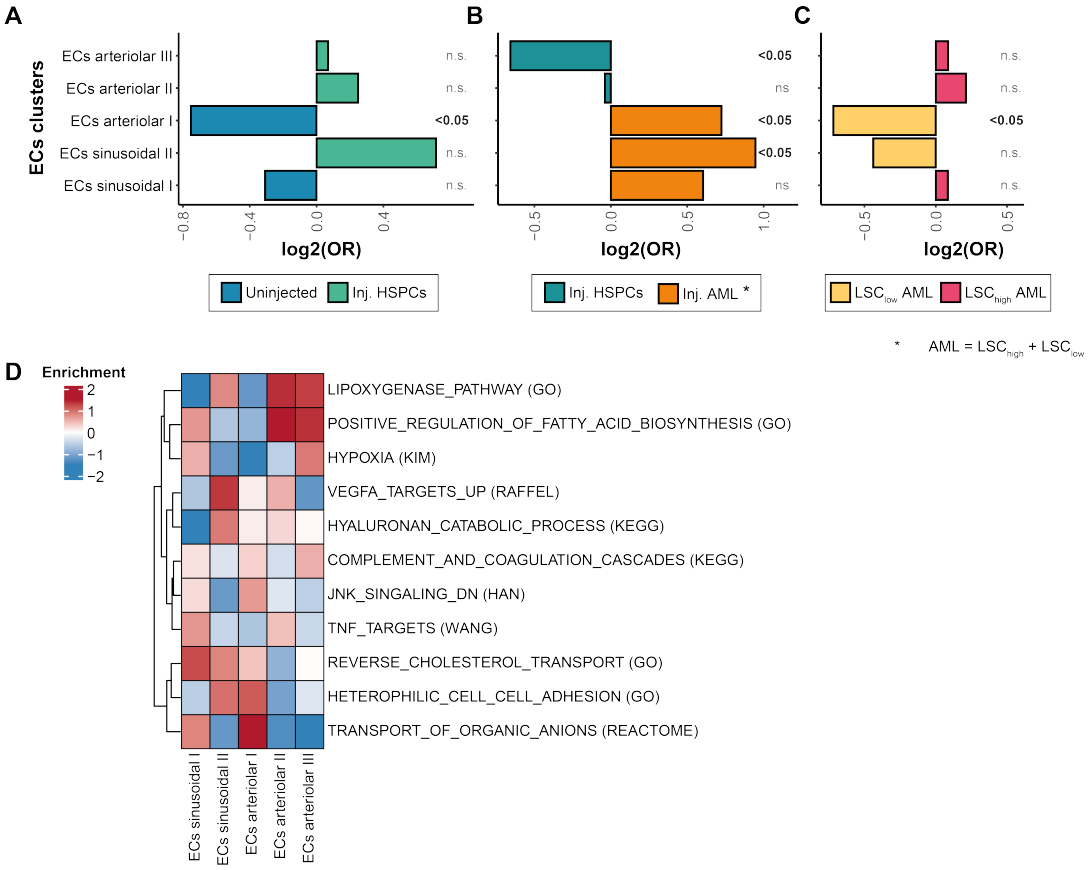

Figure 42 Human AML cells transplanted to NSGW41 mice disrupt the bone marrow vasculature.
Differential abundance analysis of the endothelial cells (ECs) clusters between: non-transplanted group (blue) and group transplanted with CB-CD34$^+$ (cyan) cells (A), group transplanted with CB-CD34$^+$ cells (cyan) and groups transplanted with AML cells (orange) (B) and group transplanted with LSC$_{high}$ (pink) and LSC$_{low}$ (yellow) AML (C). The analysis was performed per cluster using Fisher exact test. Bar plots show the log2(odds ratio) plotted, as

well as the adjusted p-value after Bonferroni correction (n.s.: not significant). (D) EC specific gene sets enriched per cluster selected from the top 20 gene sets extracted from gene set variation analysis (GSVA). Terms used for this analysis were extracted from Kalucka et al. 2020[129]. LSC: Leukemic stem cell.

Fibroblasts are frequently speculated to be highly impacted by AML, as well as support leukemia cell expansion[130]. Thus, I inspected the impact of AML cells and LSCs on the BM resident fibroblasts. AML mainly induced an imbalance between osteo- versus adipo-lineage progenitors. In detail, this analysis revealed an expansion of the AdipoCAR-II cluster and a depletion of the Osteoblasts I cluster to be associated with AML presence (**Figure 43 B**; Fisher's test, p.adj < 0.05 after Bonferroni correction). LSC_{high} AML resulted in further decline of the osteo-lineage since it led to decreased abundance of osteoblasts I (**Figure 43 B**; Fisher's test, p.adj < 0.05 after Bonferroni correction). In addition, higher LSC burden induced the expansion of several Fibroblasts subsets (**Figure 43 C**; $CD34^+$ Fibroblasts I and II, and Fibroblasts II & IV; Fisher's test, p.adj < 0.05 after Bonferroni correction). Of note, LSC burden also affected the abundance of the CAR cluster, which expressed osteo-lineage progenitor genes like *Mmp13* (**Figure 43 C**; $CD34^+$ Fibroblasts I and II, and Fibroblasts II & IV; Fisher's test, p.adj < 0.05 after Bonferroni correction).

Further investigation of selected fibroblast-associated gene sets revealed the enrichment of specific pathways in the stromal subsets enriched in AML and LSC_{high} AML. In detail, all the aforementioned clusters enriched in AML were characterized by increased TGF-β signalling, ECM glycoproteins and IFNγ response, thus suggesting that LSCs remodel the BM resident fibroblasts by inducing changes related to the extracellular matrix (Proteoglycans production, TGF-β) as well as inflammatory responses (IFN response) (**Figure 43 D, E**). Such signatures are indicative of a malignant state in several cancers, including AML[131]. For example, *TGF-β1* produced by bone marrow stromal cells regulate AML cell proliferation, and inhibiting *TGF-β*1 enhances the efficacy of cytarabine chemotherapy[132].

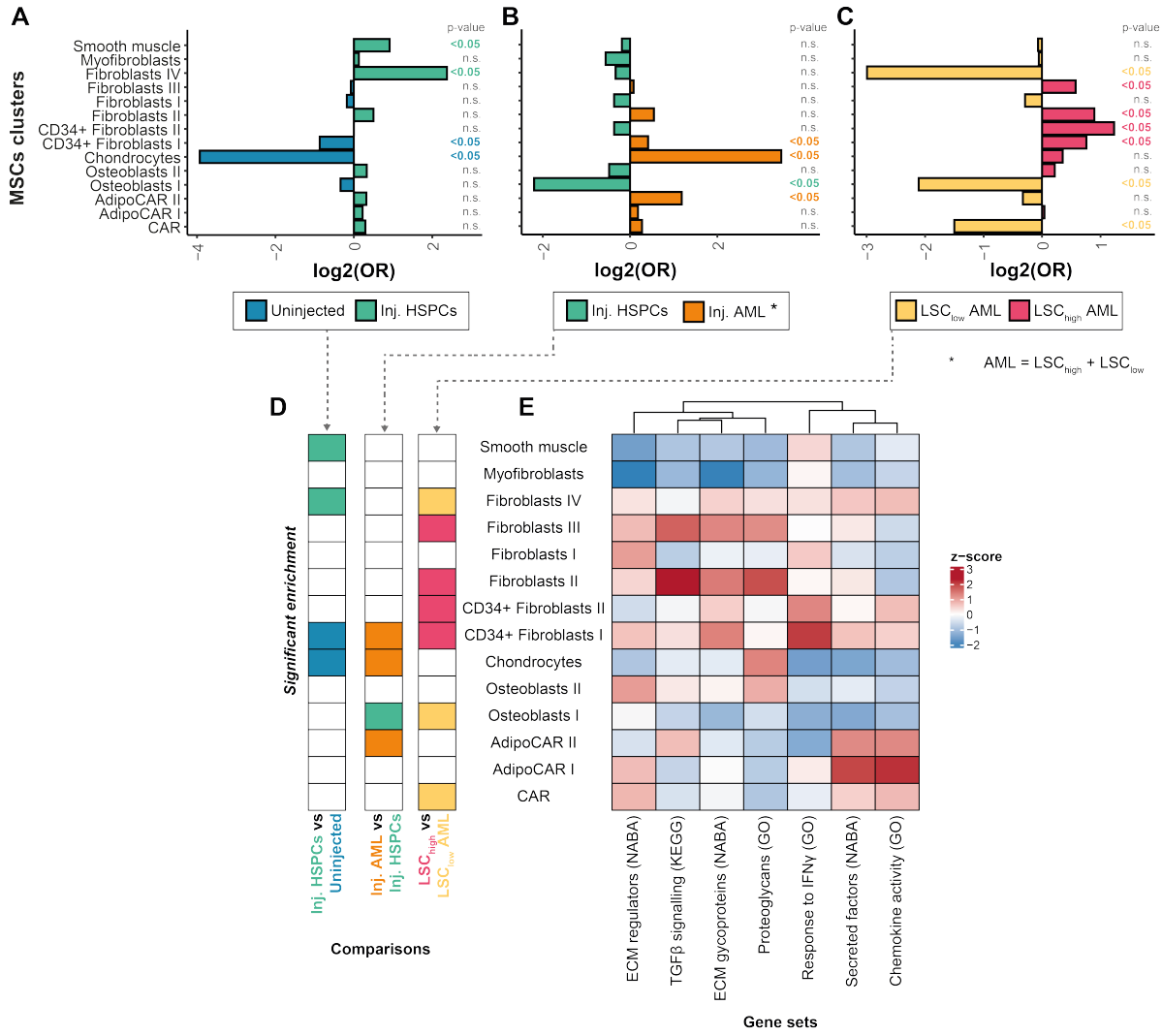

Figure 43 Human AML cells transplanted to NSGW41 mice disrupt the bone marrow fibroblast landscape.
Differential abundance analysis of mesenchymal clusters between: non-transplanted group (blue) and group transplanted with CB-CD34$^+$ (cyan) cells (A), group transplanted with CB-CD34$^+$ cells (cyan) and groups transplanted with AML cells (orange) (B) and group transplanted with LSC$_{high}$ (pink) and LSC$_{low}$ (yellow) AML (C). The analysis was performed per cluster using Fisher exact test. Bar plots show the log2(odds ratio) plotted, as well as the adjusted p-value after Bonferroni correction (n.s.: not significant). (D) Illustration summarising the comparisons from A, B and C. (E) Scaled scores of gene sets selected from multiple databases. Information regarding the origin of each gene set is GO: Gene ontology, KEGG: Kyoto Encyclopaedia of Genes and Genomes, NABA: Naba et al. 2012[133].

Since the primary focus of this study has been the non-hematopoietic microenvironment, no extensive characterisation of the impact of AML on the murine HSPCs was performed. Nevertheless, after comparing the fractions of each HSPC cluster between conditions, it is apparent that all clusters are represented by all conditions (**Figure 44**). However, a slight decrease of erythroid progenitors is observed in both AML

subsets, which suggest the paracrine role of AML cells in the myeloid, and specifically erythroid branch, previously reported in the context of AML[134].

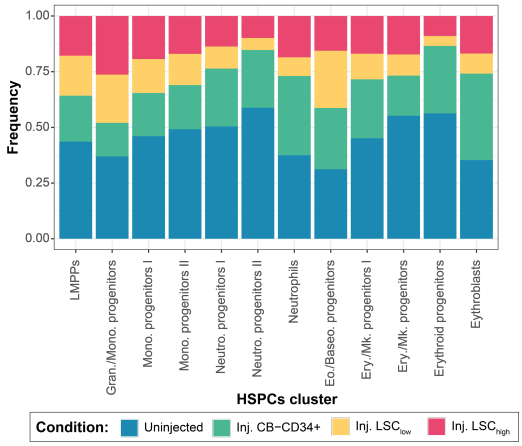

Figure 44 Bar plot depicting the representation of each condition per cluster on the CD34^{+} HSPCs compartment.

In summary, these findings highlight the impact of AML, and specifically LSCs, on endothelial cells comprising the vasculature, and various BM resident cell lineages such as fibroblasts, osteoblasts as well as adipo-lineage progenitors. Furthermore, these differentially abundant clusters were found to be linked with cancer-associated phenotypes, suggesting their potential role in AML development.

2.2.8 *Cell-to-cell* communication analysis reveals potential regulators of lineage skewing

The differential abundance analysis revealed an AML associated lineage skewing of the mesenchymal and endothelial subsets (**Figure 42, 43**), suggesting that paracrine factors secreted from AML cells influence the stromal composition. In line with this, AML is considered to induce remodelling of the BM niche through the secretion of chemokines like *Cxcl12*[130]. In addition, tumour-associated changes in the BM microenvironment in AML are hypothesized to benefit AML cells, since their interplay with neighbouring cells affects their proliferation and survival[130].

However, studies comprehensively mapping which factors induce these changes, especially those secreted by AML LSCs, are lacking. Therefore, I decided to computationally infer the cell-cell interactions occurring between AML cells and the mesenchymal cells of the BM niche. For that, I used *NicheNet*, a computational method which predicts ligand-target genes relationships between interacting cells[135] (see Methods,

Section 4.13.8). In short, *NicheNet* initially integrates prior knowledge from multiple sources on ligand-receptor interactions, signaling pathways and gene regulatory interactions into weighted networks. Using network propagation, *NicheNet* calculates the **regulatory potential** between ligands and potential target genes[135]. Per ligand, the prediction ability of target genes, termed **ligand regulatory score**, is measured as the Pearson correlation coefficient between a ligand-target regulatory potential and the observed transcriptional response. The ligand regulatory score is then used to prioritize inferred the ligands.

Here, two separate cell-cell interaction analyses were conducted, comparing mice transplanted with different groups of human cells. The first analysis compared mice transplanted with AML cells (LSC_{high} + LSC_{low}) to those injected with CB-HSPCs, while the second compared mice transplanted with LSC_{high} AML cells to those transplanted with LSC_{low} AML cells. By examining the inter-cellular interactions between AML cells and mesenchymal cells, several AML signals were identified. Among the top 20 ligands identified to regulate genes in mesenchymal cells, 16 were identified in both comparisons, which suggests a cumulative effect of LSC burden on mesenchymal cells (**Figure 45**). Of note, only the subset of previously experimentally curated ligand-receptor pairs inferred (termed as bonafide interactions, provided by *NicheNet*) was considered for further analysis (**Figure 45**).

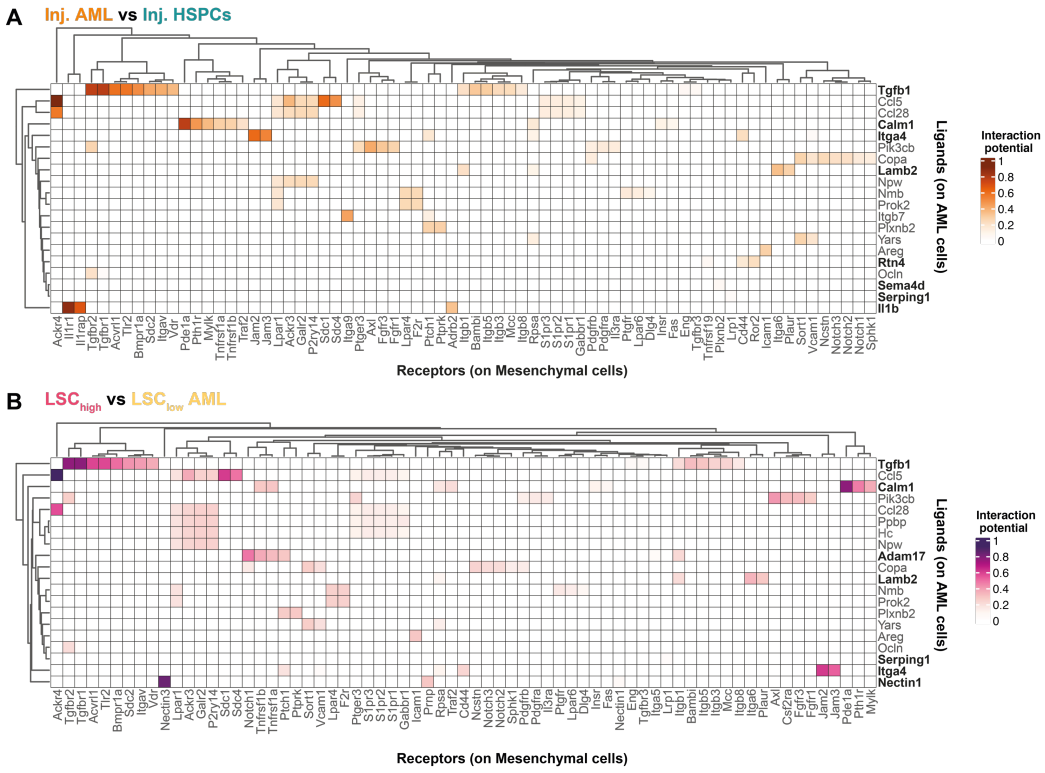

Figure 45 Outcome of NicheNet's ligand-receptor analysis, between the transplanted human cells as sender population and murine mesenchymal cells as receiver.
(A) Ligand-receptor pairs from the comparison of injected AML cells against the mice injected with CB-CD34$^+$ HSPCs. (B) Ligand-receptor pairs from the comparison of injected LSC_{high} AML cells against the mice injected with LSC_{low} AML cells. The bonafide ligands are highlighted in black bold letters. The colour represents Nichenet's prior interaction potential (see Methods, Section 4.13.8).

Among the ligands identified to contribute to the interaction of AML cells with the mesenchymal cells, *TGF-β1, IL-1β* and *CCL5* were detected and characterized by high regulatory potential (**Figure 45**). The expression levels of these ligands were higher in AML cells than CB-CD34$^+$ cells (**Figure 46 A**). Within AML cells, the expression of *TGF-β1* was further increased in cells with a high LSC burden compared to those with a low LSC burden, suggesting a correlation between their expression and LSC burden (**Figure 46 A**). These observations were further corroborated when looking into the fractions of cells positive for these ligands (**Figure 46 B**). Notably, *CCL5* is believed to play a role in protecting AML cells from tyrosine-kinase-inhibitor (TKI)-mediated cell death and contributes to treatment resistance[136]. Even though *CCL5* was considered a promising candidate due to its ability to induce collagen degradation by activating matrix metalloproteinases in fibroblasts[136], it was not selected for further analysis since it was not part of an experimentally verified interaction (bonafide, **Figure 45**).

As a follow-up step, the predicted impact of *IL-1β* and *TGF-β1* secretion from AML cells to mesenchymal cells was further investigated. To do so, *NicheNet's* Pearson correlation coefficient of predicting ability of target gene expression for every ligand was utilized (**Figure 46 C, D**). This analysis suggested that *IL-1β* and *TGF-β1* secretion from AML cells regulates the expression of *Tnfsf11, Junb, Tnfaip6* in mesenchymal cells, suggesting their role in TNF-α/NF-κB signalling (**Figure 46 C**). When comparing LSC_{high} versus LSC_{low} AML, TGF-β1 was found to regulate *Tagln, Serpine1, Cdh2* expression on mesenchymal cells exposed to LSC_{high} AML, indicating a link between LSC mediated TGF-β signalling and ECM remodelling (**Figure 46 D**). In addition, a link between *TGF-β1* and *Pdgfra/Pdgfrb* expression was identified, providing a potential link between TGF-β signalling and the fibrotic phenotype reported in Section 2.2.7 (**Figure 46 D**).

In humans, *IL-1β* is responsible for initiating inflammatory processes, and the therapeutic effect of blocking *IL-1β* activity has been explored in AML clinical trials [137]. *IL-1β* causes expansion of myeloblasts while inhibiting the growth of normal progenitors[138]. Moreover, elevated levels of *IL-1β* receptors were previously observed in patients with AML[138]. Aberrant TGF-β signalling has been implicated in hematopoiesis as well as leukemia development. The proliferation of hematopoietic cells is regulated by

the binding of *TGF-β1* to cell surface receptors through autocrine and paracrine mechanisms[139]. These receptors have been identified to be abnormally expressed in leukemia cells[140]. In AML, *TGF-β1* is overproduced by megakaryocytes and inhibits normal HSC proliferation[141]. While some studies have reported that *TGF-β1* expression is reduced in AML patients compared to healthy individuals, the evidence still remains inconclusive[139]. Therefore, it is important to consider the specific context and findings when discussing the expression levels of *TGF-β1* in AML and its impact on the microenvironment.

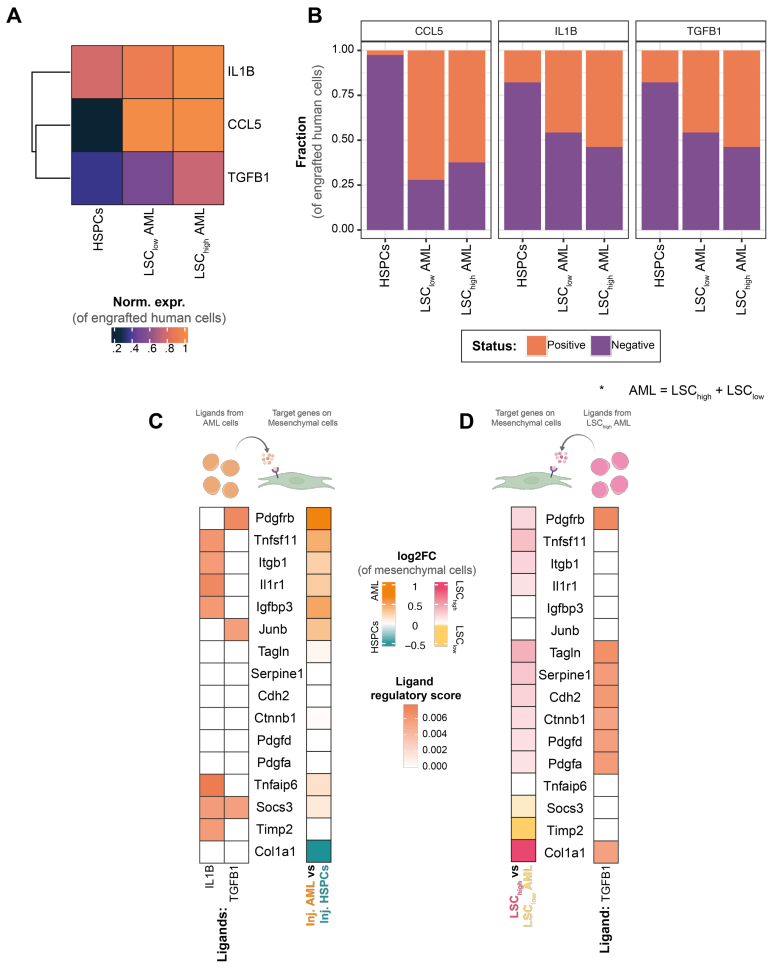

Figure 46 *IL-1β*, *TGF-β1* and *CCL5* may induce alterations on mesenchymal cells in the context of AML.
(A) Log normalized expression of the three ligands identified from NicheNet analysis on human engrafted cells. (B) Fractions of positive and negative cells for these ligands across the three conditions tested. (C) (Left) Heatmap depicting the ligand-target gene regulatory potential calculated from NicheNet. (Right) Log2 fold change (FC) of target genes on mesenchymal cells from the comparison performed between mice transplanted with AML as well as mice transplanted with HSPCs. (D) (Right) Heatmap depicting the ligand-target gene regulatory potential calculated from NicheNet. (Left) Log2FC of target genes on mesenchymal cells from the

comparison performed between mice transplanted with LSC_{high} AML and mice transplanted with LSC_{low} AML.

Collectively, these results reveal the complex landscape of cellular interactions between AML and the stromal microenvironment, potentially contributing to the altered microenvironment reported in AML patients.

2.2.9 Transcriptional changes of bone versus marrow resident AML cells

I then explored whether the localisation of AML cells in the bone lining versus the marrow might impact the phenotype of AML cells. Therefore, I performed differential expression analysis using MAST algorithm[104], between AML cells residing in the bone lining versus the BM. This assignment was done based on the information from the hashing antibodies. As illustrated in Section 2.2.3, cells harvested from the bone were incubated and tagged with different antibodies than those from the BM, thus allowing the effective identification of sample of origin (**Figure 30, 34**; see Methods, Section 4.10). I identified 19 and 15 genes in LSC_{high} and LSC_{low} AML groups, respectively, that were differentially expressed depending if the cells were localized to the bone lining or marrow (absolute log2FC $>$ 0.5, p.adj $<$ 0.05 after Bonferroni correction) (**Figure 47 A, B**). Interestingly, the log2FC of these analyses showed high correlation (**Figure 47 C**; Pearson correlation, 0.81), suggesting that these signatures are independent of the LSC burden.

Based on this analysis, several genes previously associated with AML tumorigenesis and poor disease prognosis were differentially expressed, including *CD69, JUN, FOS, CXCL2 and CXCL8*. Specifically, high *CD69* expression in human LSCs were characterized by adequate self-renewal[142]. JUN has been previously reported to support AML through the regulation of the unfolded protein response (UPR) in AML cells, through the cooperation with ATF3 which supports AML cells[143]. Amongst the genes associated with the bone lining, several chemokines were detected as well. *CXCL2* benefits AML cells while in a hypoxic environment, which is the case for the bone [144]. Moreover, *CXCL2* together with *IL-1β* increases the phosphorylation and subsequent activation of GATA-2 TF, which regulates the expression of genes responsible for proliferation[145]. *CXCL8/IL-8* and its receptors (*CXCR1, CXCR2*) have been implicated to contribute to several inflammatory diseases, through the regulation of disease associated processes like fibrosis and tumorigenesis[145]. Specifically, in AML *CXCL8/IL-8* is upregulated and induces cell growth through the activation of ERK1/2 signalling pathway[146].

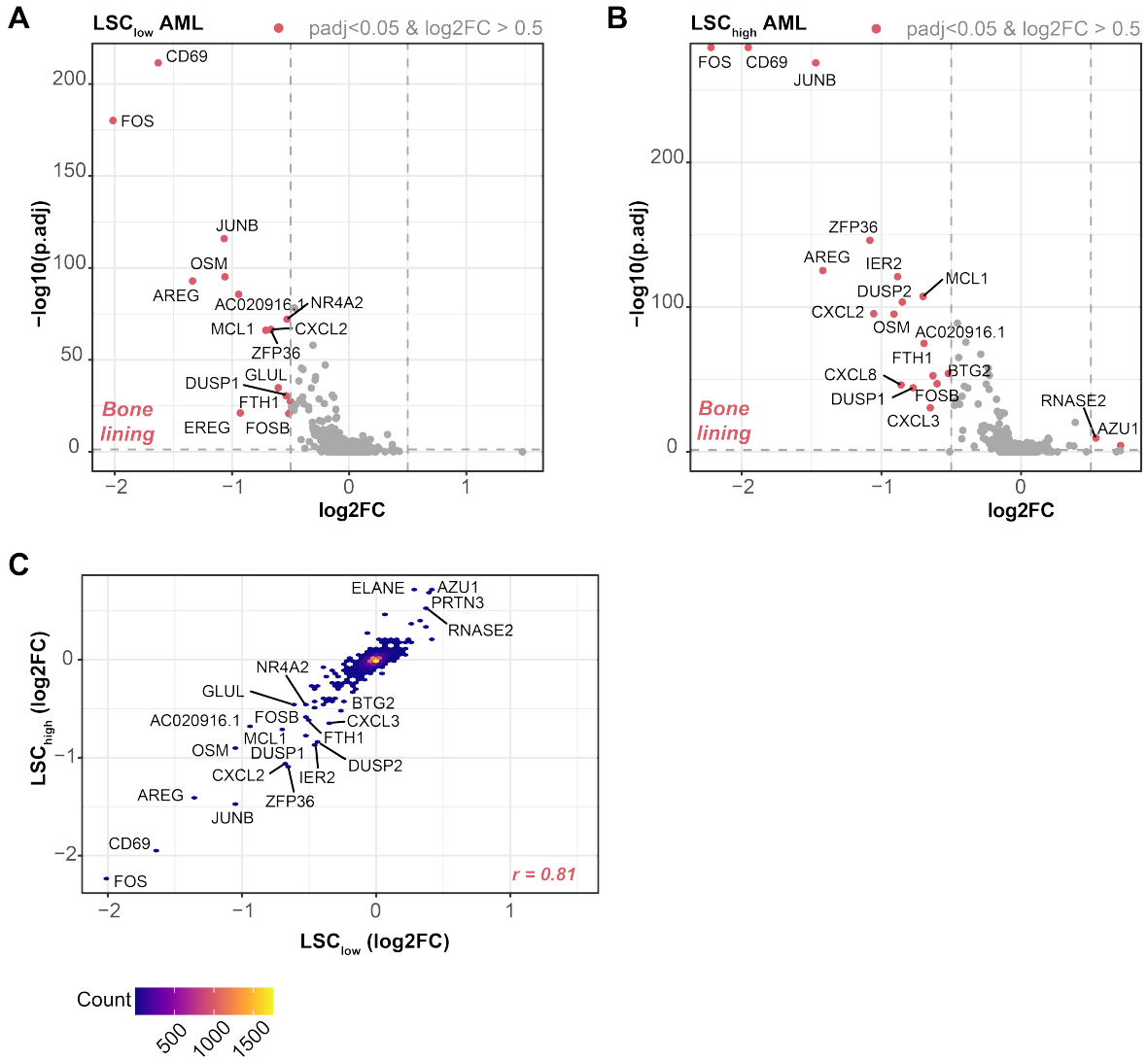

Figure 47 Differential expression analysis of AML cells residing in the bone versus the marrow. Volcano plot depicting the differentially expressed genes between bone lining and bone marrow in LSC_{low} AML cells (A) and LSC_{high} AML cells (B). Y axis denotes the p.adj (p-values were adjusted for multiple comparisons using the Bonferroni correction method), the horizontal dotted lines represent p.adj = 0.05. X axis denotes the log2FC, with the vertical dotted lines representing the absolute log2FC = 1. (C) Scatter plot displaying the log2FC between bone and marrow in the 2 different AML conditions, Pearson correlation r = 0.81.

Overall, this analysis indicate that AML cells in the BM display distinct transcriptional profiles depending on their surrounding microenvironment. AML cells in the bone lining were characterized by elevated expression of genes associated with AML self-renewal, support as well as poor disease prognosis. These findings highlight the importance of studying cancer cells while accounting for the microenvironment.

2.2.10 Distinct localization of HSPC-derived subsets in the bone

I next sought to investigate whether HSPC-derived hematopoietic subsets localize to different regions of the bone. One of the key regulators of HSPCs localisation in the bone is oxygen gradient, since HSPCs are mainly located in the hypoxic regions of the bones[148]. Moreover, stromal cells are mainly located in the bone lining and form specialized niches through the production of cytokines like *IL7* [117]*, CXCL12* [119] *and SCF*[120]*.* The secretion of such molecules leads to their interplay with the hematopoietic subsets.

While several studies have addressed HSPC localisation in the different niches, often using live imaging methods[15], little is known regarding the distinct localisation of different HSPCs maturation stages. To explore this question, I performed differential expression analysis of CB-CD34$^+$ HSPCs in the bone lining versus the marrow. When looking into the human derived HSPCs and their localisation in the bone, differential distribution of the distinct clusters was detected.

Differential expression analysis of the HSPCs revealed overexpression of neutrophil markers (*S100A8*, *S100A9*, *S100A12*) in the bone lining compared to the marrow (**Figure 48 A**). A comparable pattern was observed when comparing the fraction of cells per cluster between the two niches (**Figure 48 B**). Neutrophils have been reported to reside in the BM in quiescent reserves adjacent to blood vessels, poised to be mobilized into the bloodstream since they are the first responders to inflammation[147].

Additionally, HSPCs residing in the bone lining overexpressed B cell markers, including *IGLC1, IGLC2* and *IGKC* (**Figure 48 A**), though when looking into the fractions of the different B cell progenitors the results were ambiguous since Immature B cells were enriched in the bone lining while small preB cells (sB) were enriched in the marrow (**Figure 48 B**). Even though the understanding of BM niches in the context of B cell development has significantly improved using cell-deletion studies and microscopy, such studies have been limited to one specific cell type and may not have taken into account cellular ecosystems. Therefore, interpreting these findings is challenging due to the lack of established knowledge in the existing literature. In particular, the extravascular compartment around the vascular sinusoids in the BM has been reported to be enriched in B cells and play a critical role in positive selection of B cells[148]. Early stages of B cell maturation require osteoblasts and CAR cells, while later maturation stages are promoted by *IL-7* secreting stromal cells as well as sinusoidal ECs. Depletion of osteoblasts leads to decrease of pre-pro-B and pro-B cells[16]. On the other hand, promyelocytes and LMPPs were enriched in the marrow (**Figure 48 B**).

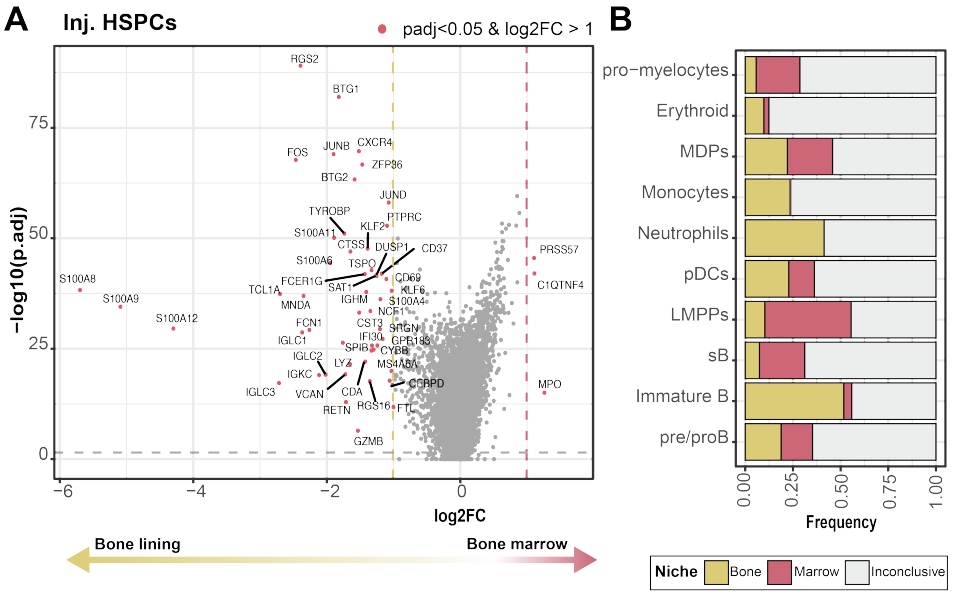

Figure 48 HSPC distribution in the bone.
(A) Volcano plot depicting the differentially expressed genes between hematopoietic stem and progenitor cells (HSPCs) residing in the bone lining versus those in the bone marrow. Y axis denotes the p.adj (p-values were adjusted for multiple comparisons using the Bonferroni correction method), the horizontal dotted lines represent p.adj = 0.05. X axis denotes the log2FC, with the vertical dotted lines representing the absolute log2FC = 1. (B) For every HSPCs cluster, the fraction of their localisation in the bone lining in comparison to the marrow is illustrated.

Overall, the distinct localization of HSPCs within the BM suggested by this analysis, highlights the importance of the BM microenvironment in regulating hematopoiesis as well as the maintenance of the immune system.

2.2.11 Effect of LSC_{high} and LSC_{low} AML cells on huMSCs

In Section 2.2.7 it was suggested that high LSC burden in AML may hinder the development of the osteo-lineage while simultaneously promote fibrotic signatures in the BM microenvironment. While the NSGW41 mouse models are appropriate for studying human hematopoiesis, evident by the expression of molecules which support hematopoiesis in the BM niche (**Figure 38**), they cannot be used to interrogate the direct interactions between human stromal cells and human tumors. Therefore, I wanted to validate these findings by examining the impact of LSCs on human derived MSCs using *in vitro* co-cultures. MSCs are multipotent and can differentiate into various cell types including osteoblasts, adipocytes and fibroblasts[11].

MSCs were isolated from primary BM aspirates of 4 AML patients after allogeneic stem cell transplantation (alloSCT; see Methods, Section 4.7). MSCs were then co-

cultured for 48 hours with either LSC_{high} or LSC_{low} AML lines prior to scRNA-seq using the 10X Genomics platform (**Figure 49**).

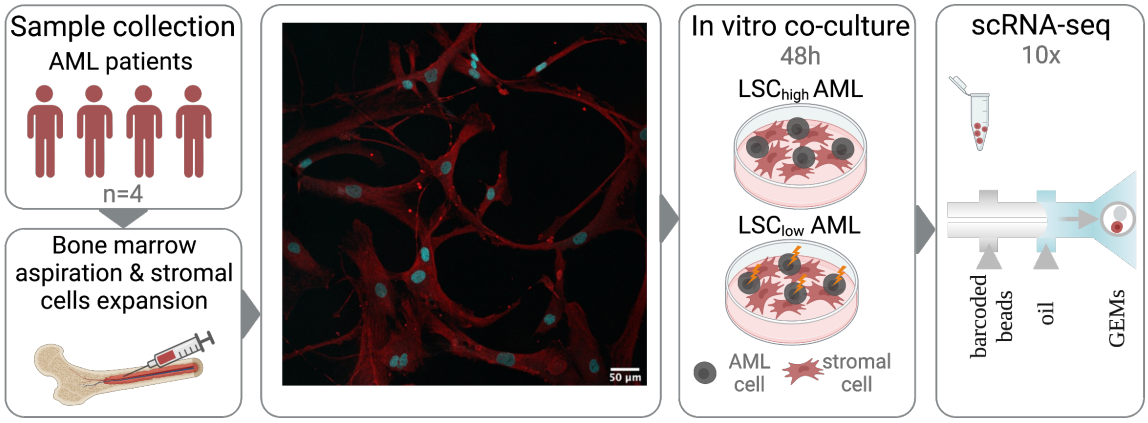

Figure 49 Schematic overview of in vitro co-cultures of BM derived MSCs with LSC_{high} and LSC_{low} AML cells.
Bone marrow aspirates were collected from 4 AML patients post alloSCT and cells were put in culture after processing (See Methods, Section 4.7). Non-adherent cells were removed while adherent cells were passaged for 3-4 passages prior to co-culture with AML cells. AML-MSCs co-culture lasted 48h and MSCs were then subjected to scRNA-seq using the 10x Genomics platform. (Middle) Immunofluorescence staining for CD90 (red) and DAPI (cyan) on primary mesenchymal stromal cells mono-cultured for 5 days. Scale bar 50µm.

This dataset consisted of 11,555 single cells, with an average number of 1,556 cells per patient and 6,146 cells per condition. Differential expression analysis between the two conditions revealed 46 DE genes (MAST [104]; p.adj < 0.05 after Bonferroni correction and log2FC > 0.25). Among the genes differentially expressed, cell migration-inducing and hyaluronan-binding protein (*CEMIP*) and matrix metalloproteinase 1 and 3 (*MMP1, MMP3)* were found upregulated on MSCs co-cultured with LSC_{high} AML. These proteins are known to degrade the extracellular matrix, thus remodelling the extracellular matrix environment[149]. Notably, MMPs have been implicated in multiple cancers, including AML[150]. These findings are in line from the in vivo observations from Section 2.2.7, which show that the BM of mice transplanted with LSC_{high} AML is enriched with remodelled fibroblasts, associated with gene sets like TGF-β signalling and ECM glycoproteins.

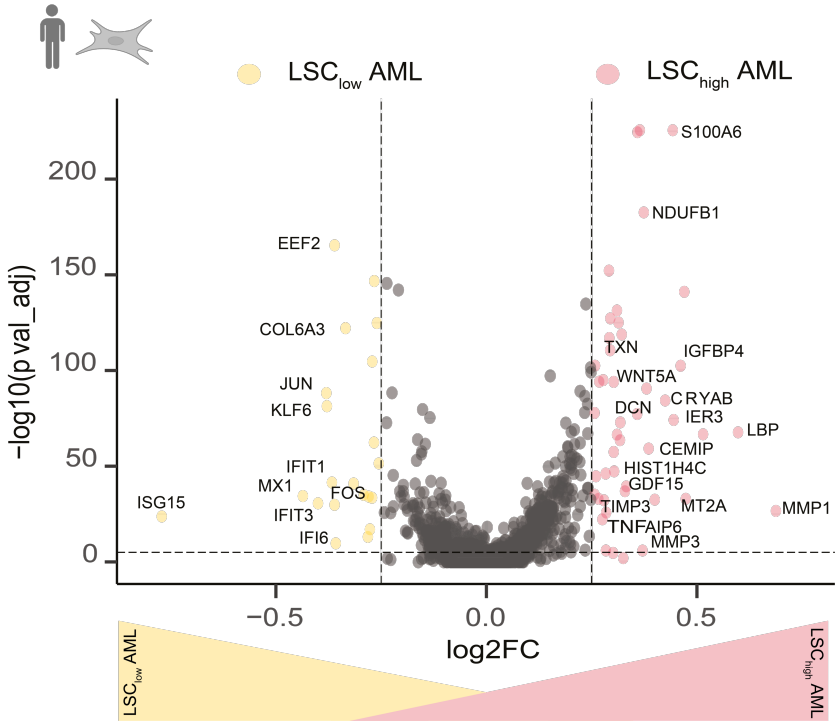

Figure 50 Impact of LSCs on huMSCs *in vitro*.
Volcano plot illustrating the differentially expressed genes of mesenchymal stromal cells (MSCs) co-cultured with LSC_{low} (yellow) and LSC_{high} AML (pink). Horizontal dotted line represent p.adj = 0.05 (p-values were adjusted for multiple comparisons using the Bonferroni correction method) and vertical dotted line represent absolute log2FC = 0.2.

2.2.11 Synopsis

Ultimately, the observations presented this Section provide a holistic view on the AML induced remodeling of the BM microenvironment, with respect to LSCs. This was accomplished by analyzing the expression patterns of thousands of individual cells using scRNA-seq, which enable the high-resolution mapping of rare BM resident populations.

By mapping these cells, I was able to detect phenotypic and compositional changes to be associated with increased LSC burden. Specifically, I observed an expansion of the adipo-lineage progenitors, decline of the osteoblasts, disruption of the vasculature and altered fibrotic landscape. These observations provide a deeper understanding of the LSC-induced pathological characteristics of the BM niche.

Section 3.2 will provide a more detailed discussion of the findings presented in Section 2.2.

CHAPTER 3: DISCUSSION

The results presented in this thesis broaden the understanding of the interplay between acute myeloid leukemia (AML) and two main components of the bone marrow (BM) microenvironment; the immune (Section 2.1) as well as the non-hematopoietic, stromal microenvironment (Section 2.2).

In Chapter 3, I further investigate these findings and elaborate on their broader implications in AML immunotherapy, the underlying molecular mechanisms of successful therapy, the role of the tumor microenvironment and its relevance not only in AML but also in other cancers.

3.1 The remission status of AML patients post alloSCT is associated with a distinct single cell signature of bone marrow T cells

Allogeneic stem cell transplantation (alloSCT) is a highly efficient immunotherapy and is often the main curative approach for AML patients. The success of this therapy strategy relies on the ability of transplanted immune cells, particularly T cells, to recognize and attack leukemia cells. However, patients often relapse potentially due to the failure of the graft immune cells to recognize AML. The identification of graft-versus-leukemia (GvL) activity in individual patients is challenging and requires monitoring of immune cell function as well as the leukemia burden. After alloSCT, T cell reconstitution is a critical step for establishing effective GvL activity. The hypothesis of this study was that the reconstitution of specific T cell subsets, may be associated with better clinical outcomes in AML patients. To investigate this hypothesis, I performed single-cell RNA-sequencing (scRNA-seq) of bone-marrow residing T and $CD34^+$ cells sorted from day +100 bone marrow aspirates of AML patients who stayed in complete remission (CR) or suffered relapse (REL) shortly after sampling.

This study established a reference map of donor-derived bone marrow T cells post alloSCT and proposes G protein–coupled receptor 56 (GPR56) dynamics as a surrogate for the extent of antigen encounters post alloSCT.

3.1.1 BM reconstitution after alloSCT and therapy outcome

Following alloSCT, HSPCs and T cells mature and repopulate the BM. The recovery of immune subsets following SCT differs, with innate immunity cells recovering earlier than T and B cells, which can take up to two years for full reconstitution[57]. Other post-transplantation factors, such as the use of immunosuppressive therapies, contribute to the delayed recovery of the immune system, which may result in relapse[57].

Even though there is an understanding in T cell reconstitution after alloSCT, there have not been many studies utilizing scRNA-seq, an unsupervised method, which could enhance our understanding in this field. Recently, using scRNA-seq researchers mapped the donors' HSPCs reconstitution after alloSCT, revealing that already one day after alloSCT, HSPCs illustrated a bias towards erythrocyte/megakaryocytic lineages[151]. In addition, a subset of neutrophil progenitors expressing elevated levels of S100 genes has been linked to a lower risk of developing acute GVHD, indicating the importance of studying HSPCs dynamics after alloSCT[151].

The results of this thesis expand our current understanding of BM reconstitution after alloSCT, since a holistic approach was followed in order to map the single cell landscape 100 days post alloSCT. It is shown that both lymphoid (preB, pre/proT) and myeloid lineage progenitors exist, with the latter being consisted of both megakaryocyte-erythrocyte progenitors (MEP) and monocyte-dendritic progenitors, suggesting that the previously reported erythrocyte/megakaryocytic lineage bias[151] is rebalanced 100 days after alloSCT. In addition, T cells were differentiated towards both $CD4^+$ and $CD8^+$ T cells. Of note, $CD4^+$ T cells were observed to be less mature and harder to assign to known cell types, potentially due to the delayed reconstitution of $CD4^+$ T cells.

After the characterisation of the BM landscape in alloSCT, further exploration of the composition between complete remission (CR) and relapse (REL) patient samples illustrated the association of altered population abundances with therapy outcomes.

Initially, the presented analysis suggested that altered BM composition may be associated with therapy outcome. In short, relapse was associated with increased naive T cell and Treg cells, as well as a decline of highly cytotoxic subsets, including $CD8^+$ effector clusters and $\gamma\delta$T cells. Relapse samples showed a decrease in MEP and the most immature MLP among donor-derived HSPCs. This decrease occurred despite no higher blast infiltration being observed at the time of sampling, which indicates that healthy hematopoiesis was hindered by competition for specific niches or paracrine signalling from rare AML cells. This phenomenon has been documented in other studies as well[152]. At the time of sampling two of the CR patients had incomplete donor chimerism, and one was minimal residual disease (MRD) positive for NPM1 mutation. However, these patients were eventually MRD negative with full donor chimerism, indicating that the GvL effect, rather than preceding chemotherapies, was responsible for preventing relapse.

An association between higher $\gamma\delta$T cell content and better outcome post alloSCT has previously been described[155]. However, in that case the $\gamma\delta$T cell content was estimated in the grafts prior to transplantation, while in this thesis their presence was detected after three months of contact with immunosuppression and patient cells.

Clinical trials with γδT cells as salvage therapy are ongoing (e.g. NCT03790072), but their safety and efficacy as cellular therapy remain to be determined. Despite γδT cells' significant role in alloSCT, there is a lack of studies mapping their reconstitution after alloSCT, with limited insights associating them with GVHD[153]. Therefore, this thesis, which successfully profiled them using scRNA-seq, is important for gaining a fundamental understanding of their role.

3.1.2 Monitoring alloSCT therapy outcome using BM gene signatures

Additionally, the results presented in this thesis highlight gene signatures, which could serve as biomarkers for monitoring therapy outcome. Interestingly, some of these signatures were specific to particular cell subsets, while others were shared across all subsets suggesting a systemic effect.

The relapse-associated T cell signatures were defined by changes in gene expression and transcription factor (TF) activity. These signatures were found to be enriched for inflammatory signaling pathways such as TNF-alpha (FOS/JUN[154,155]) and NF-κB signalling (REL, NF-κB1/NF-κB2, RELB), as well as an immunosuppressive microenvironment (*CREM*[105]). In line, NF-kB has previously been associated with dysfunctional T cells in renal cell carcinoma[156]. TFs driving these pathways were found more active in REL samples in all cell types ($CD4^+$, $CD8^+$ and HSPCs) but with distinct regulons, indicating that a general pro-inflammatory milieu might exist in the bone marrow of these patients potentially hampering GvL and favouring leukemia promoting clues. *STAT1*, *IRF1* and *ELF1* were found to be more active in REL $CD8^+$ T cells and HSPCs and regulate IFN response genes, suggesting a generalisable inflammatory response to correspond to poor outcome samples. Inflammatory signatures in the bone marrow have been previously associated with poor survival in AML patients[157]. Importantly, the identification of such signatures can be utilized to further stratify patients who may need alternative treatment, since inflammation is critical for several aspects of AML, including chemoresistance[158].

In contrast, CR $CD8^+$ T cells were mainly characterized by a highly cytotoxic gene signature, potentially regulated at a transcriptional level by the TF *TBX21*. Among the target genes of *TBX21*, the adhesion GPCR GPR56 was detected. In addition to its implications in lymphocyte cytotoxicity[159], GPR56's presence as a cell surface molecule indicating its potential for further clinical research. One of the advantages of focusing on specific surface molecules for clinical research is the ease of measurement through flow cytometry.

3.1.3 Discrepancies between RNA and protein

Single-cell technologies have revolutionized our understanding of cell identity, diversity, and function. However, the transcriptome does not always correlate well with the proteome, and many genes show a poor correlation between RNA and protein levels[160]. This discrepancy may arise from multiple factors, including post-transcriptional regulation, protein turnover, and technical noise in scRNA-seq. Therefore, it is important to validate scRNA-seq results with protein-level measurements using techniques such as flow cytometry or immunohistochemistry. Combining RNA and protein data is critical for accurately characterising cell phenotypes and functions, and for translating scRNA-seq findings into clinical applications.

In the present thesis, scRNA-seq analysis revealed *ADGRG1*/GPR56 at the RNA level to be associated with therapy outcome. To further corroborate these findings, GPR56 levels were analysed at the protein level using flow cytometry analysis, which revealed similar patterns between both modalities.

Protein levels directly reflect cellular functions, since proteins are the main effectors in cellular processes. A direct way of studying these discrepancies between RNA and protein is CITE-seq (cellular indexing of transcriptomes and epitopes by sequencing), which is a single cell proteomics technique that combines transcriptome sequencing with the detection of cell surface proteins or other epitopes using oligonucleotide-labelled antibodies[77]. In addition, single-cell proteomics using mass spectrometry (MS) based methods have made significant advancements in the past years, allowing for the quantification of proteins at the single-cell level. Single-cell proteomics still faces several limitations, including technical challenges such as low throughput and high costs, as well as difficulties in obtaining high-quality single cells and accurately quantifying low-abundance proteins.

3.1.4 Exploring the mechanism and clinical implications of GPR56

One of the signatures that was found to be associated with complete remission following alloSCT was GPR56 expression on $CD8^+$ effector memory T cells. GPR56 is known to be lowly expressed in the lymphocytes of murine models, hence making mechanistic studies challenging[159]. Though, various studies have identified its association with specific functions. Combining these findings with the results presented in this thesis represents an important first step in understanding its underlying mechanism.

GPR56, which is encoded by the *ADGRG1* gene, is a typical G protein coupled receptor, known to act together with CD81[164]. It is involved in several developmental

processes, including those in the central nervous system, male reproductive system, the immune system as well as tumorigenesis[159]. In addition, GPR56 has been previously identified as an LSC marker in AML[36], where its expression on leukemic blasts is associated with poor outcome.

There have been several studies of GPR56 in the context of the immune system, though such studies have been mainly conducted in murine models. Here, gene regulatory network analysis suggested TBX21 as a potential TF regulating *ADGRG1*/GPR56 expression, as well as several cytotoxic molecules. To this date, no direct link has been made between TBX21 and *ADGRG1*. The HOBIT/*ZNF683* TF, which according to literature is expressed in terminally differentiated lymphocytes, has been previously proposed to regulate GPR56 expression in human NK cells[112]. HOBIT is expressed in $CD4^+$ and $CD8^+$ T cells[112]. Knockdown of *ZNF683* resulted in GPR56 downregulation in NK-92 cells, while ectopic expression of HOBIT induced GPR56 expression in Jurkat cells[112]. In addition, TGF-β signalling and HOBIT induce the expression of GPR56 in cytotoxic lymphocytes[159]. In the present thesis, HOBIT was detected as a marker of relapse enriched $CD8^+$ hobit subset which expressed low levels of *ADGRG1*, as well as CR enriched $CD8^+$ mem. 2 cluster which expressed high levels of *ADGRG1*. Gene regulatory network analysis did not reveal any direct connection between Hobit and *ADGRG1*. However, it is unclear whether this observation is due to biological factors or limitations of the method used to infer gene regulatory networks[85]. *SCENIC* infers TF to gene connections based on co-expression networks and upstream TF motif analysis. Therefore, it is possible that a connection may not appear due to technical limitations such as insufficient expression levels of a given TF.

The exact role and mechanism of GPR56 in T lymphocytes is still under investigation[161,162]. Here, I showed that GPR56 is mainly expressed in $CD8^+$ and $CD4^+$ TEMRA cells, which is in line with previous studies[111]. In addition, at both RNA and protein level GPR56 is co-expressed with cytotoxic molecules like GZMB and PRF1, but not classical exhaustion markers like PD-1. In addition, in collaboration with the Schmidt group, me and my colleagues present evidence that GPR56 expression on T cells occurs only upon antigen encounter in a CD33-directed CAR-T/HL-60 model, making GPR56 an excellent candidate for marking alloreactivity. However, in order to directly associate GPR56 with higher killing potential further mechanistic studies are needed. One possible strategy would be to co-culture AML cells with either $GPR56^+$ or $GPR56^-$ $CD8^+$ T cells and compare the fraction of remaining AML cells between these two conditions.

The role of GPR56 has been studied in various cell types. GPR56 has been implicated in the control of neural progenitors and melanoma cell migration[163] and is expressed in microglia, which are myeloid lineage derived macrophages of the central nervous system[164]. Knockout mice with a conditional deletion of microglial *Adgrg1* showed increased TNF production in microglia[164]. In line with these previous findings, a TNF signalling signature was detected to be higher in REL across all cell lineages, while *ADGRG1*/GPR56 was downregulated. In addition, GPR56 overexpression has been proposed to reduce the migratory capacity of NK-92 and primary T cells[111]. Mechanistically, it has been found that GPR56 enhances reciprocally inhibiting signalling pathways and thus drives both, a highly proliferative $Wnt^{high}TGF\text{-}\beta^{low}$ and a stem cell-enriched, slowly cycling, $Wnt^{low}TGF\text{-}\beta^{high}$ compartment, thus maintaining self-renewal and preventing exhaustion of the LSC pool[165]. Whether GPR56 functions in a similar way in T cells to maintain self-renewal, would have to be addressed in functional T cell studies.

In order to substantiate clinical claims of these findings, analysis in larger patient cohorts is necessary. Flow cytometry analysis performed on an independent cohort of 139 AML patients revealed the temporal progression of GPR56 levels after alloSCT. This analysis showed that the fraction of $GPR56^+$ $CD4^+$ and $CD8^+$ T cells increases after transplantation, beyond the levels of non-transplanted individuals, suggesting GPR56 as a hallmark of the alloreactivity. This increase already occurs early on after alloSCT, during immunosuppression, even though on a lower level compared to later time points, when immunosuppression is reduced. Donor-derived T cells recognize “foreign” cells regardless of whether they are healthy or leukemia cells. LSCs may be recognized and eradicated more efficiently when more antigen encounters occur. This indicates that an increase in the number of $GPR56^+$ T cells could help facilitate this process. Here it is shown that distinct thresholds are necessary for CMV-sero-positive and -negative patients, given that CMV-positive patients elicit an increase of GPR56 expression. This is in line with previous findings which report that CMV-specific T cells become and maintain GPR56 positivity[111].

Identification of the factors that promote GPR56 upregulation as well as those that create the inflammatory and immunosuppressive environment present in REL samples could aid in the development of small molecules that enhance GvL effects while minimizing the risk of relapse. Importantly, the classification of relapse versus remission poses a complex challenge due to the dynamic nature of the disease, where a patient's current relapse status may not persist indefinitely. This complexity can hinder the possibility of drawing confident conclusions regarding clinical outcomes.

This study implicates GPR56 in $CD8^+$ cytotoxicity, therapy outcome as well as the ability to recognize and eliminate foreign cells. Therefore, monitoring GPR56 expression on T cells may serve as a valuable tool to assess the level of recognition of non-self-antigens following alloSCT.

3.1.5 Balancing GvL Effect and GVHD after alloSCT

AlloSCT is based on the principle of using immune cells from a healthy donor to attack and eliminate AML cells in the patient. However, this approach can lead to a situation where transplanted immune cells not only attack cancer cells but also healthy tissues of the recipient, resulting in graft-versus-host disease (GVHD). The severity and frequency of GVHD are major limiting factors for alloSCT, and many efforts have been made to address this complication[56]. The paradox between the GvL effect and GVHD arises because the same immune cells that target AML cells can also cause tissue damage in the host. Thus, in order to achieve optimal clinical outcomes after alloSCT a balance between GvL and GVHD needs to be achieved. Several strategies, such as T-cell depletion, pharmacologic prophylaxis, and regulatory T cell therapy, have been developed to reduce GVHD without compromising the GvL effect[56].

Various populations of regulatory cells have been studied to gain a better understanding of the immunologic aspects of alloSCT and their translation to clinical practice[166]. Tregs and NK-T cells have been studied extensively in the context of GVHD modulation, while myeloid-derived suppressor cells, mesenchymal stem cells, and regulatory B cells are also believed to play a significant role in post-transplant immune regulation[167]. For example, in mice it was shown that $CD4^+$ $CD25^+$ $FoxP3^+$ Tregs suppress lethal GVHD while maintaining graft-versus-tumour activity[168].

Tregs are frequently implicated in immune modulation, since their main function is to suppress the activation and proliferation of other T cells[99]. In this thesis, it was shown that Tregs are enriched in patients that suffered from early relapse. Alongside with this enrichment, $CD8^+$ T cells were found to be characterized by an immunosuppressive phenotype, suggesting that the complex interplay between these immune components are instrumental in therapy outcome.

In this thesis, *TBX21* was identified as a driving TF of alloSCT T cells in CR, regulating the expression of cytotoxicity (*GZMB*, *NKG7*) as well as T cell activation (*CXC3R1*) molecules. Recent studies have established a link between TBX21 and graft outcome and organ rejection. After kidney transplant, patients that suffered from acute cellular rejection, were characterized by elevated expression of *TBX21* in alloreactive $CD8^+$ T cells[169]. This example illustrates that in order to improve the outcomes of

alloSCT, researchers should draw their attention to other transplantation models and organ rejection. By doing so, researchers may gain a better understanding into the mechanisms of immune rejection and to devise strategies to prevent or treat such complications.

Even though this study provides initial indications that elevated expression of GPR56 may be associated with therapy outcome, further investigation using larger patient cohorts with longitudinal monitoring is necessary to ascertain the significance of GPR56 dynamics in clinical diagnosis. It would be beneficial to understand the factors that directly contribute to GPR56 upregulation as well as to the inflammatory and immunosuppressive environment observed in relapse patients, which could ultimately aid in developing small molecules which promote GvL while preventing relapse.

3.2 Single cell profiling of xenograft mouse models unveil the bone marrow microenvironment remodeling in Acute Myeloid Leukemia

Acute myeloid leukemia (AML) resides in the bone marrow and promotes remodelling of the niche towards a leukemia permissive microenvironment. While several studies have attempted to study the impact of AML to the BM niche, such studies are often limited in one cell type, relying on *in vitro* models or AML mouse models. To date, there have been limited studies studying the potential influence of leukemic stem cells (LSCs), a fraction of which can escape therapy, on the bone marrow (BM) niche, highlighting a need for further research in this area.

Therefore, the second aim of this thesis was to profile changes in the BM stromal microenvironment induced by LSC burden during AML disease establishment and progression. To do so, I performed scRNA-seq on the BM niche of xenograft mouse models transplanted with human AML cells, characterized by either high or low LSC burden (LSC_{high} and LSC_{low} respectively). In order to account for differences associated with the presence of human cells in the mouse, mice transplanted with CB derived $CD34^+$ HSPCs were included in the analysis. CB derived $CD34^+$ HSPCs are characterized by increased proliferation capacity than those derived from BM or PBMCs, and therefore were advantageous for obtaining sufficient amounts of cells for downstream assays[170]. Lastly, the findings in the xenograft models were later validated using in vitro co-cultures of patient derived stromal cells.

3.2.1 Challenges of humanized mouse models

The immunocompromised NSGW41 mouse strain used in this study lacks functional B cells, T cells, and NK cells[62]. This strain has been widely used as a recipient for human cell engraftment, such as human HSPCs as well as AML cells[62], and, along with other NSG strains, are often utilized in preclinical research for hematopoetic malignancies. However, these models have several limitations, which may influence their ability to accurately recapitulate human disease.

For instance, the mouse microenvironment may not be capable of fully supporting the growth of human cells. In addition, the use of immunocompromized mice can limit the study of immunotherapeutic agents that depend on a functional immune system. Despite these limitations, xenograft mouse models remain a valuable tool in AML research, and ongoing efforts to refine these models will continue to enhance their utility in preclinical research.

In order to address some of these challenges, I compared the BM landscape of the NSGW41 strain to that derived from the commonly used C57BL/6J strain, the latter of

which has been previously profiled and is the basis of our current understanding of the BM microenvironment. For that, I first characterized the single cell landscape of the BM niche of NSGW41 mice and then performed horizontal integration with publicly available C57BL/6J datasets, which determined that there was striking overlap between the stromal cell populations of the different mouse strains. These results suggest that lack of immune cells in NSGW41 mice does not significantly affect the cellular states and identity of the BM stromal niche. Of note, to this date no published scRNA-seq dataset has outlined the non-hematopoietic landscape of the BM niche in NSGW41 mice.

3.2.2 Implication of ECs in the AML microenvironment

ECs are their main component of the BM vasculature and are subcategorized to arteriolar, which are tightly packed around the blood vessels and sinusoidal, which are less dense and are responsible for the trafficking of large molecules[175].

In this thesis I showed that transplantation of human CB-CD34$^+$ HSPCs had a minor impact on the EC compartment, which only showed a slight decrease of the EC arteriolar clusters. On the other hand, AML resulted in expansion of sinusoidal ECs as well as depletion of arteriolar subsets, while LSCs did not have significant impact. These results are consistent with previous studies, which have demonstrated that AML cells can infiltrate blood vessel walls and disrupt their structure, causing weakened or leaky blood vessels[171]. Moreover, AML cells can also secrete cytokines and other signaling molecules, which may contribute to the abnormal formation of blood vessels[171].

ECs have been implicated in disease progression and metastasis of several cancers other than AML[171]. In a healthy situation they are responsible for angiogenesis, immunity as well as trafficking of small molecules and cells, which can be exploited during tumorigenesis and metastatic conditions[171]. The link between tumour growth, metastasis and angiogenesis has been first proposed by Judah Folkman in 1971[172]. Importantly, ECs research has been exploding in the recent years due to the advances in isolating these cells from primary tissue and the advancement of single cell technologies.

Studies focusing on the ECs in several cancers have pinpointed several cancer associated signatures. In lung cancer, ECs have been found to downregulate genes involved in pro-inflammatory stimulation, chemotaxis like *CCL2* and *CCL18* as well as immune cell homing and recruitment of immune cells (*ICAM1*)[173]. *VEGF*, which regulates angiogenesis, has been proposed as a potential therapeutic target for several cancers[174]. In a clinical trial for metastatic colorectal cancer, the use of Bevacizumab which targets VEGF led to increased survival, although it was discontinued due to toxicity and hypertension, a well-established side effect of VEGF blockers[174].

In this thesis, several of the cancer associated EC signatures were identified to be associated with EC subsets enriched in AML, including *VEGFA* signalling[174]. This finding highlights the recurrent nature of the tumor microenvironment (TME) across different cancers. Signatures related to immune regulation from ECs, like recruitment of immune cells, were not detected in the present dataset, potentially due to one of the underlying challenges of the model which lacks a mature adaptive immune system.

3.2.3 Molecular signatures of cancer associated fibroblasts in AML

Fibroblasts are present across all tissues and one of their most common functions is the production of extracellular matrix (ECM), which facilitates cytokine and growth factor trafficking and is essential for the maintenance of tissue structure[175]. In addition, they are versatile and able to quickly respond to environmental cues, including the presence of cancer cells. In this case, the crosstalk between cancer cells and fibroblasts leads to the activation of the latter, which are often termed as cancer associated fibroblasts (CAFs)[175]. Even though the existence of CAFs in the tumour microenvironment is widely accepted, their exact role still remains debatable. They can secrete growth factors and cytokines which stimulate cancer cell proliferation, promote angiogenesis, lead to immune suppression and induce fibrosis, which may lead to scarring and functional loss of the affected area[175].

In this thesis, when investigating the impact of AML and LSCs on mesenchymal cells and fibroblasts, an AML-induced imbalance between the osteo- versus the adipo-lineage was detected. Adipo-lineage bias was determined based on the presence of adipo-CARs (adipocyte progenitors), since the high lipid content of adipocytes makes it challenging to isolate them. These results are complementary with previous studies, since adipocytes have been reported to support AML blast proliferation *in vivo* and *in vitro*[176].

Remarkably, the presented results illustrate the relationship between LSC burden and the remodelled stromal niche. In short, increased LSC burden was associated with further decrease of the osteo-lineage as well as the expansion of fibroblast clusters characterized by increased TGF-β signalling, ECM glycoproteins, suggesting an LSC induced ECM remodelling in the leukemic bone marrow (**Figure 51**).

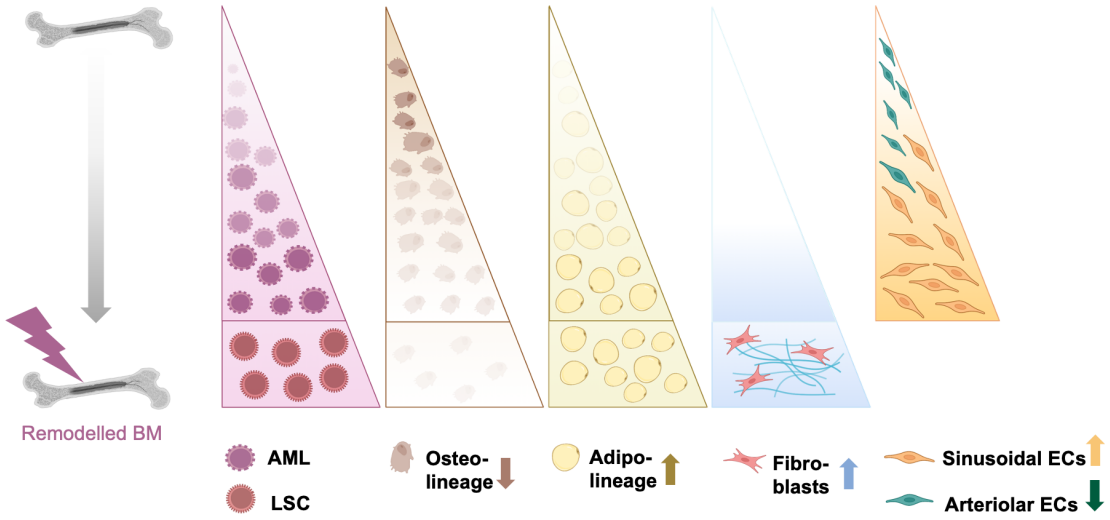

Figure 51 Schematic summarizing the proposed model of the altered BM stromal landscape in AML.
Acute myeloid leukemia (AML) remodels the bone marrow (BM) microenvironment. The presence of AML results in a notable increase in sinusoidal endothelial cells (ECs) and a reduction in arteriolar subsets as well as an imbalance between osteo- and adipo-lineage. While leukemic stem cells (LSCs) did not extensively impact the ECs, their presence resulted in further decline of the osteo-lineage and expansion of fibroblasts characterized by an altered fibrotic phenotype.

In order to define the driver of these changes, the cellular communication landscape was inferred, and detected *IL-1β* and *TGF-β* as top candidates to induce transcriptional changes to the fibroblasts in AML. In the bone marrow, fibroblast activation protein (FAP) expression is upregulated after stimulation by *IL-1β* and *TGF-β*[177]. Cancer associated fibroblasts (CAFs) upregulate FAP in several cancers, which is a protease located on the cell-surface, known to influence components of the ECM. It has often been utilized as a marker for stroma that supports tumour growth[178] and it has been shown in mice with lung carcinoma that depletion of FAP$^+$ CAFs can lead to tumour necrosis and tumour eradication by the immune system[179].

However, in the dataset presented here FAP expression was not detected, which limited the ability to draw conclusions regarding its presence in AML CAFs. One of the identified downstream targets of TGF-β here was transgelin (*Tagln*), which is known to regulate osteoblastic and adipogenic differentiation[180]. In addition, *Tagln* regulates ECM stiffness during ovarian cancer progression[181]. Thus, the link between TGF-β and *Tagln* may explain the observed phenotypes of BM resident fibroblasts in AML.

It is important to note that the interaction between AML cells and the BM stromal microenvironment is bidirectional. Here, the impact of AML cells to the stroma was primarily investigated since the ultimate goal was to understand how AML remodels the

microenvironment. In future studies it would be valuable to address whether and how the altered stroma may benefit AML, and examine whether the BM stroma could be targeted in order to improve therapy outcome.

To build on the findings presented in Section 2.2.7, further research could validate the altered population abundances outlined in this section using fluorescent microscopy. Several studies have successfully used imaging to spatially map the BM resident cell types in normal mice[26,122] as well as AML models[182,123].

3.2.4 Validating humanized mouse model findings *in vitro*

While studies in primary human samples are clinically more relevant, it is challenging to establish causal relationships due to the variability associated with low control over the experimental conditions. Thus, researchers in hematology mainly rely on alternative models, including humanized mouse models[62], *in vitro* co-cultures[66] and more recently 3D human BM organoids[67].

After delineating the altered stromal landscape of the non-hematopoietic BM niche using humanized mouse models, I sought to investigate the resemblance of these findings in the human setting. The motivation behind this approach was that mouse models do not fully replicate the complexity of human disease. Therefore, I performed *in vitro* co-cultures of patient derived MSCs with LSC_{high} or LSC_{low} AML. The results presented in Section 2.2.11, illustrate that MSCs exposed to AML cells with higher LSC burden decrease their potential towards osteo-lineage, thus validating the findings presented in Section 2.2.7. In addition, MSCs exposed to LSC_{high} AML overexpress matrix metalloproteinases *MMP1* and *MMP9*, known to degrade the ECM[149]. In addition, inhibition of *MMPs* including *MMP9*, reduces the expansion of AML and improves the response to chemotherapy[150]. Showing that this phenotype is caused by therapy resistant LSCs, warrants further investigation of AML-induced phenotypes while considering LSC burden.

A strategy to determine the underlying signalling causing these transcription alterations as well as the changes in population abundances is to quantify the secretion of cytokines upon coculture of AML cells with the BM stroma. Secreted cytokines, which include interferons, interleukins as well as chemokines modulate the behaviour of cell types, making them excellent candidates for the LSC induced BM remodelling. The landscape of the secreted cytokines, known as secretome, can be quantified by several methods, including colorimetric or fluorescent ELISA based methods (enzyme-linked immunosorbent assay), such as Luminex assays[183,184].

3.2.5 "To bone or to marrow, that is the question." HAMLet

As proposed in Section 2.2.8, AML cells express cytokines associated with increased LSC burden that may be secreted and thus impact the stroma (e.g., IL-1β, TGF-β1). In addition, depending on their localization in different bone regions they exhibit altered transcriptional profiles. These two points emphasize the significant relationship between AML cells and their stromal microenvironment.

Distinct stromal BM cell types may contribute to the immune escape of AML cells[185]. Since these cell types are more abundant in the bone lining, the phenotype of differentially localized AML cells was assessed. For example, bone lining residing AML was characterized by higher expression of *CD69*, known to regulate the self-renewal of AML cells[142] as well as *CXCL2* which together with *IL-1β* regulates cell proliferation[186].

As a follow up, associating the aforementioned signatures with the overall survival of AML patients may assist in estimating their clinical implications in prognosis. Such questions can be addressed by performing survival analysis using curated, publicly available bulk RNAseq datasets, from databases like TCGA Research Network: https://www.cancer.gov/tcga. In short, these analyses aim to examine the differences between transcriptional differences and time of death (overall survival, OS) or the time of AML relapse (relapse free survival, RFS)[187]. Kaplan-Meier plots can be used to visualize survival curves, while Cox proportional hazards regression analysis to identify the effect of certain variables on the estimated survival. An important aspect while selecting a dataset to perform this analysis is the number of individuals included in the study, in order to achieve sufficient statistical power, as well as how many individuals have information about the follow-up time and time of deaths[187]. In addition, the preparation of the sample and the feature tested warrants detailed examination, since different cell types are in different abundance in a tissue. For example, the BM consists primarily of hematopoietic cells. Thus, when assessing the prognostic power of a fibroblast derived feature, which is part of the 0.1% of non-hematopoietic cells in the BM[3], the non-specific expression into other cell types should be considered.

Despite these challenges, numerous studies have successfully used survival analysis to evaluate the broader implications of their findings. A recent study utilized single cell technologies as well as survival analysis of 2 bulk RNA-seq datasets of adult (Alliance) and pediatric (TARGET-AML) patients. This approach demonstrated a subset of inflammation genes, termed iScore, to be associated with AML survival risk[157]. In addition, single cell analysis of the T cell landscape of AML patients detected $GZMK^+$ $CD8^+$ T cells to be enriched in patients responding to PD-1 blockade therapy[92]. This

signature was associated with better outcomes in AML patients from the TCGA cohort, illustrating again the power of these analyses[92].

3.2.6 Leveraging cross-cancer insights to advance cancer treatment

Studying the different niches or microenvironments in which cancer cells reside, is crucial for understanding the complex nature of cancer. Cancer cells do not exist in isolation but rather interact with the surrounding cells, tissues, and structures. Each niche presents a unique set of environmental conditions that can influence tumour growth, survival, therapy response, even metastasis. Recently, Lomakin *et al.* 2022 marked a pivotal moment in understanding the impact of niches in cancer, by elucidating the spatial organisation of metastatic subclones in the TME, showing that genetically distinct clones locate in separate niches with distinct properties in the lymph node[188].

The underlying genetic as well as phenotypic heterogeneity of cancer cells is a well-established concept across several cancer types. Cancer cells exist in different states in the tumour. A recent study from the Yanai group defined recurrent cancer cell states among different cancers, subdivided into 16 transcriptional modules consisting of several genes[131]. These signatures include interferon production, hypoxia and epithelial to mesenchymal transition (EMT)[131]. Cancer states induce changes to their environment. For example, IFNγ production from cancer cells leads to interferon response and inflammation but also suppresses the expression of matrix metalloproteinases (MMPs), which degrade the ECM[149]. Moreover, it mitigates fibrosis through the inhibition of TGF-βR signalling[149]. However, chronic inflammation often leads to fibrosis, which occurs when the synthesis of new collagens is faster than its degradation[189]. Therefore, a link between inflammation and ECM remodelling has been hypothesized, though it needs to be further corroborated.

Another example of a cancer associated state is hypoxia. Cancer cells in the hypoxic region of a tumour are often more resistant to radiation and chemotherapy than cells in areas with sufficient oxygen supply[190]. Importantly, such phenotypes are shared between very distinct cancer types. In glioblastoma, which is the most aggressive brain tumour, cancer cells reside in peri-arteriolar niches in the brain, similarly with the HSCs in the BM[194]. Researchers have identified 17 biomarkers to be indicative of the hypoxic periarteriolar ecosystem, both in the HSCs niches in the BM as well as in the glioblastoma stem cell niches in the brain, providing an unexpected link between the two[191].

Single cell analysis is a powerful tool which allows researchers to better understand the heterogeneity within a given tissue. The difficulty of scaling up and analysing a

larger number of cells still remains a challenge. In contrast, imaging techniques like multiplexed immunofluorescence can provide an overview of a million cells or more in a single sample[192]. Such methods are also able to provide spatial information about the cells within the tissue, which is lost in scRNA-seq experiments. This spatial information can be critical in understanding the interactions between cells within a tissue and their roles in disease processes, like cancer. Several computational approaches allow the inference of cellular communications from scRNA-seq data[135], but such inferences would have to be further validated.

Importantly, solid tumors have well defined tumor boundaries, while circulating cancers like AML lack a specific location and are dispersed throughout the blood, making their analyses in the spatial context more challenging. Hence, the spatial landscape of the TME has been mainly characterized in solid tumors. For example, imaging mass cytometry on patient samples with glioblastoma revealed distinct cellular neighborhoods associated with survival, which lead to the identification of a specific subset of myeloperoxidase (MPO)-positive macrophages which related to long term survival[193].

While imaging-based techniques have significant advantages in the analysis of cellular profiling, they also face several challenges, including the number of features which can be simultaneously profiled[192]. Another challenge is the segmentation of individual cells, which can be particularly challenging for cells with complex morphology (i.e. neurons), or in tissue sections with high cell density (i.e. high tumour burden) or in sections with high levels of background immunofluorescence[192].

In conclusion, studying the TME is critical to gain a deeper understanding of malignancy and subsequently develop more effective therapies. Owing to the recent advancements in single-cell and multiplexed imaging technologies as well as computational analyses, we have the opportunity to gain novel insights into this complex ecosystem and unravel the interactions between tumor cells with the TME, which may lead to new targets for intervention, potentially improving outcomes for cancer patients.

CHAPTER 4: MATERIALS AND METHODS

The methods outlined in this Chapter were used in both Results' sections. Project specific procedures are specified as Section 2.1 or Section 2.2.

4.1 Freezing and thawing of cells

Cells were frozen in freezing medium (Appendix **Table 10**, 10% DMSO in fetal bovine serum, FBS; Sigma-Aldrich, #F5724) at a concentration of 10^6/ml and aliquoted into vials. The cells were placed in a freezing container at a cooling rate of 1°C/min in a -80°C freezer.

Samples were thawed in a 37°C water bath and cells were recovered in warm thawing medium containing Iscove's modified Dulbecco's medium (IMDM; Thermo Fisher Scientific, #21980065) supplemented with 20% fetal bovine serum (FBS; Sigma-Aldrich, #F5724) and 10µg/mL DNase I (Sigma-Aldrich, #DN25) in the case of AML samples (Section 2.2.11) as well as primary bone marrow (BM) samples (Section 2.1). In the case of mesenchymal stromal cells (MSC) samples, thawing medium contained Dulbecco's Modified Eagle Medium (DMEM, Thermo Fisher Scientific, #21885108) instead of IMDM. The protocol for MSC cell isolation is outlined in Section 4.7.

4.2 BM cell isolation for scRNA-seq (Section 2.1)

Frozen ficoll-processed primary BM samples of 6 AML patients (**Table 4**) were thawed and then stained using CD45-Pacific blue (Biolegend #304029), CD3-PerCP (Biolegend #344814), CD34-APC (BD biosciences #555824). Cells were sorted using BD FACSAria Fusion into $CD34^+$ and $CD3^+$ fractions. Prior to sorting, cells were stained with Caspase 3-FITC and collected in the FACS tubes coated with 10% FBS and collection buffer with 1x PBS and 0.04% RNAse-free Bovine Serum Albumin (BSA; Invitrogen #AM2616).

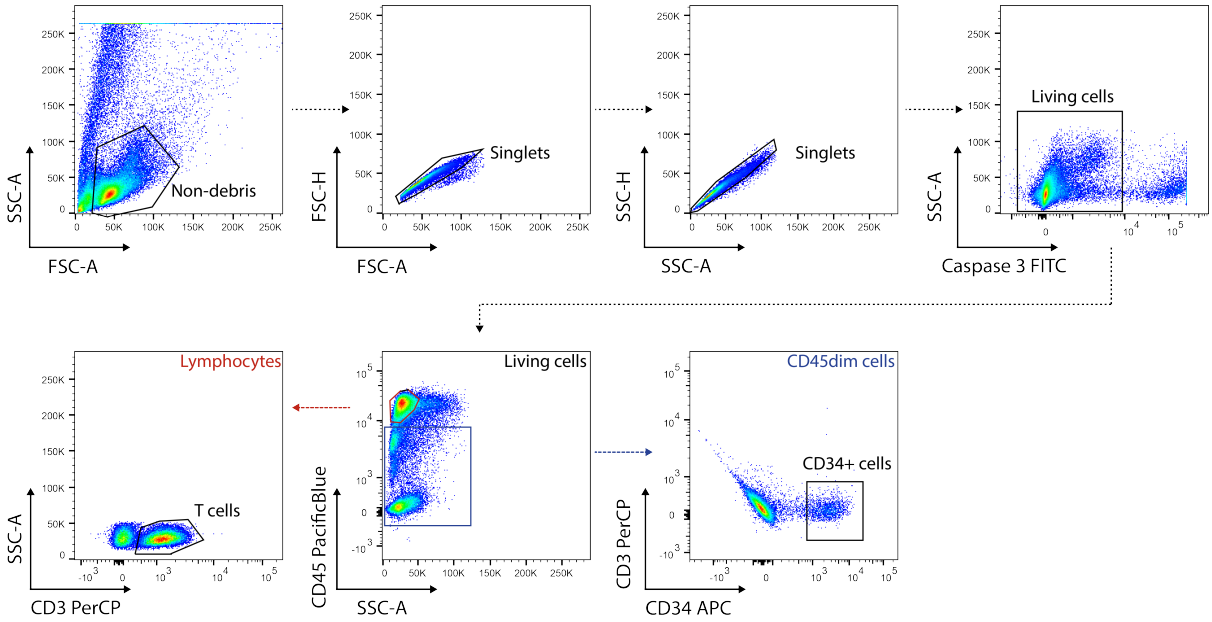

Figure 52 Representative gating scheme for FACS prior to the scRNA-seq analysis. Strategy used to isolate $CD3^+$ T cells and $CD34^+$ HSPCs from bone marrow aspirates of AML patients prior to scRNA-seq using the 10x Genomics platform.

Table 4: Patient characteristics of samples that underwent scRNA-seq. CR: Complete Remission; REL: Relapse.

Group	Sex	Donor sex	Days post alloSCT	BM blasts%	Chimerism (BM)	AML_type	Conditioning chemotherapy
REL	M	M	107	5%	94% donor	de novo	Thio/Flu/Treo/ ATG
REL	F	F	101	5%	96% donor	de novo	HAM/TBI 4Gy/Cy/Flu
REL	F	M	106	5%	90% donor	sAML	Treo/Flu/ATG
CR	M	M	95	<5%	100% donor	de novo	Thio/Flu/Treo/ ATG
CR	M	M	99	<5%	100% donor	AML MRC	Treo/Flu/ATG
CR	F	F	98	<5%	98% donor	de novo	Treo/Flu/ATG

4.3 Intracellular flow cytometry analysis (Section 2.1)

For the intracellular analysis of GZMB and PRF1 (Section 2.1.9), PBMCs of 10 AML patients post alloSCT were used. Unless stated otherwise, after each step samples were washed with FACS buffer (Appendix **Table 10**, 1x PBS supplemented with 2% FBS; 1200rpm/5min/RT). After thawing, cells were incubated for 15 minutes in room temperature with Zombie yellow (Zombie Yellow: Biolegend, #423103) according to manufacturer's protocol), followed by a 15 minute at room temperature staining with cell surface antibodies (GPR56-PE, Biolegend #358204; CD3-BUV750, CD8-BUV396).

Cells were then permeabilized using BD Permeabilizing solution 2 (BD Biosciences #340973) for 10 minutes at room temperature and then washed with FACS buffer (1000 g/5 minutes/room temperature). Finally, samples were stained for GZMB (GZMB-PE-Cy5, Biolegend, #372226) and PRF1 (PRF1-PacBlue, Biolegend, #353305) for 15 minutes at room temperature. Flow cytometry analysis was done at FACSymphony.

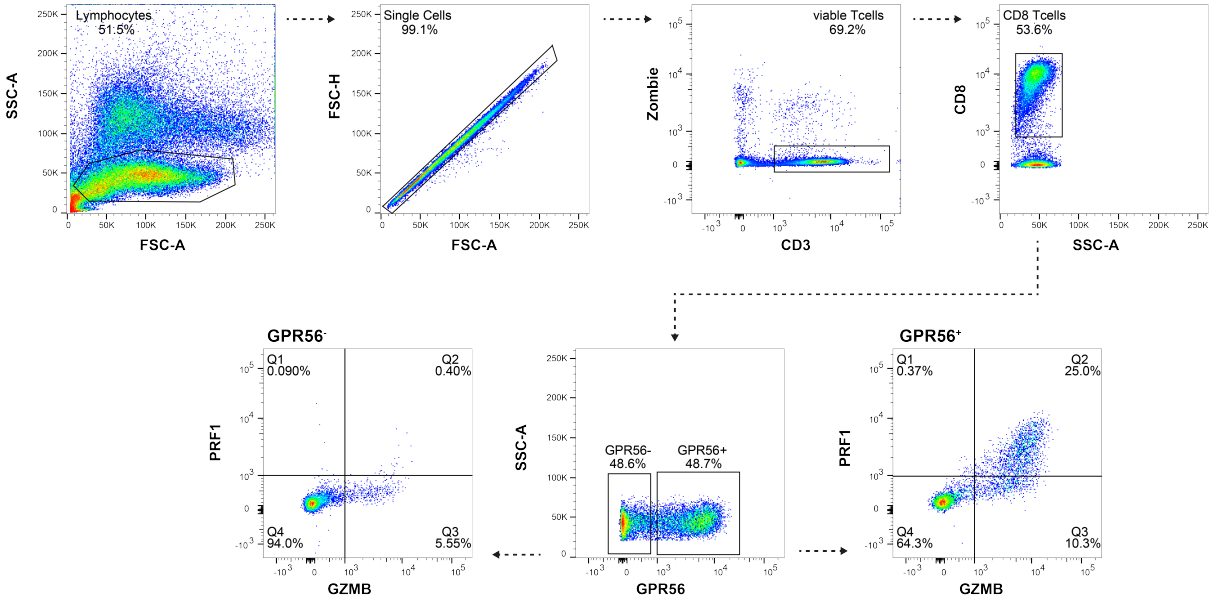

Figure 53 Representative gating scheme for flow cytometry analysis of intracellular cytotoxic markers on peripheral blood mononuclear cells (PBMCs) samples of 10 AML patients (Section 2.1.9).

4.4 Extracellular flow cytometry analysis (Section 2.1)

Staining for flow cytometry analysis was performed similar to Section 4.2. Data were acquired using BD LSRII, Canto and FACSymphony flow cytometers. Data were analysed using BD FACS Diva 8.0/9.0 and Flowjo X (Treestar Inc.) softwares.

For the analysis presented in Section 2.1.9: GPR56 dynamics after alloSCT, the antibody overview is in **Table 5**. The gating strategy for this analysis is in **Figure 54**. For the analysis presented in Section 2.1.9: GPR56 is co-expressed with cytotoxic molecules, the antibody overview is in **Table 6**. The gating strategy for this analysis is in **Figure 55**.

Table 5: List of antibodies used for flow cytometry analysis presented in Section 2.1.9.

Target	Fluorophore	Company	Cat. No.
CD3	FITC	BD biosciences	555916
CD8	APC	BD biosciences	555369
CD4	APCH7	BD biosciences	560158
GPR56	PE	Biolegend	358204

CD27	BV510	Biolegend	302835
CD45RA	PECy7	Biolegend	304126
CD56	PECy7	Biolegend	362509
CCR7	Pacific blue	Biolegend	353210

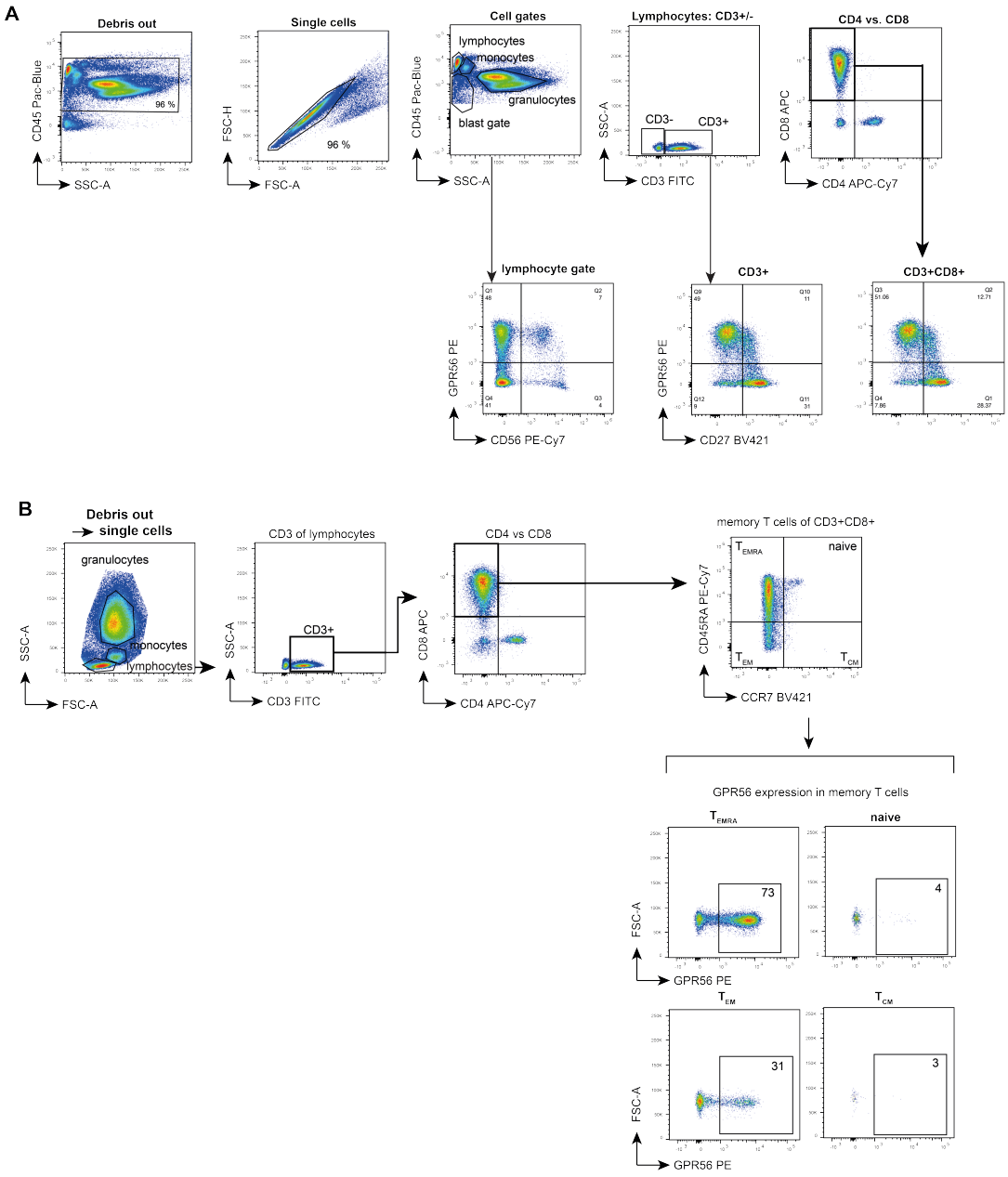

Figure 54 Flow cytometry analysis gating scheme for results presented in Section 2.1.9. Figure provided by Xizhe Wang. TCM: Central memory T cellls, TEM: CD45RA+ effector memory T cells, TEMRA: CD45RA+ effector memory T cells.

Table 6: List of antibodies used for flow cytometry analysis of extracellular activation markers presented in Section 2.1.9.

Target	Fluorophore	Company	Cat. No.
CD3	BV750	BD biosciences	747058
CD8	BUV395	BD Horizon	563795
CD4	BUV496	BD Horizon	612936
GPR56	PE	Biolegend	358204
CD27	BUV737	BD Horizon	612829
CD107a	PE-Cy7	Biolegend	328618
PD-1	APC	Biolegend	329908

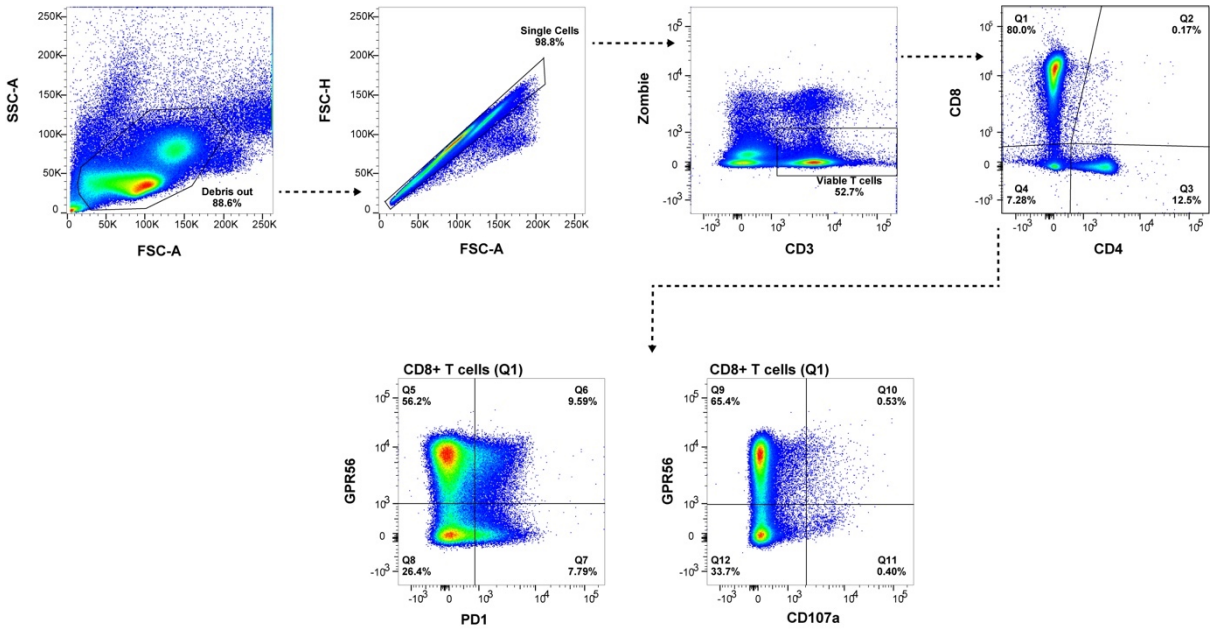

Figure 55 Representative gating scheme for flow cytometry analysis of activation markers on Peripheral blood mononuclear cells (PBMCs) samples of 10 AML patients (Section 2.1.9).

4.5 scRNA-seq sample and library preparation

Single-cell RNA sequencing (scRNA-seq) was performed using the 10x Genomics platform, which enables high-throughput profiling of gene expression in individual cells. 10x technology utilizes microfluidics and a gel bead-in-emulsion (GEM) approach to capture RNA from single cells. These beads are uniquely barcoded since they are coated with oligonucleotides that contain a cell barcode as well as unique molecular identifiers (UMI), which are used for distinguishing PCR duplicates. Single cells are encapsulated with the uniquely barcoded beads as well as lysis buffer, allowing the capture of the cell's mRNA. Once captured, the beads undergo reverse transcription (RT) to generate cDNA libraries that are further amplified and sequenced using Illumina sequencing technology.

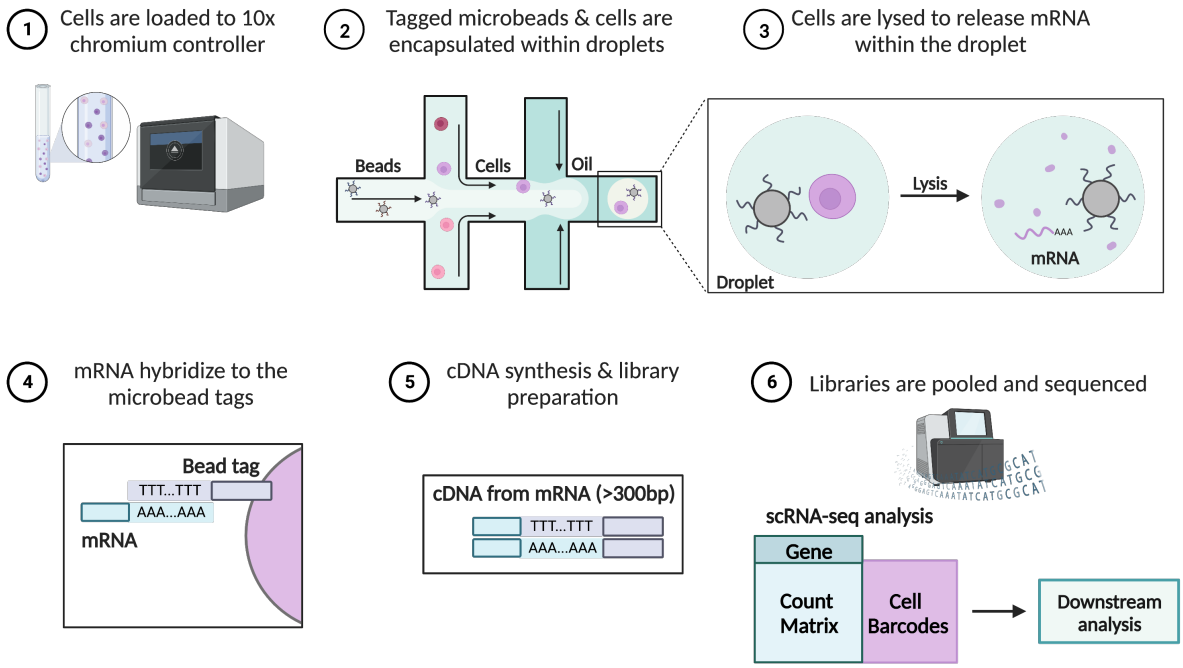

Figure 56 Overview of the 10X Genomics scRNA-seq protocol.
Single cell suspension is loaded onto the 10x Genomics Chromium microfluidics device, which partitions the cells into individual gel beads in emulsion (GEMs). Each GEM contains a single cell, a unique barcode, and a bead coated with oligonucleotides specific to that barcode. GEMs are processed to lyse the cells and capture the RNA transcripts. The captured RNA is reverse transcribed into cDNA. The resulting cDNA is then amplified using PCR. The amplified cDNA library is sequenced using Illumina sequencing technology. The resulting sequencing data are analysed to identify and quantify the expression levels of genes in each individual cell. Created using Biorender.com.

Single cells per sample were used as an input to 10X Genomics single-cell 3′ Gene Expression v3 assay. Libraries were prepared based on manufacturer's instructions. Sequencing was performed using Illumina NextSeq 500 (Section 2.1) and Illumina NextSeq 2000 (Section 2.2).

4.6 Isolation of cord blood HSPCs (Section 2.2)

Cord blood (CB) samples were incubated for 10 minutes at room temperature with 1:1000 DNase 1 (Sigma-Aldrich, #DN25) and were then resuspended with CB-HSPCs resuspension buffer (Appendix **Table 10, 11**; 1% BSA & 10mM EDTA in 1xPBS) by a factor 1:2. They were then subjected to mononuclear cell (MNC) isolation using 15ml Ficoll Hypaque density gradient (Thermo, #GE17-1440-02) per Sepmate isolation tube (STEMCELL, #85450). The diluted blood was added and centrifuged (800g/15min/RT). After centrifugation, the supernatant was further diluted with

resuspension buffer and centrifuged again (300g/5min/RT). The pellet, which contained the CB-MNCs, was resuspended CB-HSPCs resuspension buffer.

After MNC isolation, $CD34^+$ HSPCs were isolated by immunomagnetic separation using $CD34^+$ MicroBeads (Miltenyi Biotec, #130-046-502). The method is based on the positive selection of HSPCs using columns coated with CD34 antibody. Every 3 x 10^8 MNCs, 100ul of FcR blocking reagent was added along with 100ul CD34 magnetic microbeads. The suspension was incubated for 30min/4°C, diluted with CB-HSPCs resuspension buffer to 10ml and centrifuged (300g/10min/RT). Afterwards, the pellet was resuspended in 500ul CB-HSPCs buffer. The column previously got rinsed 2 times with 500ul CB-HSPCs buffer. The $CD34^-$ MNCs suspension was run through the column, while the $CD34^+$ remain bound. Elution of the $CD34^+$ cells happened in the absence of a magnetic field, using 1ml of CB-HSPCs buffer.

4.7 Bone Marrow MSC isolation & In vitro co-cultures (Section 2.2)

Bone Marrow aspirates were subjected to a ficoll density gradient (similarly to Section 4.6) in order to obtain mononuclear cells (MNC)[194]. Bone marrow MNCs were plated using MSC medium (Appendix **Table 10**) containing Dulbecco's Modified Eagle Medium (DMEM, Thermo Fisher Scientific, #21885108) supplemented with 10% human platelet lysate (hPL; PanBiotech, #P40-29050). Non-adherent cells were removed from the culture after 3 days and MSCs were allowed to expand till they reached 90% confluency (1-2 weeks in passage 1). MSCs were then trypsinized (Trypsin 10x, Sigma-Aldrich, #59427c) and passaged at a seeding density 20.000 cells/cm^2.

MSCs were seeded into 6-well tissue-culture plates at a seeding density of 40.000cells/cm^2. After reaching 80-90% confluency, AML cells were added on top at a seeding density of 20.000cells/cm^2. AML cells were cultured in IMDM (Thermo Fisher Scientific, #21980065) supplemented with 15% BIT (bovine serum albumin insulin transferrin, Stem Cell Technologies, #09500), SCF 100ng/mL (Stem cell factor; Shenandoah, #100-04), FLT-3 Ligand 50ng/mL (Fms-related tyrosine kinase 3 ligand; Shenandoah, #100-21), IL-3 20 ng/mL (Interleukin 3; Shenandoah, #100-80), G-CSF 10 ng/mL (Granulocyte-colony stimulating factor; Shenandoah, #100-72), 100 μM β-mercaproethanol (Gibco, #21985023), 50 μg/ml Gentamicin (Thermo Fisher Scientific #15750060) and 10 μg/mL Ciproflaxin (GenHunter #Q902-10ML). *In vitro* co-culture lasted for 48 hours. MSCs were then trypsinized and were subjected to scRNA-seq, similarly to Section 4.5.

4.8 Mouse Xenotransplantation (Section 2.2)

Mouse experiments were approved by Animal Care and Use Committees of the German Regierungspräsidium Karlsruhe für Tierschutz und Arzneimittelüberwachung. Mice were maintained in individually ventilated cages at the German Cancer Research Center (DKFZ), Heidelberg, Germany. NOD.Cg-*Kit*$^{W\text{-}41J}$*Prkdc*^{scid}*Il2rg*^{tm1Wjl}/WaskJ (NSGW41) mice were used for xenotransplantation. AML cells and CB-CD34^{+} HSPCs were intravenously injected in female NSGW41 mice, of age between 8-10 weeks old. AML cells were provided by Dr. Swati Garg and CRISPR/Cas9 knock out lines were generated as previously described in Garg et al., 2019[38].

4.9 Mouse bone preparation and cell isolation for scRNA-seq (Section 2.2)

Hips, femurs and tibiae were isolated and cleaned from surrounding tissue. Then, the bone marrow (BM) was flushed out using Roswell Park Memorial Institute (RPMI 1640, Sigma-Aldrich, #R8758) supplemented with 2% FBS and the bone linings were then digested.

To isolate the cells from the bone lining, bones were crushed in RPMI 1640 supplemented with 2% FBS and were then digested with 2 ml Digestion Buffer containing 2mg/ml collagenase IV (ThermoFisher Scientific, #17104-019) and 1 mg/ml dispase in Hank's Balanced Salt Solution (HBSS; ThermoFisher, #14175-053) for 10 minutes at 37 °C. The digestion was later blocked using FACS buffer (1X PBS with 2% FBS, Appendix **Table 10**). This step was repeated twice and the digested medium was then centrifuged. In order to discard red blood cells, pelleted cells were treated with ACK lysing buffer (Gibco, #A1049201) for 10 minutes at room temperature. After adding 2 ml of FACS buffer, lysis was stalled and hematopoietic cell depletion was performed using CD45 microbeads (Miltenyi Biotech, #130-052-301) for 15 minutes on ice and were then filtered through magnetic LS columns. Cells were then incubated with a staining mix containing specific antibodies described in **Table 7** before being FACS sorted into collection tubes containing 1x PBS and 0.4% BSA (Appendix **Table 11**).

In addition, flushed BM samples were used to isolate human HSPCs, human AML and mouse HSPCs. Similarly to the bone lining, cells were subjected to red blood cell lysis using the ACK lysing buffer (Gibco, #A1049201) and hematopoietic cell depletion using CD45 microbeads (Miltenyi Biotech, #130-052-301). The resulting cells were split and incubated with separate staining cocktails (**Table 7**), depending on the target population. FACS was performed with BD FACSAria Fusion Cell Sorter.

Table 7: List of antibodies used for the isolation of cells from mouse bones, prior to scRNA-seq.

Source	Sorted population	Target	Fluorophore	Company	Cat. No.
Bone Lining	**Mouse stroma**	mPDFGR	PE-Cy7	eBioscience	2071282
		mTer119	FITC	Invitrogen	1989148
		mCD41	FITC	BD Pharmingen	9281961
		mCD45	FITC	eBioscience	11045182
		mCD200	Alexa700	BD Horizon	565546
		mCD31	BV421	Invitrogen	562939
		mCD71	PE	eBioscience	12071183
		hCD45	BV711	Biolegend	563685
		hCD34	APC-Cy7	Invitrogen	47034942
Bone Marrow	**Mouse blood**	CD150	PE-Cy7	Biolegend	115914
		CD48	FITC	Biolegend	103404
		lineage: Ter119	Alexa700	Biolegend	475921
		lineage: CD11b	Alexa700	Invitrogen	2106095
		lineage: B220	Alexa700	eBioscience	2283174
		lineage: CD4	Alexa700	Invitrogen	4313129
		lineage: Gr1	Alexa700	Invitrogen	4313597
		lineage: CD8a	Alexa700	Invitrogen	4329739
		Sca1	APC-Cy7	BD Pharmingen	560654
		CD34	BV421	BD Horizon	624336
		cKit	BV711	Biolegend	105835
	Human HSPCs/AML	CD19	PE-Cy7	BD Biosciences	557835
		CD3	FITC	BD Biosciences	555332
		CD45	Alexa700	Biolegend	304023
		CD45RA	APC-Cy7	Biolegend	304128
		CD34	BV421	BD Biosciences	562577
		CD33	PE	BD Biosciences	555450
All	**All**	Zombie	BV570	Biolegend	423103

4.10 Hashing using TotalSeq-B Ab (Section 2.2)

TotalSeq-B antibodies with oligonucleotides conjugates contain a capture sequence ("Capture Sequence 1"). This sequence is compatible with the capture sequence of the Single Cell 3' v3 Gel Bead oligos. The barcode sequence conjugated to the antibody is then sequenced, allowing the extraction of sample origin information. The TotalSeqB antibodies and their sequence used in this project are listed in **Table 8**. Hashing antibodies were added to the FACS staining cocktail from Section 4.9, in order to achieve sample deconvolution.

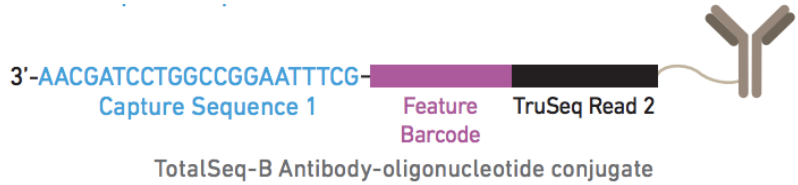

Figure 57 Cartoon illustrating TotalSeq-B antibody conjugated with an oligonucleotide. Provided by 10X Genomics.

Table 8: List of TotalSeqB hashing antibodies used in this study.

Species	Product	Company	Cat No.	Barcode Sequence
Mouse	TotalSeq™-B0301 anti-mouse Hashtag 1	Biolegend	394631	ACCCACCAGTAAGAC
Mouse	TotalSeq™-B 0304 anti-mouse Hashtag 4	Biolegend	155839	AAAGCATTCTTCACG
Mouse	TotalSeq™-B 0305 anti-mouse Hashtag 5	Biolegend	155839	CTTTGTCTTTGTGAG
Human	TotalSeq™-B0251 anti-human Hashtag 1	Biolegend	394631	GTCAACTCTTTAGCG
Human	TotalSeq™-B0252 anti-human Hashtag 2	Biolegend	394633	TGATGGCCTATTGGG
Human	TotalSeq™-B0253 anti-human Hashtag 3	Biolegend	394635	TTCCGCCTCTCTTTG
Human	TotalSeq™-B0254 anti-human Hashtag 4	Biolegend	394637	AGTAAGTTCAGCGTA
Human	TotalSeq™-B0255 anti-human Hashtag 5	Biolegend	394639	AAGTATCGTTTCGCA
Human	TotalSeq™-B0256 anti-human Hashtag 6	Biolegend	394641	GGTTGCCAGATGTCA

4.11 Immunofluorescence (Section 2.2)

MSCs cultured on coverslips were allowed to fix for 10min/RT using 4% PFA, then permeabilized using 0.1% TritonX (Sigma-Aldrich, #T8787-50ML) for 10min in RT. Cells were blocked using 2% BSA for 45min/RT and afterwards probed with primary antibody for 1h/RT. After washing with PBS, cells were probed with secondary antibody dye-conjugated antibodies for 45min/dark/RT and mounted using Prolong Gold anti-fade mounting medium. Slides were analysed using Olympus FV3000 confocal microscope and Fiji software was used for creating the projections of the z-stacks. The reagents used are listed in **Table 9**.

Table 9: Materials used for immunofluorescence experiments.

	Company	Cat. No.
Anti-CD90 / Thy1 antibody	Abcam	ab181469
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 594	Invitrogen	A-11032
Pierce ™ 16% formaldehyde (w / v), methanol-free	Thermo	28908
prolong diamond mounting medium with DAPI	Invitrogen	P36966
Thermo Scientific™ Frosted Microscope Slides, Cut	Thermo	AAAA000001##12E
Poly-L-Lysine Solution (0.01%)	EMD Millipore	A-005-C

4.12 Western Blot (Section 2.2)

Cells were thawed and centrifuged (1200rpm/5min/RT), then resuspended in 1xPBS and centrifuged again. Then, they were incubated in RIPA lysis buffer (ThermoFisher, #89900) containing protease inhibitors (Sigma-Aldrich, #11836170001) for 30 minutes on ice. Following centrifugation, the supernatant, which contained the cells' protein lysate was transferred into a new tube and resuspended in 4x NuPage LDS Sample Buffer (Invitrogen, #NP0007). Samples were then incubated at 95°C for 10 minutes before separation on a NuPage 4-12% Bis-Tris gel (Invitrogen, #NP0322BOX) using electrophoresis (10 min at 80 V followed by 1h at 120 V). The separated proteins were then transferred to a nitrocellulose blotting membrane using a blotting chamber (100V/1h/4°C) and stained with Ponceau. Membrane was washed with PBST and then incubated overnight with primary antibody diluted in 5% milk powder in PBST (GAPDH: GeneTex 101 #GTX627408; HLF: Abnova #H00003131-M04). The membrane was then washed three times with PBST and incubated with secondary antibody diluted in 5% milk powder in PBST for 1 hour at RT (HRP tagged anti-mouse Dianova, #115-036-062 and anti-rabbit Dianova, #111-165-144). Finally, the membrane was washed three more times with PBST prior the analysis of the resulting data. The exposed membranes were imaged using GE Healthcare Life Sciences, Amersham™ Imager 600. Detailed overview of all the buffers used can be found in Appendix **Table 10, 11**.

4.13 scRNA-seq data analysis

4.13.1 Preprocessing and quality control

Reads were aligned to the GRCh38 (v.2020-A), mm10 (v.2020-A-2.0.0) and barnyard (GRCh38 and mm10, v.2020-A) reference genome (depending on the dataset) and quantified using ***cellranger count*** (10x Genomics, v.3.0.1). For the downstream analysis I used Seurat v3[195]. Cells with less than 200 genes and more than 10-15% mitochondrial genes per cell were excluded from the downstream analysis. In the case of Section 2.1, the expression data across cells were corrected for ambient RNA using soupX[196].

4.13.2 Normalization, dimensionality reduction and clustering

Following quality control and prior to dimensionality reduction, raw counts were normalized in order to account for the sequencing depth per cell and per sample. For that I used ***SCTransform***, a method for normalisation and variance stabilisation for scRNA-seq data[195]. Briefly, UMI counts are modelled under a regularized negative binomial model in order to remove the variation due to sequencing depth. Variances are adjusted based on pooling information across genes with similar expression levels. The model outputs residuals which are the normalized expression.

On the SCT assay (*SCTransform* output), I performed principal component analysis (PCA) and then uniform manifold approximation and projection (UMAP) on the first 50 principal components. Ribosomal, mitochondrial, sex chromosome genes and transcripts were excluded from the variable features. Cells were then grouped into clusters using the Louvain algorithm, using FindClusters function. To define the clustering resolution of the aforementioned function Clustree was used (resolution = 1.5)[197].

Marker genes of the unsupervised clusters were identified using Seurat's ***FindMarkers*** function on the RNA assay. Genes considered were detected in at least 50% of cells per cluster (min.pct=0.5). Differentially expressed genes between clusters were identified using Wilcoxon Rank Sum test.

In Section 2.2, hashing antibody barcodes were used in order to multiplex multiple samples in one experiment. Seurat's ***HTOdemux*** function with default parameters was used to determine sample origin.

4.13.3 Comparison with publicly available data

In order to further validate that the unsupervised scRNA-seq clusters' annotation was accurate, I performed ***LabelTransfer***[103] analysis between the in house datasets produced with high quality publicly available ones. This approach involves executing ***SCTransform*** function for separately normalising each dataset, followed by the execution of ***PrepSCTIntegration*** which ensures all Pearson residuals are calculated. Lastly, data integration is performed by running with ***FindIntegrationAnchors*** and ***IntegrateData*** functions, with normalization.method = 'SCT'.

For Section 2.1 the publicly available dataset used was a CITE-seq dataset of peripheral blood mononuclear cells (PBMCs)[102]. For Section 2.2, the dataset discussed in Dolgalev and Tikhonova 2021[124] was used, which consists of three previously published C57BL/6J datasets[26,122,123].

4.13.4 Demultiplexing single cells based on genotypes

In order to cluster cells by genotype and distinguish individuals I used souporcell[81]. This method relies on variants identified in scRNA-seq or scATAC-seq data, without prior knowledge of the genotypes. These were then mapped to common variants based on common variants from the 1000 genomes project, filtered for variants of minor allele frequency (MAF) greater than $5x10^{-4}$ [84]. In order to determine matching individuals from different experiments, souporcell's shared_samples.py script was used.

In the scRNA-seq data, the sex of the individuals was defined qualitatively based on Xist (females) and RPS4Y1 (males) expression.

4.13.5 SCENIC TF activity analysis

The *pySCENIC* workflow[85] was run using an in-house constructed Snakemake pipeline. For gene regulatory network (GRN) inference, GRNBoost2 algorithm from the Arboreto package was used[198]. *SCENIC* analysis was performed on the raw scRNA-seq data. For predicting the transcription factor (TF) regulons, human v9 motif collection was used, hg38__refseq-r80__10kb_up_and_down_tss.mc9nr.feather and hg38__refseq-r80__500bp_up_and_100bp_down_tss.mc9nr.feather databases from cisTarget (https://resources.aertslab.org/cistarget/). The output AUC scores per cell and GRN were used for visualization and downstream analysis.

Assignment of target genes to known functions in **Figure 20** was performed using publicly available gene sets (IFN: HALLMARK_INTERFERON_GAMMA_RESPONSE, HALLMARK_INTERFERON_ALPHA_RESPONSE; Activation: Gene ontology, cell

activation involved in immune response and regulation of immune effector process, TNF: HALLMARK_TNFA_SIGNALING_VIA_NFKB).

4.13.6 Differential expression & Differential TF activity analysis from scRNA-seq data

For differential expression (DE) analysis between conditions, Seurat's FindMarkers function was used on the RNA assay, with the method MAST, an algorithm suitable for DE analysis from single cell data[104]. The analysis identified DE genes between conditions in all cell populations (p.adj < 0.05 & log2FC > 0.5, p-values were adjusted for multiple comparisons using the Bonferroni correction method).

To identify differentially active TFs, the output of the *SCENIC* algorithm was utilized (Section 4.13.5). *SCENIC* reconstructed a GRN and inferred TF activity at a single cell level. Since the inferred *SCENIC* TF activity scores consist of low values, following variable distributions per TF, it was challenging to conclude on a meaningful statistical test to detect differences in these scores. Thus, I followed an alternative approach outlined in **Figure 19 A**, where I used Fisher's exact test to first test for the enrichment of condition specific DE genes over all the TF target genes, extracted from the *SCENIC* GRN (p.adj < 0.05 after Bonferonni correction).

Functional analysis of DE genes was performed using ClusterProfiler[199]. The ClusterProfiler's functions applied (with default parameters) were ***enrichGO*** for Gene ontology enrichment analysis, ***compareCluster*** for KEGG pathway analysis, ***enricher*** function for Molecular Signatures Database MSigDB[200,201] (Hallmark collection) and ***enrichPathway*** for pathway annotation from ReactomePA[205]. The background gene-set was defined as all the genes expressed in the dataset. P values were adjusted using Benjamini-Hochberg and the cutoff was set to 0.05.

For endothelial cell specific functional enrichment analysis, gene set variation analysis (GSVA) [202] on 615 endothelial cells' related gene sets selected from the MSigDB database. This gene set was first presented in Kallucka et al. 2020 was used[129].

4.13.7 Pseudotime analysis (Section 2.1)

Pseudotime was calculated on the conventional CD8$^+$ T cells subsets (CD8$^+$ NV, CD8$^+$ eff. 1 and 2, CD8$^+$ mem. 1, 2 and 3) using *Monocle3*[100]. Pseudotime was calculated using the SCT assay, which contains the normalized and variance stabilized RNA count data[195]. Prior to the analysis, single cells across patients were aligned using align_cds function. Afterwards, the function learn_graph was executed with the use_partition argument set to True. CD8$^+$ NV cells were set as the starting point.

Diffusion maps[203], which is an alternative dimensionality reduction algorithm, were computed on the SCT assay using Partition-based graph abstraction (PAGA), a python based package for single cell analysis[204].

4.13.8 Cell to cell communication analysis (for Section 2.2)

NicheNet method was used in order to determine intercellular communication between cell clusters[135]. This method predicts associations between ligands (sender cells) and target genes (receiving cells) of interacting cells, by combining the expression data with prior knowledge regarding ligand-receptor interactions as well as downstream GRN.

To create a prior model of **ligand-target gene regulatory potential**, *NicheNet* integrates ligand-receptor and signaling data sources into a ligand-signaling network and it similarly generates a separate gene regulatory network. A weight of each data source is assigned automatically based on its contribution to the final model, using mlrMBO's model-based optimization[135]. The ligand-signaling network is used for calculating the importance for every target gene, Personalized PageRank[135]. Ligand-target genes regulatory potential scores are acquired by multiplying the ligand-regulator signaling scores with the adjacency matrix of the weighted gene regulatory network[135].

After the generation of the prior model, **ligand activity scores** are calculated as the Pearson correlation coefficient between the ligand-target regulatory potential and the target gene response. That way, *NicheNet* prioritizes ligand–target links in the query dataset (**Figure 58**)[135].

In Section 2.2.8, the analysis was performed using ***nichenet_seuratobj_aggregate()*** function, on the RNA assay, with the parameters expression_pct = 0.05 and lfc_cutoff = 0.2. Ligand-target, ligand-receptor and weighted networks matrices used were the default ones. In the case of inferred interactions between human and mouse cells, species gene conversion was performed using *Nichenet's* function ***convert_human_to_mouse_symbols***.

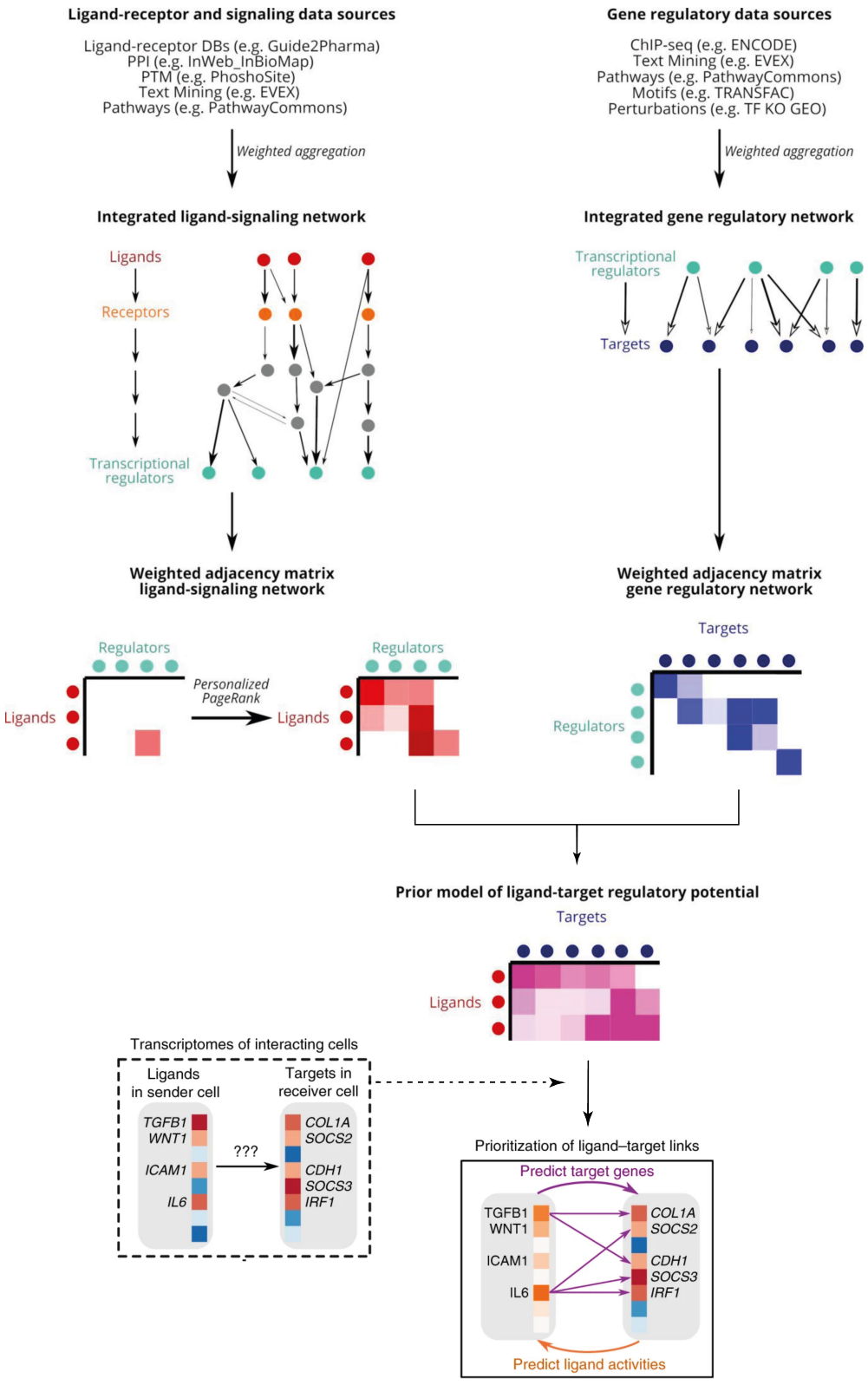

Figure 58 NicheNet workflow. Figure modified from Browaeys et al. 2020[135].

4.13.9 Cell type composition analysis

Apart from changes in gene expression and TF activity, analysis concerning the cell type compositional changes between conditions was performed. Per sorted population, or gate, the enrichment of every unsupervised cluster was estimated using Fisher's exact test. P-values were adjusted for multiple comparisons using the Bonferroni correction method. Per condition, a cluster was considered enriched when log2 odds ratio (log2OR) was greater than 0 and with a p.adj < 0.05.

BIBLIOGRAPHY

1. Weinberg, R.A. (2013). The biology of cancer (W.W. Norton & Company) 10.1201/9780429258794.

2. Schofield, R. (1978). The relationship between the spleen colony-forming cell and the haemopoietic stem cell. Blood Cells *4*, 7–25.

3. Nombela-Arrieta, C., and Manz, M.G. (2017). Quantification and three-dimensional microanatomical organization of the bone marrow. Blood Adv. *1*, 407–416. 10.1182/bloodadvances.2016003194.

4. Kaushansky, K. (2006). Lineage-specific hematopoietic growth factors. N. Engl. J. Med. *354*, 2034–2045. 10.1056/NEJMra052706.

5. Morrison, S.J., Wandycz, A.M., Hemmati, H.D., Wright, D.E., and Weissman, I.L. (1997). Identification of a lineage of multipotent hematopoietic progenitors. Development *124*, 1929–1939. 10.1242/dev.124.10.1929.

6. Cheng, H., Zheng, Z., and Cheng, T. (2020). New paradigms on hematopoietic stem cell differentiation. Protein Cell *11*, 34–44. 10.1007/s13238-019-0633-0.

7. Nombela-Arrieta, C., Pivarnik, G., Winkel, B., Canty, K.J., Harley, B., Mahoney, J.E., Park, S.-Y., Lu, J., Protopopov, A., and Silberstein, L.E. (2013). Quantitative imaging of haematopoietic stem and progenitor cell localization and hypoxic status in the bone marrow microenvironment. Nat. Cell Biol. *15*, 533–543. 10.1038/ncb2730.

8. Silberstein, L.E., and Lin, C.P. (2013). A new image of the hematopoietic stem cell vascular niche. Cell Stem Cell *13*, 514–516. 10.1016/j.stem.2013.10.012.

9. Kunisaki, Y., Bruns, I., Scheiermann, C., Ahmed, J., Pinho, S., Zhang, D., Mizoguchi, T., Wei, Q., Lucas, D., Ito, K., et al. (2013). Arteriolar niches maintain haematopoietic stem cell quiescence. Nature *502*, 637–643. 10.1038/nature12612.

10. Pinho, S., Marchand, T., Yang, E., Wei, Q., Nerlov, C., and Frenette, P.S. (2018). Lineage-Biased Hematopoietic Stem Cells Are Regulated by Distinct Niches. Dev. Cell *44*, 634-641.e4. 10.1016/j.devcel.2018.01.016.

11. Frenette, P.S., Pinho, S., Lucas, D., and Scheiermann, C. (2013). Mesenchymal stem cell: keystone of the hematopoietic stem cell niche and a stepping-stone for regenerative medicine. Annu. Rev. Immunol. *31*, 285–316. 10.1146/annurev-immunol-032712-095919.

12. Gottschling, S., Saffrich, R., Seckinger, A., Krause, U., Horsch, K., Miesala, K., and Ho, A.D. (2007). Human mesenchymal stromal cells regulate initial self-renewing divisions of hematopoietic progenitor cells by a beta1-integrin-dependent mechanism. Stem Cells *25*, 798–806. 10.1634/stemcells.2006-0513.

13. Jing, D., Fonseca, A.-V., Alakel, N., Fierro, F.A., Muller, K., Bornhauser, M., Ehninger, G., Corbeil, D., and Ordemann, R. (2010). Hematopoietic stem cells in co-culture with mesenchymal stromal cells--modeling the niche compartments in vitro. Haematologica *95*, 542–550. 10.3324/haematol.2009.010736.

14. Calvi, L.M., Adams, G.B., Weibrecht, K.W., Weber, J.M., Olson, D.P., Knight, M.C., Martin, R.P., Schipani, E., Divieti, P., Bringhurst, F.R., et al. (2003). Osteoblastic cells regulate the haematopoietic stem cell niche. Nature *425*, 841–846. 10.1038/nature02040.

15. Lo Celso, C., Fleming, H.E., Wu, J.W., Zhao, C.X., Miake-Lye, S., Fujisaki, J., Côté, D., Rowe, D.W., Lin, C.P., and Scadden, D.T. (2009). Live-animal tracking of individual haematopoietic stem/progenitor cells in their niche. Nature *457*, 92–96. 10.1038/nature07434.

16. Zhu, J., Garrett, R., Jung, Y., Zhang, Y., Kim, N., Wang, J., Joe, G.J., Hexner, E., Choi, Y., Taichman, R.S., et al. (2007). Osteoblasts support B-lymphocyte commitment and differentiation from hematopoietic stem cells. Blood *109*, 3706–3712. 10.1182/blood-2006-08-041384.

17. Naveiras, O., Nardi, V., Wenzel, P.L., Hauschka, P.V., Fahey, F., and Daley, G.Q. (2009). Bone-marrow adipocytes as negative regulators of the haematopoietic microenvironment. Nature *460*, 259–263. 10.1038/nature08099.

18. Bilwani, F.A., and Knight, K.L. (2012). Adipocyte-derived soluble factor(s) inhibits early stages of B lymphopoiesis. J. Immunol. *189*, 4379–4386. 10.4049/jimmunol.1201176.

19. Romero Marquez, S., Hettler, F., Hausinger, R., Schreck, C., Landspersky, T., Henkel, L., Angerpointner, C., Demir, I.E., Schiemann, M., Bassermann, F., et al. (2021). Secreted factors from mouse embryonic fibroblasts maintain repopulating function of single cultured hematopoietic stem cells. Haematologica *106*, 2633–2640. 10.3324/haematol.2020.249102.

20. Winkler, I.G., Barbier, V., Nowlan, B., Jacobsen, R.N., Forristal, C.E., Patton, J.T., Magnani, J.L., and Lévesque, J.-P. (2012). Vascular niche E-selectin regulates hematopoietic stem cell dormancy, self renewal and chemoresistance. Nat. Med. *18*, 1651–1657. 10.1038/nm.2969.

21. Hooper, A.T., Butler, J.M., Nolan, D.J., Kranz, A., Iida, K., Kobayashi, M., Kopp, H.-G., Shido, K., Petit, I., Yanger, K., et al. (2009). Engraftment and reconstitution of hematopoiesis is dependent on VEGFR2-mediated regeneration of sinusoidal endothelial cells. Cell Stem Cell *4*, 263–274. 10.1016/j.stem.2009.01.006.

22. Itkin, T., Gur-Cohen, S., Spencer, J.A., Schajnovitz, A., Ramasamy, S.K., Kusumbe, A.P., Ledergor, G., Jung, Y., Milo, I., Poulos, M.G., et al. (2016). Distinct bone marrow blood vessels differentially regulate haematopoiesis. Nature *532*, 323–328. 10.1038/nature17624.

23. Crisan, M., Yap, S., Casteilla, L., Chen, C.-W., Corselli, M., Park, T.S., Andriolo, G., Sun, B., Zheng, B., Zhang, L., et al. (2008). A perivascular origin for mesenchymal stem cells in multiple human organs. Cell Stem Cell *3*, 301–313. 10.1016/j.stem.2008.07.003.

24. Sugiyama, T., Kohara, H., Noda, M., and Nagasawa, T. (2006). Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. Immunity *25*, 977–988. 10.1016/j.immuni.2006.10.016.

25. Omatsu, Y., Sugiyama, T., Kohara, H., Kondoh, G., Fujii, N., Kohno, K., and Nagasawa, T. (2010). The essential functions of adipo-osteogenic progenitors as

the hematopoietic stem and progenitor cell niche. Immunity *33*, 387–399. 10.1016/j.immuni.2010.08.017.

26. Baccin, C., Al-Sabah, J., Velten, L., Helbling, P.M., Grünschläger, F., Hernández-Malmierca, P., Nombela-Arrieta, C., Steinmetz, L.M., Trumpp, A., and Haas, S. (2020). Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. Nat. Cell Biol. *22*, 38–48. 10.1038/s41556-019-0439-6.

27. Zhao, E., Xu, H., Wang, L., Kryczek, I., Wu, K., Hu, Y., Wang, G., and Zou, W. (2012). Bone marrow and the control of immunity. Cell. Mol. Immunol. *9*, 11–19. 10.1038/cmi.2011.47.

28. Luckheeram, R.V., Zhou, R., Verma, A.D., and Xia, B. (2012). $CD4^+$T cells: differentiation and functions. Clin. Dev. Immunol. *2012*, 925135. 10.1155/2012/925135.

29. Dong, C. (2009). Differentiation and function of pro-inflammatory Th17 cells. Microbes Infect. *11*, 584–588. 10.1016/j.micinf.2009.04.001.

30. Raskov, H., Orhan, A., Christensen, J.P., and Gögenur, I. (2021). Cytotoxic CD8+ T cells in cancer and cancer immunotherapy. Br. J. Cancer *124*, 359–367. 10.1038/s41416-020-01048-4.

31. Basu, R., Whitlock, B.M., Husson, J., Le Floc'h, A., Jin, W., Oyler-Yaniv, A., Dotiwala, F., Giannone, G., Hivroz, C., Biais, N., et al. (2016). Cytotoxic T cells use mechanical force to potentiate target cell killing. Cell *165*, 100–110. 10.1016/j.cell.2016.01.021.

32. Khwaja, A., Bjorkholm, M., Gale, R.E., Levine, R.L., Jordan, C.T., Ehninger, G., Bloomfield, C.D., Estey, E., Burnett, A., Cornelissen, J.J., et al. (2016). Acute myeloid leukaemia. Nat. Rev. Dis. Primers *2*, 16010. 10.1038/nrdp.2016.10.

33. Trumpp, A., and Haas, S. (2022). Cancer stem cells: The adventurous journey from hematopoietic to leukemic stem cells. Cell *185*, 1266–1270. 10.1016/j.cell.2022.03.025.

34. Menssen, A.J., and Walter, M.J. (2020). Genetics of progression from MDS to secondary leukemia. Blood *136*, 50–60. 10.1182/blood.2019000942.

35. Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M.A., and Dick, J.E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature *367*, 645–648. 10.1038/367645a0.

36. Pabst, C., Bergeron, A., Lavallée, V.-P., Yeh, J., Gendron, P., Norddahl, G.L., Krosl, J., Boivin, I., Deneault, E., Simard, J., et al. (2016). GPR56 identifies primary human acute myeloid leukemia cells with high repopulating potential in vivo. Blood *127*, 2018–2027. 10.1182/blood-2015-11-683649.

37. Ng, S.W.K., Mitchell, A., Kennedy, J.A., Chen, W.C., McLeod, J., Ibrahimova, N., Arruda, A., Popescu, A., Gupta, V., Schimmer, A.D., et al. (2016). A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature *540*, 433–437. 10.1038/nature20598.

38. Garg, S., Reyes-Palomares, A., He, L., Bergeron, A., Lavallée, V.-P., Lemieux, S., Gendron, P., Rohde, C., Xia, J., Jagdhane, P., et al. (2019). Hepatic leukemia factor is a novel leukemic stem cell regulator in DNMT3A, NPM1, and FLT3-ITD triple-mutated AML. Blood *134*, 263–276. 10.1182/blood.2018862383.

39. Juliusson, G., Antunovic, P., Derolf, A., Lehmann, S., Möllgård, L., Stockelberg, D., Tidefelt, U., Wahlin, A., and Höglund, M. (2009). Age and acute myeloid leukemia: real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry. Blood *113*, 4179–4187. 10.1182/blood-2008-07-172007.

40. Puumala, S.E., Ross, J.A., Aplenc, R., and Spector, L.G. (2013). Epidemiology of childhood acute myeloid leukemia. Pediatr. Blood Cancer *60*, 728–733. 10.1002/pbc.24464.

41. Kennedy, A.L., and Shimamura, A. (2019). Genetic predisposition to MDS: clinical features and clonal evolution. Blood *133*, 1071–1085. 10.1182/blood-2018-10-844662.

42. Rio-Machin, A., Vulliamy, T., Hug, N., Walne, A., Tawana, K., Cardoso, S., Ellison, A., Pontikos, N., Wang, J., Tummala, H., et al. (2020). The complex

genetic landscape of familial MDS and AML reveals pathogenic germline variants. Nat. Commun. *11*, 1044. 10.1038/s41467-020-14829-5.

43. Goldin, L.R., Kristinsson, S.Y., Liang, X.S., Derolf, A.R., Landgren, O., and Björkholm, M. (2012). Familial aggregation of acute myeloid leukemia and myelodysplastic syndromes. J. Clin. Oncol. *30*, 179–183. 10.1200/JCO.2011.37.1203.

44. Döhner, H., Wei, A.H., Appelbaum, F.R., Craddock, C., DiNardo, C.D., Dombret, H., Ebert, B.L., Fenaux, P., Godley, L.A., Hasserjian, R.P., et al. (2022). Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. Blood *140*, 1345–1377. 10.1182/blood.2022016867.

45. Duarte, D., Hawkins, E.D., and Lo Celso, C. (2018). The interplay of leukemia cells and the bone marrow microenvironment. Blood *131*, 1507–1511. 10.1182/blood-2017-12-784132.

46. Duarte, D., Hawkins, E.D., Akinduro, O., Ang, H., De Filippo, K., Kong, I.Y., Haltalli, M., Ruivo, N., Straszkowski, L., Vervoort, S.J., et al. (2018). Inhibition of endosteal vascular niche remodeling rescues hematopoietic stem cell loss in AML. Cell Stem Cell *22*, 64-77.e6. 10.1016/j.stem.2017.11.006.

47. Passaro, D., Di Tullio, A., Abarrategi, A., Rouault-Pierre, K., Foster, K., Ariza-McNaughton, L., Montaner, B., Chakravarty, P., Bhaw, L., Diana, G., et al. (2017). Increased vascular permeability in the bone marrow microenvironment contributes to disease progression and drug response in acute myeloid leukemia. Cancer Cell *32*, 324-341.e6. 10.1016/j.ccell.2017.08.001.

48. Boyd, A.L., Reid, J.C., Salci, K.R., Aslostovar, L., Benoit, Y.D., Shapovalova, Z., Nakanishi, M., Porras, D.P., Almakadi, M., Campbell, C.J.V., et al. (2017). Acute myeloid leukaemia disrupts endogenous myelo-erythropoiesis by compromising the adipocyte bone marrow niche. Nat. Cell Biol. *19*, 1336–1347. 10.1038/ncb3625.

49. Kumar, B., Garcia, M., Weng, L., Jung, X., Murakami, J.L., Hu, X., McDonald, T., Lin, A., Kumar, A.R., DiGiusto, D.L., et al. (2018). Acute myeloid leukemia

transforms the bone marrow niche into a leukemia-permissive microenvironment through exosome secretion. Leukemia *32*, 575–587. 10.1038/leu.2017.259.

50. Li, Z., Philip, M., and Ferrell, P.B. (2020). Alterations of T-cell-mediated immunity in acute myeloid leukemia. Oncogene *39*, 3611–3619. 10.1038/s41388-020-1239-y.

51. Ustun, C., Miller, J.S., Munn, D.H., Weisdorf, D.J., and Blazar, B.R. (2011). Regulatory T cells in acute myelogenous leukemia: is it time for immunomodulation? Blood *118*, 5084–5095. 10.1182/blood-2011-07-365817.

52. Noviello, M., Manfredi, F., Ruggiero, E., Perini, T., Oliveira, G., Cortesi, F., De Simone, P., Toffalori, C., Gambacorta, V., Greco, R., et al. (2019). Bone marrow central memory and memory stem T-cell exhaustion in AML patients relapsing after HSCT. Nat. Commun. *10*, 1065. 10.1038/s41467-019-08871-1.

53. Takami, A. (2018). Hematopoietic stem cell transplantation for acute myeloid leukemia. Int. J. Hematol. *107*, 513–518. 10.1007/s12185-018-2412-8.

54. Kantarjian, H., Kadia, T., DiNardo, C., Daver, N., Borthakur, G., Jabbour, E., Garcia-Manero, G., Konopleva, M., and Ravandi, F. (2021). Acute myeloid leukemia: current progress and future directions. Blood Cancer J. *11*, 41. 10.1038/s41408-021-00425-3.

55. Zilberberg, J., Feinman, R., and Korngold, R. (2015). Strategies for the identification of T cell-recognized tumor antigens in hematological malignancies for improved graft-versus-tumor responses after allogeneic blood and marrow transplantation. Biol. Blood Marrow Transplant. *21*, 1000–1007. 10.1016/j.bbmt.2014.11.001.

56. Negrin, R.S. (2015). Graft-versus-host disease versus graft-versus-leukemia. Hematology Am Soc Hematol Educ Program *2015*, 225–230. 10.1182/asheducation-2015.1.225.

57. Velardi, E., Tsai, J.J., and van den Brink, M.R.M. (2021). T cell regeneration after immunological injury. Nat. Rev. Immunol. *21*, 277–291. 10.1038/s41577-020-00457-z.

58. Li, Y., and Xu, L. (2015). Evaluation of TCR repertoire diversity in patients after hematopoietic stem cell transplantation. Stem Cell Investig. *2*, 17. 10.3978/j.issn.2306-9759.2015.09.01.

59. Kamel-Reid, S., and Dick, J.E. (1988). Engraftment of immune-deficient mice with human hematopoietic stem cells. Science *242*, 1706–1709. 10.1126/science.2904703.

60. Shultz, L.D., Ishikawa, F., and Greiner, D.L. (2007). Humanized mice in translational biomedical research. Nat. Rev. Immunol. *7*, 118–130. 10.1038/nri2017.

61. Shultz, L.D., Lyons, B.L., Burzenski, L.M., Gott, B., Chen, X., Chaleff, S., Kotb, M., Gillies, S.D., King, M., Mangada, J., et al. (2005). Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. J. Immunol. *174*, 6477–6489. 10.4049/jimmunol.174.10.6477.

62. Rahmig, S., Kronstein-Wiedemann, R., Fohgrub, J., Kronstein, N., Nevmerzhitskaya, A., Bornhäuser, M., Gassmann, M., Platz, A., Ordemann, R., Tonn, T., et al. (2016). Improved human erythropoiesis and platelet formation in humanized NSGW41 mice. Stem Cell Reports *7*, 591–601. 10.1016/j.stemcr.2016.08.005.

63. Yarden, Y., and Ullrich, A. (1988). Growth factor receptor tyrosine kinases. Annu. Rev. Biochem. *57*, 443–478. 10.1146/annurev.bi.57.070188.002303.

64. Hsu, Y.R., Wu, G.M., Mendiaz, E.A., Syed, R., Wypych, J., Toso, R., Mann, M.B., Boone, T.C., Narhi, L.O., Lu, H.S., et al. (1997). The majority of stem cell factor exists as monomer under physiological conditions. Implications for dimerization mediating biological activity. J. Biol. Chem. *272*, 6406–6415. 10.1074/jbc.272.10.6406.

65. Dexter, T.M., Allen, T.D., and Lajtha, L.G. (1977). Conditions controlling the proliferation of haemopoietic stem cells in vitro. J. Cell. Physiol. *91*, 335–344. 10.1002/jcp.1040910303.

66. Brenner, A.K., Nepstad, I., and Bruserud, Ø. (2017). Mesenchymal Stem Cells Support Survival and Proliferation of Primary Human Acute Myeloid Leukemia Cells through Heterogeneous Molecular Mechanisms. Front. Immunol. *8*, 106. 10.3389/fimmu.2017.00106.

67. Khan, A.O., Rodriguez-Romera, A., Reyat, J.S., Olijnik, A.-A., Colombo, M., Wang, G., Wen, W.X., Sousos, N., Murphy, L.C., Grygielska, B., et al. (2023). Human bone marrow organoids for disease modeling, discovery, and validation of therapeutic targets in hematologic malignancies. Cancer Discov. *13*, 364–385. 10.1158/2159-8290.CD-22-0199.

68. Cohen, Y.C., Zada, M., Wang, S.-Y., Bornstein, C., David, E., Moshe, A., Li, B., Shlomi-Loubaton, S., Gatt, M.E., Gur, C., et al. (2021). Identification of resistance pathways and therapeutic targets in relapsed multiple myeloma patients through single-cell sequencing. Nat. Med. *27*, 491–503. 10.1038/s41591-021-01232-w.

69. Fluorescence-Activated Cell Sorting (2011). In SpringerReference (Springer-Verlag). 10.1007/SpringerReference_32217.

70. Trapnell, C. (2015). Defining cell types and states with single-cell genomics. Genome Res. *25*, 1491–1498. 10.1101/gr.190595.115.

71. Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat. Methods *6*, 377–382. 10.1038/nmeth.1315.

72. Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. Nature *472*, 90–94. 10.1038/nature09807.

73. Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C., and Shendure, J. (2015). Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. Science *348*, 910–914. 10.1126/science.aab1601.

74. Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. Nature *523*, 486–490. 10.1038/nature14590.

75. Guo, H., Zhu, P., Wu, X., Li, X., Wen, L., and Tang, F. (2013). Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. Genome Res. *23*, 2126–2135. 10.1101/gr.161679.113.

76. Peterson, V.M., Zhang, K.X., Kumar, N., Wong, J., Li, L., Wilson, D.C., Moore, R., McClanahan, T.K., Sadekova, S., and Klappenbach, J.A. (2017). Multiplexed quantification of proteins and transcripts in single cells. Nat. Biotechnol. *35*, 936–939. 10.1038/nbt.3973.

77. Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat. Methods *14*, 865–868. 10.1038/nmeth.4380.

78. Nam, A.S., Kim, K.-T., Chaligne, R., Izzo, F., Ang, C., Taylor, J., Myers, R.M., Abu-Zeinah, G., Brand, R., Omans, N.D., et al. (2019). Somatic mutations and cell identity linked by Genotyping of Transcriptomes. Nature *571*, 355–360. 10.1038/s41586-019-1367-0.

79. Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B.Z., Mauck, W.M., Smibert, P., and Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biol. *19*, 224. 10.1186/s13059-018-1603-1.

80. Gehring, J., Hwee Park, J., Chen, S., Thomson, M., and Pachter, L. (2020). Highly multiplexed single-cell RNA-seq by DNA oligonucleotide tagging of cellular proteins. Nat. Biotechnol. *38*, 35–38. 10.1038/s41587-019-0372-z.

81. Heaton, H., Talman, A.M., Knights, A., Imaz, M., Gaffney, D.J., Durbin, R., Hemberg, M., and Lawniczak, M.K.N. (2020). Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. Nat. Methods *17*, 615–620. 10.1038/s41592-020-0820-1.

82. Huang, Y., McCarthy, D.J., and Stegle, O. (2019). Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. Genome Biol. *20*, 273. 10.1186/s13059-019-1865-2.

83. de Lima, M., Porter, D.L., Battiwalla, M., Bishop, M.R., Giralt, S.A., Hardy, N.M., Kröger, N., Wayne, A.S., and Schmid, C. (2014). Proceedings from the National Cancer Institute's Second International Workshop on the Biology, Prevention, and Treatment of Relapse After Hematopoietic Stem Cell Transplantation: part III. Prevention and treatment of relapse after allogeneic transplantation. Biol. Blood Marrow Transplant. *20*, 4–13. 10.1016/j.bbmt.2013.08.012.

84. 1000 Genomes Project Consortium, Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061–1073. 10.1038/nature09534.

85. Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat. Protoc. *15*, 2247–2276. 10.1038/s41596-020-0336-2.

86. Triana, S., Vonficht, D., Jopp-Saile, L., Raffel, S., Lutz, R., Leonce, D., Antes, M., Hernández-Malmierca, P., Ordoñez-Rueda, D., Ramasz, B., et al. (2021). Single-cell proteo-genomic reference maps of the hematopoietic system enable the purification and massive profiling of precisely defined cell states. Nat. Immunol. *22*, 1577–1589. 10.1038/s41590-021-01059-0.

87. Gorczyca, W., Sun, Z.-Y., Cronin, W., Li, X., Mau, S., and Tugulea, S. (2011). Immunophenotypic pattern of myeloid populations by flow cytometry analysis. In Recent Advances in Cytometry, Part B - Advances in Applications. Elsevier Methods in cell biology. (Elsevier), pp. 221–266. 10.1016/B978-0-12-385493-3.00010-3.

88. Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D., et al. (2017). Human haematopoietic stem

cell lineage commitment is a continuous process. Nat. Cell Biol. *19*, 271–281. 10.1038/ncb3493.

89. Braun, J., Frentsch, M., and Thiel, A. (2015). Hobit and human effector T-cell differentiation: The beginning of a long journey. Eur. J. Immunol. *45*, 2762–2765. 10.1002/eji.201545959.

90. Walunas, T.L., Lenschow, D.J., Bakker, C.Y., Linsley, P.S., Freeman, G.J., Green, J.M., Thompson, C.B., and Bluestone, J.A. (1994). CTLA-4 can function as a negative regulator of T cell activation. Immunity *1*, 405–413. 10.1016/1074-7613(94)90071-x.

91. Pizzolato, G., Kaminski, H., Tosolini, M., Franchini, D.-M., Pont, F., Martins, F., Valle, C., Labourdette, D., Cadot, S., Quillet-Mary, A., et al. (2019). Single-cell RNA sequencing unveils the shared and the distinct cytotoxic hallmarks of human TCRVδ1 and TCRVδ2 γδ T lymphocytes. Proc Natl Acad Sci USA *116*, 11906–11915. 10.1073/pnas.1818488116.

92. Abbas, H.A., Hao, D., Tomczak, K., Barrodia, P., Im, J.S., Reville, P.K., Alaniz, Z., Wang, W., Wang, R., Wang, F., et al. (2021). Single cell T cell landscape and T cell receptor repertoire profiling of AML in context of PD-1 blockade therapy. Nat. Commun. *12*, 6071. 10.1038/s41467-021-26282-z.

93. Ramesh, R., Kozhaya, L., McKevitt, K., Djuretic, I.M., Carlson, T.J., Quintero, M.A., McCauley, J.L., Abreu, M.T., Unutmaz, D., and Sundrud, M.S. (2014). Pro-inflammatory human Th17 cells selectively express P-glycoprotein and are refractory to glucocorticoids. J. Exp. Med. *211*, 89–104. 10.1084/jem.20130301.

94. Weston, B.R., Li, L., and Tyson, J.J. (2018). Mathematical Analysis of Cytokine-Induced Differentiation of Granulocyte-Monocyte Progenitor Cells. Front. Immunol. *9*, 2048. 10.3389/fimmu.2018.02048.

95. Willinger, T., Freeman, T., Herbert, M., Hasegawa, H., McMichael, A.J., and Callan, M.F.C. (2006). Human naive CD8 T cells down-regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in vivo. J. Immunol. *176*, 1439–1446. 10.4049/jimmunol.176.3.1439.

96. Hori, S., Nomura, T., and Sakaguchi, S. (2003). Control of regulatory T cell development by the transcription factor Foxp3. Science *299*, 1057–1061. 10.1126/science.1079490.

97. Fontenot, J.D., Gavin, M.A., and Rudensky, A.Y. (2003). Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. Nat. Immunol. *4*, 330–336. 10.1038/ni904.

98. Van Acker, H.H., Versteven, M., Lichtenegger, F.S., Roex, G., Campillo-Davo, D., Lion, E., Subklewe, M., Van Tendeloo, V.F., Berneman, Z.N., and Anguille, S. (2019). Dendritic Cell-Based Immunotherapy of Acute Myeloid Leukemia. J. Clin. Med. *8*. 10.3390/jcm8050579.

99. Vignali, D.A.A., Collison, L.W., and Workman, C.J. (2008). How regulatory T cells work. Nat. Rev. Immunol. *8*, 523–532. 10.1038/nri2343.

100. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386. 10.1038/nbt.2859.

101. Hamann, D., Baars, P.A., Rep, M.H., Hooibrink, B., Kerkhof-Garde, S.R., Klein, M.R., and van Lier, R.A. (1997). Phenotypic and functional separation of memory and effector human CD8+ T cells. J. Exp. Med. *186*, 1407–1418. 10.1084/jem.186.9.1407.

102. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zagar, M., et al. (2020). Integrated analysis of multimodal single-cell data. BioRxiv. 10.1101/2020.10.12.335331.

103. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21. 10.1016/j.cell.2019.05.031.

104. Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A.K., Slichter, C.K., Miller, H.W., McElrath, M.J., Prlic, M., et al. (2015). MAST: a flexible statistical framework for assessing transcriptional changes and characterizing

heterogeneity in single-cell RNA sequencing data. Genome Biol. *16*, 278. 10.1186/s13059-015-0844-5.

105. Yu, K., Kuang, L., Fu, T., Zhang, C., Zhou, Y., Zhu, C., Zhang, Q., Zhang, Z., and Le, A. (2021). CREM Is Correlated With Immune-Suppressive Microenvironment and Predicts Poor Prognosis in Gastric Adenocarcinoma. Front. Cell Dev. Biol. *9*, 697748. 10.3389/fcell.2021.697748.

106. Karl, F., Stoll, A., Böttcher-Loschinski, R., Böttcher, M., Baur, R., Jacobs, B., Völkl, S., Jitschin, R., Rösler, W., Mackensen, A., et al. (2021). Impact of Nrf2 expression in reconstituting T-cells of allogeneic hematopoietic stem cell transplanted patients. Leukemia *35*, 910–915. 10.1038/s41375-020-0956-0.

107. Kamal, A., Arnold, C., Claringbould, A., Moussa, R., Daga, N., Nogina, D., Kholmatov, M., Servaas, N., Mueller-Dott, S., Reyes-Palomaresv, A., et al. (2021). GRaNIE and GRaNPA: Inference and evaluation of enhancer-mediated gene regulatory networks applied to study macrophages. BioRxiv. 10.1101/2021.12.18.473290.

108. Chaix, J., Nish, S.A., Lin, W.-H.W., Rothman, N.J., Ding, L., Wherry, E.J., and Reiner, S.L. (2014). Cutting edge: CXCR4 is critical for CD8+ memory T cell homeostatic self-renewal but not rechallenge self-renewal. J. Immunol. *193*, 1013–1016. 10.4049/jimmunol.1400488.

109. Hintzen, R.Q., de Jong, R., Lens, S.M., and van Lier, R.A. (1994). CD27: marker and mediator of T-cell activation? Immunol. Today *15*, 307–311. 10.1016/0167-5699(94)90077-9.

110. Ma, P., Magut, M., Faller, D.V., and Chen, C.-Y. (2002). The role of Ras in T lymphocyte activation. Cell. Signal. *14*, 849–859. 10.1016/s0898-6568(02)00029-3.

111. Peng, Y.-M., van de Garde, M.D.B., Cheng, K.-F., Baars, P.A., Remmerswaal, E.B.M., van Lier, R.A.W., Mackay, C.R., Lin, H.-H., and Hamann, J. (2011). Specific expression of GPR56 by human cytotoxic lymphocytes. J. Leukoc. Biol. *90*, 735–740. 10.1189/jlb.0211092.

112. Chang, G.-W., Hsiao, C.-C., Peng, Y.-M., Vieira Braga, F.A., Kragten, N.A.M., Remmerswaal, E.B.M., van de Garde, M.D.B., Straussberg, R., König, G.M.,

Kostenis, E., et al. (2016). The Adhesion G Protein-Coupled Receptor GPR56/ADGRG1 Is an Inhibitory Receptor on Human NK Cells. Cell Rep. *15*, 1757–1770. 10.1016/j.celrep.2016.04.053.

113. Gruber, T., Kremenovic, M., Sadozai, H., Rombini, N., Baeriswyl, L., Maibach, F., Modlin, R.L., Gilliet, M., von Werdt, D., Hunger, R.E., et al. (2020). IL-32γ potentiates tumor immunity in melanoma. JCI Insight *5*. 10.1172/jci.insight.138772.

114. Liu, Y., Wang, S., Schubert, M.-L., Lauk, A., Yao, H., Blank, M.F., Cui, C., Janssen, M., Schmidt, C., Göllner, S., et al. (2022). CD33-directed immunotherapy with third-generation chimeric antigen receptor T cells and gemtuzumab ozogamicin in intact and CD33-edited acute myeloid leukemia and hematopoietic stem and progenitor cells. Int. J. Cancer *150*, 1141–1155. 10.1002/ijc.33865.

115. Hanamsagar, R., Reizis, T., Chamberlain, M., Marcus, R., Nestle, F.O., de Rinaldis, E., and Savova, V. (2020). An optimized workflow for single-cell transcriptomics and repertoire profiling of purified lymphocytes from clinical samples. Sci. Rep. *10*, 2219. 10.1038/s41598-020-58939-y.

116. Ding, L., Saunders, T.L., Enikolopov, G., and Morrison, S.J. (2012). Endothelial and perivascular cells maintain haematopoietic stem cells. Nature *481*, 457–462. 10.1038/nature10783.

117. Aiello, F.B., Keller, J.R., Klarmann, K.D., Dranoff, G., Mazzucchelli, R., and Durum, S.K. (2007). IL-7 induces myelopoiesis and erythropoiesis. J. Immunol. *178*, 1553–1563. 10.4049/jimmunol.178.3.1553.

118. Young, K., Eudy, E., Bell, R., Loberg, M.A., Stearns, T., Sharma, D., Velten, L., Haas, S., Filippi, M.-D., and Trowbridge, J.J. (2021). Decline in IGF1 in the bone marrow microenvironment initiates hematopoietic stem cell aging. Cell Stem Cell *28*, 1473-1482.e7. 10.1016/j.stem.2021.03.017.

119. Hattori, K., Heissig, B., Tashiro, K., Honjo, T., Tateno, M., Shieh, J.H., Hackett, N.R., Quitoriano, M.S., Crystal, R.G., Rafii, S., et al. (2001). Plasma elevation of stromal cell-derived factor-1 induces mobilization of mature and immature

hematopoietic progenitor and stem cells. Blood *97*, 3354–3360. 10.1182/blood.v97.11.3354.

120. Kaur, S., Sehgal, A., Wu, A.C., Millard, S.M., Batoon, L., Sandrock, C.J., Ferrari-Cestari, M., Levesque, J.-P., Hume, D.A., Raggatt, L.J., et al. (2021). Stable colony-stimulating factor 1 fusion protein treatment increases hematopoietic stem cell pool and enhances their mobilisation in mice. J. Hematol. Oncol. *14*, 3. 10.1186/s13045-020-00997-w.

121. Goldman, D.C., Bailey, A.S., Pfaffle, D.L., Al Masri, A., Christian, J.L., and Fleming, W.H. (2009). BMP4 regulates the hematopoietic stem cell niche. Blood *114*, 4393–4401. 10.1182/blood-2009-02-206433.

122. Tikhonova, A.N., Dolgalev, I., Hu, H., Sivaraj, K.K., Hoxha, E., Cuesta-Domínguez, Á., Pinho, S., Akhmetzyanova, I., Gao, J., Witkowski, M., et al. (2019). The bone marrow microenvironment at single-cell resolution. Nature *569*, 222–228. 10.1038/s41586-019-1104-8.

123. Baryawno, N., Przybylski, D., Kowalczyk, M.S., Kfoury, Y., Severe, N., Gustafsson, K., Kokkaliaris, K.D., Mercier, F., Tabaka, M., Hofree, M., et al. (2019). A cellular taxonomy of the bone marrow stroma in homeostasis and leukemia. Cell *177*, 1915-1932.e16. 10.1016/j.cell.2019.04.040.

124. Dolgalev, I., and Tikhonova, A.N. (2021). Connecting the dots: resolving the bone marrow niche heterogeneity. Front. Cell Dev. Biol. *9*, 622519. 10.3389/fcell.2021.622519.

125. Bapat, A., Keita, N., and Sharma, S. (2019). Pan-myeloid Differentiation of Human Cord Blood Derived CD34+ Hematopoietic Stem and Progenitor Cells. J. Vis. Exp. 10.3791/59836.

126. Matsuoka, Y., Sumide, K., Kawamura, H., Nakatsuka, R., Fujioka, T., Sasaki, Y., and Sonoda, Y. (2015). Human cord blood-derived primitive CD34-negative hematopoietic stem cells (HSCs) are myeloid-biased long-term repopulating HSCs. Blood Cancer J. *5*, e290. 10.1038/bcj.2015.22.

127. Drolle, H., Wagner, M., Vasold, J., Kütt, A., Deniffel, C., Sotlar, K., Sironi, S., Herold, T., Rieger, C., and Fiegl, M. (2015). Hypoxia regulates proliferation of

acute myeloid leukemia and sensitivity against chemotherapy. Leuk. Res. *39*, 779–785. 10.1016/j.leukres.2015.04.019.

128. Li, H.Y., Appelbaum, F.R., Willman, C.L., Zager, R.A., and Banker, D.E. (2003). Cholesterol-modulating agents kill acute myeloid leukemia cells and sensitize them to therapeutics by blocking adaptive cholesterol responses. Blood *101*, 3628–3634. 10.1182/blood-2002-07-2283.

129. Kalucka, J., de Rooij, L.P.M.H., Goveia, J., Rohlenova, K., Dumas, S.J., Meta, E., Conchinha, N.V., Taverna, F., Teuwen, L.-A., Veys, K., et al. (2020). Single-Cell Transcriptome Atlas of Murine Endothelial Cells. Cell *180*, 764-779.e20. 10.1016/j.cell.2020.01.015.

130. Zhang, H., Liesveld, J.L., Calvi, L.M., Lipe, B.C., Xing, L., Becker, M.W., Schwarz, E.M., and Yeh, S.-C.A. (2023). The roles of bone remodeling in normal hematopoiesis and age-related hematological malignancies. Bone Res. *11*, 15. 10.1038/s41413-023-00249-w.

131. Barkley, D., Moncada, R., Pour, M., Liberman, D.A., Dryg, I., Werba, G., Wang, W., Baron, M., Rao, A., Xia, B., et al. (2022). Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. Nat. Genet. *54*, 1192–1201. 10.1038/s41588-022-01141-9.

132. Tabe, Y., Shi, Y.X., Zeng, Z., Jin, L., Shikami, M., Hatanaka, Y., Miida, T., Hsu, F.J., Andreeff, M., and Konopleva, M. (2013). TGF-β-Neutralizing Antibody 1D11 Enhances Cytarabine-Induced Apoptosis in AML Cells in the Bone Marrow Microenvironment. PLoS ONE *8*, e62785. 10.1371/journal.pone.0062785.

133. Naba, A., Clauser, K.R., Hoersch, S., Liu, H., Carr, S.A., and Hynes, R.O. (2012). The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. Mol. Cell. Proteomics *11*, M111.014647. 10.1074/mcp.M111.014647.

134. Trino, S., Laurenzana, I., Lamorte, D., Calice, G., De Stradis, A., Santodirocco, M., Sgambato, A., Caivano, A., and De Luca, L. (2022). Acute myeloid leukemia cells functionally compromise hematopoietic stem/progenitor cells inhibiting normal hematopoiesis through the release of extracellular vesicles. Front. Oncol. *12*, 824562. 10.3389/fonc.2022.824562.

135. Browaeys, R., Saelens, W., and Saeys, Y. (2020). NicheNet: modeling intercellular communication by linking ligands to target genes. Nat. Methods *17*, 159–162. 10.1038/s41592-019-0667-5.

136. Waldeck, S., Rassner, M., Keye, P., Follo, M., Herchenbach, D., Endres, C., Charlet, A., Andrieux, G., Salzer, U., Boerries, M., et al. (2020). CCL5 mediates target-kinase independent resistance to FLT3 inhibitors in FLT3-ITD-positive AML. Mol. Oncol. *14*, 779–794. 10.1002/1878-0261.12640.

137. Arranz, L., Arriero, M.D.M., and Villatoro, A. (2017). Interleukin-1β as emerging therapeutic target in hematological malignancies and potentially in their complications. Blood Rev. *31*, 306–317. 10.1016/j.blre.2017.05.001.

138. Carey, A., Edwards, D.K., Eide, C.A., Newell, L., Traer, E., Medeiros, B.C., Pollyea, D.A., Deininger, M.W., Collins, R.H., Tyner, J.W., et al. (2017). Identification of Interleukin-1 by Functional Screening as a Key Mediator of Cellular Expansion and Disease Progression in Acute Myeloid Leukemia. Cell Rep. *18*, 3204–3218. 10.1016/j.celrep.2017.03.018.

139. Dong, M., and Blobe, G.C. (2006). Role of transforming growth factor-beta in hematologic malignancies. Blood *107*, 4589–4596. 10.1182/blood-2005-10-4169.

140. Wu, Y., Su, M., Zhang, S., Cheng, Y., Liao, X.Y., Lin, B.Y., and Chen, Y.Z. (2017). Abnormal expression of TGF-beta type II receptor isoforms contributes to acute myeloid leukemia. Oncotarget *8*, 10037–10049. 10.18632/oncotarget.14325.

141. Gong, Y., Zhao, M., Yang, W., Gao, A., Yin, X., Hu, L., Wang, X., Xu, J., Hao, S., Cheng, T., et al. (2018). Megakaryocyte-derived excessive transforming growth factor β1 inhibits proliferation of normal hematopoietic stem cells in acute myeloid leukemia. Exp. Hematol. *60*, 40-46.e2. 10.1016/j.exphem.2017.12.010.

142. Sachs, K., Sarver, A.L., Noble-Orcutt, K.E., LaRue, R.S., Antony, M.L., Chang, D., Lee, Y., Navis, C.M., Hillesheim, A.L., Nykaza, I.R., et al. (2020). Single-Cell Gene Expression Analyses Reveal Distinct Self-Renewing and Proliferating Subsets in the Leukemia Stem Cell Compartment in Acute Myeloid Leukemia. Cancer Res. *80*, 458–470. 10.1158/0008-5472.CAN-18-2932.

143. Sykes, S.M., Di Marcantonio, D., Martinez, E., Huhn, J., Gupta, A., and Mistry, R. (2018). JUN and ATF3 regulate the transcriptional output of the unfolded protein response to support acute myeloid leukemia. Blood *132*, 1327–1327. 10.1182/blood-2018-99-119742.

144. Li, L., Zhao, L., Man, J., and Liu, B. (2021). CXCL2 benefits acute myeloid leukemia cells in hypoxia. Int. J. Lab. Hematol. *43*, 1085–1092. 10.1111/ijlh.13512.

145. Russo, R.C., Garcia, C.C., Teixeira, M.M., and Amaral, F.A. (2014). The CXCL8/IL-8 chemokine family and its receptors in inflammatory diseases. Expert Rev. Clin. Immunol. *10*, 593–619. 10.1586/1744666X.2014.894886.

146. Li, Y., Cheng, J., Li, Y., Jiang, Y., Ma, J., Li, Q., and Pang, T. (2018). CXCL8 is associated with the recurrence of patients with acute myeloid leukemia and cell proliferation in leukemia cell lines. Biochem. Biophys. Res. Commun. *499*, 524–530. 10.1016/j.bbrc.2018.03.181.

147. De Filippo, K., and Rankin, S.M. (2020). The secretive life of neutrophils revealed by intravital microscopy. Front. Cell Dev. Biol. *8*, 603230. 10.3389/fcell.2020.603230.

148. Pillai, S., and Cariappa, A. (2009). The bone marrow perisinusoidal niche for recirculating B cells and the positive selection of bone marrow-derived B lymphocytes. Immunol. Cell Biol. *87*, 16–19. 10.1038/icb.2008.89.

149. Hu, X., and Ivashkiv, L.B. (2009). Cross-regulation of signaling pathways by interferon-gamma: implications for immune responses and autoimmune diseases. Immunity *31*, 539–550. 10.1016/j.immuni.2009.09.002.

150. Pirillo, C., Birch, F., Tissot, F.S., Anton, S.G., Haltalli, M., Tini, V., Kong, I., Piot, C., Partridge, B., Pospori, C., et al. (2022). Metalloproteinase inhibition reduces AML growth, prevents stem cell loss, and improves chemotherapy effectiveness. Blood Adv. *6*, 3126–3141. 10.1182/bloodadvances.2021004321.

151. Huo, Y., Wu, L., Pang, A., Li, Q., Hong, F., Zhu, C., Yang, Z., Dai, W., Zheng, Y., Meng, Q., et al. (2023). Single-cell dissection of human hematopoietic reconstitution after allogeneic hematopoietic stem cell transplantation. Sci. Immunol. *8*, eabn6429. 10.1126/sciimmunol.abn6429.

152. Zhang, T.Y., Dutta, R., Zhao, F., and Majeti, R. (2018). Human Acute Myeloid Leukemia Inhibits Normal Erythroid Differentiation through the Paracrine Effects of IL-6. Blood *132*, 911–911. 10.1182/blood-2018-99-119297.

153. Viale, M., Ferrini, S., and Bacigalupo, A. (1992). TCR gamma/delta positive lymphocytes after allogeneic bone marrow transplantation. Bone Marrow Transplant. *10*, 249–253.

154. Zenz, R., Eferl, R., Scheinecker, C., Redlich, K., Smolen, J., Schonthaler, H.B., Kenner, L., Tschachler, E., and Wagner, E.F. (2008). Activator protein 1 (Fos/Jun) functions in inflammatory bone and skin disease. Arthritis Res. Ther. *10*, 201. 10.1186/ar2338.

155. Xiao, G., Deng, A., Liu, H., Ge, G., and Liu, X. (2012). Activator protein 1 suppresses antitumor T-cell function via the induction of programmed death 1. Proc Natl Acad Sci USA *109*, 15419–15424. 10.1073/pnas.1206370109.

156. Kourtis, N., Wang, Q., Wang, B., Oswald, E., Adler, C., Cherravuru, S., Malahias, E., Zhang, L., Golubov, J., Wei, Q., et al. (2022). A single-cell map of dynamic chromatin landscapes of immune cells in renal cell carcinoma. Nat. Cancer *3*, 885–898. 10.1038/s43018-022-00391-0.

157. Lasry, A., Nadorp, B., Fornerod, M., Nicolet, D., Wu, H., Walker, C.J., Sun, Z., Witkowski, M.T., Tikhonova, A.N., Guillamot-Ruano, M., et al. (2023). An inflammatory state remodels the immune microenvironment and improves risk stratification in acute myeloid leukemia. Nat. Cancer *4*, 27–42. 10.1038/s43018-022-00480-0.

158. Récher, C. (2021). Clinical implications of inflammation in acute myeloid leukemia. Front. Oncol. *11*, 623952. 10.3389/fonc.2021.623952.

159. Hsiao, C.-C., Vos, E., van Gisbergen, K.P.J.M., and Hamann, J. (2023). The adhesion G protein-coupled receptor GPR56/ADGRG1 in cytotoxic lymphocytes. Basic Clin. Pharmacol. Toxicol. 10.1111/bcpt.13841.

160. Edfors, F., Danielsson, F., Hallström, B.M., Käll, L., Lundberg, E., Pontén, F., Forsström, B., and Uhlén, M. (2016). Gene-specific correlation of RNA and

protein levels in human cells and tissues. Mol. Syst. Biol. *12*, 883. 10.15252/msb.20167144.

161. Bilemjian, V., Vlaming, M.R., Álvarez Freile, J., Huls, G., De Bruyn, M., and Bremer, E. (2022). The Novel Immune Checkpoint GPR56 Is Expressed on Tumor-Infiltrating Lymphocytes and Selectively Upregulated upon TCR Signaling. Cancers (Basel) *14*. 10.3390/cancers14133164.

162. Argyriou, A., Wadsworth, M.H., Lendvai, A., Christensen, S.M., Hensvold, A.H., Gerstner, C., van Vollenhoven, A., Kravarik, K., Winkler, A., Malmström, V., et al. (2022). Single cell sequencing identifies clonally expanded synovial CD4+ TPH cells expressing GPR56 in rheumatoid arthritis. Nat. Commun. *13*, 4046. 10.1038/s41467-022-31519-6.

163. Chiang, N.-Y., Peng, Y.-M., Juang, H.-H., Chen, T.-C., Pan, H.-L., Chang, G.-W., and Lin, H.-H. (2017). GPR56/ADGRG1 Activation Promotes Melanoma Cell Migration via NTF Dissociation and CTF-Mediated Gα12/13/RhoA Signaling. J. Invest. Dermatol. *137*, 727–736. 10.1016/j.jid.2016.10.031.

164. Yu, D., Li, T., Delpech, J.-C., Zhu, B., Kishore, P., Koshi, T., Luo, R., Pratt, K.J.B., Popova, G., Nowakowski, T.J., et al. (2022). Microglial GPR56 is the molecular target of maternal immune activation-induced parvalbumin-positive interneuron deficits. Sci. Adv. *8*, eabm2545. 10.1126/sciadv.abm2545.

165. He, L., Arnold, C., Thoma, J., Rohde, C., Kholmatov, M., Garg, S., Hsiao, C.-C., Viol, L., Zhang, K., Sun, R., et al. (2022). CDK7/12/13 inhibition targets an oscillating leukemia stem cell network and synergizes with venetoclax in acute myeloid leukemia. EMBO Mol. Med. *14*, e14990. 10.15252/emmm.202114990.

166. Biernacki, M.A., Sheth, V.S., and Bleakley, M. (2020). T cell optimization for graft-versus-leukemia responses. JCI Insight *5*. 10.1172/jci.insight.134939.

167. Schneidawind, D., Pierini, A., and Negrin, R.S. (2013). Regulatory T cells and natural killer T cells for modulation of GVHD following allogeneic hematopoietic cell transplantation. Blood *122*, 3116–3121. 10.1182/blood-2013-08-453126.

168. Edinger, M., Hoffmann, P., Ermann, J., Drago, K., Fathman, C.G., Strober, S., and Negrin, R.S. (2003). CD4+CD25+ regulatory T cells preserve graft-versus-

tumor activity while inhibiting graft-versus-host disease after bone marrow transplantation. Nat. Med. *9*, 1144–1150. 10.1038/nm915.

169. Perez-Gutierrez, A., Metes, D.M., Lu, L., Hariharan, S., Thomson, A.W., and Ezzelarab, M.B. (2018). Characterization of eomesodermin and T-bet expression by allostimulated CD8+ T cells of healthy volunteers and kidney transplant patients in relation to graft outcome. Clin. Exp. Immunol. *194*, 259–272. 10.1111/cei.13162.

170. Hordyjewska, A., Popiołek, Ł., and Horecka, A. (2015). Characteristics of hematopoietic stem cells of umbilical cord blood. Cytotechnology *67*, 387–396. 10.1007/s10616-014-9796-y.

171. Carmeliet, P., and Jain, R.K. (2000). Angiogenesis in cancer and other diseases. Nature *407*, 249–257. 10.1038/35025220.

172. Folkman, J. (1971). Tumor angiogenesis: therapeutic implications. N. Engl. J. Med. *285*, 1182–1186. 10.1056/NEJM197111182852108.

173. Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., Bassez, A., Decaluwé, H., Pircher, A., Van den Eynde, K., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. Nat. Med. *24*, 1277–1289. 10.1038/s41591-018-0096-5.

174. Crawford, Y., and Ferrara, N. (2009). VEGF inhibition: insights from preclinical and clinical studies. Cell Tissue Res. *335*, 261–269. 10.1007/s00441-008-0675-8.

175. Caligiuri, G., and Tuveson, D.A. (2023). Activated fibroblasts in cancer: Perspectives and challenges. Cancer Cell *41*, 434–449. 10.1016/j.ccell.2023.02.015.

176. Shafat, M.S., Oellerich, T., Mohr, S., Robinson, S.D., Edwards, D.R., Marlein, C.R., Piddock, R.E., Fenech, M., Zaitseva, L., Abdul-Aziz, A., et al. (2017). Leukemic blasts program bone marrow adipocytes to generate a protumoral microenvironment. Blood *129*, 1320–1332. 10.1182/blood-2016-08-734798.

177. Chung, K.-M., Hsu, S.-C., Chu, Y.-R., Lin, M.-Y., Jiaang, W.-T., Chen, R.-H., and Chen, X. (2014). Fibroblast activation protein (FAP) is essential for the

migration of bone marrow mesenchymal stem cells through RhoA activation. PLoS ONE *9*, e88772. 10.1371/journal.pone.0088772.

178. Puré, E., and Blomberg, R. (2018). Pro-tumorigenic roles of fibroblast activation protein in cancer: back to the basics. Oncogene *37*, 4343–4357. 10.1038/s41388-018-0275-3.

179. Kraman, M., Bambrough, P.J., Arnold, J.N., Roberts, E.W., Magiera, L., Jones, J.O., Gopinathan, A., Tuveson, D.A., and Fearon, D.T. (2010). Suppression of antitumor immunity by stromal cells expressing fibroblast activation protein-alpha. Science *330*, 827–830. 10.1126/science.1195300.

180. Elsafadi, M., Manikandan, M., Dawud, R.A., Alajez, N.M., Hamam, R., Alfayez, M., Kassem, M., Aldahmash, A., and Mahmood, A. (2016). Transgelin is a TGFβ-inducible gene that regulates osteoblastic and adipogenic differentiation of human skeletal stem cells through actin cytoskeleston organization. Cell Death Dis. *7*, e2321. 10.1038/cddis.2016.196.

181. Wei, X., Lou, H., Zhou, D., Jia, Y., Li, H., Huang, Q., Ma, J., Yang, Z., Sun, C., Meng, Y., et al. (2021). TAGLN mediated stiffness-regulated ovarian cancer progression via RhoA/ROCK pathway. J. Exp. Clin. Cancer Res. *40*, 292. 10.1186/s13046-021-02091-6.

182. Hawkins, E.D., Duarte, D., Akinduro, O., Khorshed, R.A., Passaro, D., Nowicka, M., Straszkowski, L., Scott, M.K., Rothery, S., Ruivo, N., et al. (2016). T-cell acute leukaemia exhibits dynamic interactions with bone marrow microenvironments. Nature *538*, 518–522. 10.1038/nature19801.

183. Cook, D.B., McLucas, B.C., Montoya, L.A., Brotski, C.M., Das, S., Miholits, M., and Sebata, T.H. (2019). Multiplexing protein and gene level measurements on a single Luminex platform. Methods *158*, 27–32. 10.1016/j.ymeth.2019.01.018.

184. Chiswick, E.L., Duffy, E., Japp, B., and Remick, D. (2012). Detection and quantification of cytokines and other biomarkers. Methods Mol. Biol. *844*, 15–30. 10.1007/978-1-61779-527-5_2.

185. Tettamanti, S., Pievani, A., Biondi, A., Dotti, G., and Serafini, M. (2022). Catch me if you can: how AML and its niche escape immunotherapy. Leukemia *36*, 13–22. 10.1038/s41375-021-01350-x.

186. Katsumura, K.R., Ong, I.M., DeVilbiss, A.W., Sanalkumar, R., and Bresnick, E.H. (2016). GATA Factor-Dependent Positive-Feedback Circuit in Acute Myeloid Leukemia Cells. Cell Rep. *16*, 2428–2441. 10.1016/j.celrep.2016.07.058.

187. Suresh, K., and Psoter, K.J. (2023). Post hoc survival analyses using RNAseq data: handle with care. Am. J. Physiol. Lung Cell. Mol. Physiol. *324*, L1–L4. 10.1152/ajplung.00037.2022.

188. Lomakin, A., Svedlund, J., Strell, C., Gataric, M., Shmatko, A., Rukhovich, G., Park, J.S., Ju, Y.S., Dentro, S., Kleshchevnikov, V., et al. (2022). Spatial genomics maps the structure, nature and evolution of cancer clones. Nature *611*, 594–602. 10.1038/s41586-022-05425-2.

189. Wynn, T.A. (2008). Cellular and molecular mechanisms of fibrosis. J. Pathol. *214*, 199–210. 10.1002/path.2277.

190. Muz, B., de la Puente, P., Azab, F., and Azab, A.K. (2015). The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. Hypoxia (Auckl) *3*, 83–92. 10.2147/HP.S93413.

191. Hira, V.V.V., Breznik, B., Vittori, M., Loncq de Jong, A., Mlakar, J., Oostra, R.-J., Khurshed, M., Molenaar, R.J., Lah, T., and Van Noorden, C.J.F. (2020). Similarities between stem cell niches in glioblastoma and bone marrow: rays of hope for novel treatment strategies. J. Histochem. Cytochem. *68*, 33–57. 10.1369/0022155419878416.

192. Elhanani, O., Ben-Uri, R., and Keren, L. (2023). Spatial profiling technologies illuminate the tumor microenvironment. Cancer Cell *41*, 404–420. 10.1016/j.ccell.2023.01.010.

193. Karimi, E., Yu, M.W., Maritan, S.M., Perus, L.J.M., Rezanejad, M., Sorin, M., Dankner, M., Fallah, P., Doré, S., Zuo, D., et al. (2023). Single-cell spatial immune landscapes of primary and metastatic brain tumours. Nature *614*, 555–563. 10.1038/s41586-022-05680-3.

194. Jaatinen, T., and Laine, J. (2007). Isolation of mononuclear cells from human cord blood by Ficoll-Paque density gradient. Curr. Protoc. Stem Cell Biol. *Chapter 2*, Unit 2A.1. 10.1002/9780470151808.sc02a01s1.

195. Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. *20*, 296. 10.1186/s13059-019-1874-1.

196. Young, M.D., and Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. Gigascience *9*. 10.1093/gigascience/giaa151.

197. Zappia, L., and Oshlack, A. (2018). Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience *7*. 10.1093/gigascience/giy083.

198. Moerman, T., Aibar Santos, S., Bravo González-Blas, C., Simm, J., Moreau, Y., Aerts, J., and Aerts, S. (2019). GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. Bioinformatics *35*, 2159–2161. 10.1093/bioinformatics/bty916.

199. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb) *2*, 100141. 10.1016/j.xinn.2021.100141.

200. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA *102*, 15545–15550. 10.1073/pnas.0506580102.

201. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739–1740. 10.1093/bioinformatics/btr260.

202. Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics *14*, 7. 10.1186/1471-2105-14-7.

203. Haghverdi, L., Buettner, F., and Theis, F.J. (2015). Diffusion maps for high-dimensional single-cell analysis of differentiation data. Bioinformatics *31*, 2989–2998. 10.1093/bioinformatics/btv325.

204. Wolf, F.A., Hamey, F.K., Plass, M., Solana, J., Dahlin, J.S., Göttgens, B., Rajewsky, N., Simon, L., and Theis, F.J. (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol. *20*, 59. 10.1186/s13059-019-1663-x.

APPENDIX

Appendix Table 10: Buffers and media composition.

Buffer	Composition
Resuspension buffer (CB-HSPCs isolation)	1% BSA & 10mM EDTA in 1xPBS
PBST buffer	9.55 g PBS to 1 liter of water + 1ml Tween 20
Transfer buffer (Western blot)	3.03 g Tris, 14.3 g of glycine, 200 ml of MeOH to 1 liter of water
Mesenchymal Stromal cells (MSC) medium	DMEM low glucose, 10% hPL, 0.1% Ciproflaxin, 0.1% Gentamicin, 0.2% Heparin
Thaw medium	20% FBS in DMEM
Freezing medium	10% DMSO in FBS

Appendix Table 11: List of commonly used chemicals.

Product	Company	Cat. No.
Bovine Serum Albumin (BSA)	Sigma-Aldrich	A9647
Ethylenediaminetetraacetic Acid (EDTA)	Sigma-Aldrich	E5134
10X Dulbecco's Phosphate Buffered Saline (PBS)	Sigma-Aldrich	56064C-50L
Tween 20	Gerbu Biotechnik	2001
Tris	Sigma-Aldrich	17132101
Glycine	Carl Roth	39082
MeOH	Carl Roth	46272
Powdered Milk	Carl Roth	T1452

Appendix Table 12: List of reagents used for cell culture.

DMEM: Dulbecco's Modified Eagle Medium, IMDM: Iscove's modified Dulbecco's medium, RPMI: Roswell Park Memorial Institute, FBS: Fetal Bovine Serum, hPL: human platelet lysate.

Product	Company	Cat. No.
DMEM low glucose, GlutaMAX ™ Supplement, pyruvate	ThermoFisher	21885108
IMDM	ThermoFisher	21980065
RPMI	Sigma-Aldrich	R8758
FBS	Sigma-Aldrich	F5724
hPL	PanBiotech	P40-29050
Ciproflaxin	Genhunter	Q902
Gentamicin	ThermoFisher	15750060
Heparin Sodium 25000 ratiopharm (vials)	Ratiopharm	
Trypsin solution 10x	Sigma-Aldrich	59427c
$CD34^+$ Microbead kit	Miltenyi Biotec	130-046-702
PBMC (Lympho) Spin Medium	Pluriselect	60-00092-10
SepMate	STEMCELL Technologies	85450
Cell Count Kit with trypan blue	Bio-Rad	1450003

Cryo 1°C Freezing container	Nalgene	5100-0001