Single-cell Sequencing of Acute Leukemia

Thorsson, Hanna

2025

Document Version:
Publisher's PDF, also known as Version of record

Single-cell Sequencing of Acute Leukemia

HANNA THORSSON | DIVISION OF CLINICAL GENETICS
FACULTY OF MEDICINE | LUND UNIVERSITY

Single-cell Sequencing of Acute Leukemia

Hanna Thorsson

LUND
UNIVERSITY

DOCTORAL DISSERTATION

Doctoral dissertation for the degree of Doctor of Philosophy (PhD) at the Faculty of Medicine at Lund University to be publicly defended on Friday 2nd of May 2025 at 09.00 in Belfragesalen, BMC D15, Sölvegatan 19, Lund.

Main supervisor
Professor Thoas Fioretos, MD, PhD
Lund University, Sweden

Faculty opponent
Professor Brian Huntly, MD, PhD
University of Cambridge, England

Organization: LUND UNIVERSITY, Faculty of Medicine, Division of Clinical Genetics

Document name: Doctoral dissertation **Date of issue** 2025-06-02

Author: Hanna Thorsson **Sponsoring organization:**

Title and subtitle: Single-cell Sequencing of Acute Leukemia

Abstract:

Acute leukemia affects individuals of all ages and encompasses a heterogeneous group of malignancies caused by genetic alterations that give rise to leukemia stem cells (LSCs), which sustain self-renewal and drive the expansion of abnormal cells. Based on how these alterations affect maturation, acute leukemia is classified into acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL). However, leukemia is influenced not only by mutations but also by altered transcriptional programs and dynamic cellular ecosystems that shape disease evolution and treatment response. In this thesis, the aim was to investigate the cellular and molecular landscape of acute leukemia to enhance disease understanding, improve diagnostics, and identify novel therapeutic targets. Taken together, more than 800,000 single cells from 130 patients diagnosed with acute leukemia and 20 healthy donors were analyzed.

In **Article I,** a multimodal single-cell approach provided novel insights into the pathogenesis of childhood ALL, revealing distinct maturation patterns across subtypes and highlighting the particular vulnerability of *DUX4*-rearranged ALL to targeted therapies, including phosphoinositide 3-kinase (PI3K) inhibitors and CD371 chimeric antigen receptor (CAR) T cell therapy.

In **Article II**, the fine-grained detail of the single-cell data led to the identification of a novel stem cell marker, C3AR, on *NPM1*-mutated AML, the most common subtype of AML. We demonstrated that targeting C3AR with antibodies effectively eliminated AML cells while sparing normal progenitor cells, making it a promising therapeutic strategy.

Article III demonstrates that bulk AML gene expression is shaped by diverse cellular signatures better resolved with single-cell technologies, which also uncover unexpected heterogeneity beyond current genomic classification systems. Notably, *NPM1*-mutated AML could be stratified into two novel, clinically relevant subclasses, with different immune evasion properties.

The aim of **Article IV** was to investigate age-related characteristics in adult and pediatric AML using single-cell RNA-sequencing data. The analysis revealed that cellular maturation did not differ between the age groups but was strongly correlated with genetic profiles. We successfully identified putative LSC populations, and while no differences were found between adult and pediatric AML in terms of maturation, distinct transcriptional programs were observed, including higher levels of inflammatory signaling and bone marrow remodeling in pediatric AML. Preliminary findings suggest that incorporating maturation status could improve AML classification and highlight potential therapeutic opportunities for both adult and pediatric patients.

In conclusion this thesis demonstrates the power of single-cell technologies in unraveling the cellular and molecular landscape of acute leukemias, leading to the identification of novel leukemia subtypes, deeper insights into LSC biology, and the discovery of potential therapeutic targets.

Key words: acute lymphoblastic leukemia, acute myeloid leukemia, pediatric leukemia, *NPM1*-mutated AML, *DUX4*-rearranged ALL, single-cell sequencing, bioinformatics, cellular hierarchies, leukemia stem cell, target therapies, immunotherapies, CD371, C3AR, PI3K

Classification system and/or index terms (if any) Supplementary bibliographical information

Language English **Number of pages**: 105

ISSN and key title: 1652-8220 **ISBN:** 978-91-8021-695-1

Recipient's notes Price Security classification

Signature  Date 2025-03-24

Single-cell Sequencing of Acute Leukemia

Hanna Thorsson

LUND
UNIVERSITY

Cover art by Hanna Thorsson, Gunilla Westergren-Thorsson and Fanny Westergren

Department of Laboratory Medicine, Lund
Division of Clinical Genetics
Faculty of Medicine
Lund University

ISBN 978-91-8021-695-1

ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University

Lund 2025

Both of us, you and I, began as single cells
Siddhartha Mukherjee

Table of Contents

Original Articles

The thesis is based on the following articles:

Article I
Thorsson H, Henningsson R, Puente-Moncada N, Sjöström L, Ågerstam H, Peña-Martínez P, Sandén C, Rissler M, Castor A, Vibeke Marquart H, Modvig S, Paulsson K, Pronk CJ, Schmiegelow K, Hyrenius-Wittsten A, Orsmark-Pietras C, Lilljebjörn H, Fioretos T. **Single-cell genomics details the maturation block in BCP-ALL and identifies therapeutic vulnerabilities in *DUX4*-r cases**. *Blood, 2024;144(13):1399-1411.*

Article II
von Palffy S, Thorsson H, Peña-Martínez P, Puente-Moncada N, Sandén C, Blom AM, Henningsson R, Juliusson G, King B, Landberg N, Lazarevic V, Orsmark-Pietras C, Rissler M, Rissler V, Ågerstam H, Järås M, Lilljebjörn H, Fioretos T. **The complement receptor C3AR constitutes a novel therapeutic target in *NPM1*-mutated AML**. *Blood Advances, 2023;7(7):1204-1218.*

Article III
Lilljebjörn H, Peña-Martínez P, Thorsson H, Henningsson R, Rissler M, Landberg N, Puente-Moncada N, von Palffy S, Rissler V, Stanek P, Desponds J, Zhong X, Juliusson G, Lazarevic V, Lehmann S, Fontes M, Ågerstam H, Sandén C, Orsmark-Pietras C, Fioretos T. **The cellular state space of AML unveils novel *NPM1* subtypes with distinct clinical outcomes and immune evasion properties.** *Manuscript in revision*. Preprint at bioRxiv: https://doi.org/10.1101/2025.02.12.63782.

Article IV
Thorsson H, Henningsson R, Peña-Martínez P, Rissler M, Sandén C, Ågerstam H, Orsmark-Pietras C, Lilljebjörn H, Fioretos T. **Mapping the cellular state space of pediatric and adult AML unveils complex differentiation patterns and defines the subtype-specific maturation of the leukemia stem cell compartment.** *Manuscript*.

Article not included in the thesis:

Sandén C*, Landberg N*, Peña-Martínez P, Thorsson H, Daga S, Puente-Moncada N, Rodriguez-Zabala M, von Palffy S, Rissler M, Lazarevic V, Juliusson G, Hyrenius-Wittsten A, Orsmark-Pietras C, Lilljebjörn H, Ågerstam H, Fioretos T. **Aberrant expression of SLAMF6 constitutes a novel and targetable immune escape mechanism in acute myeloid leukemia.** *Nature Cancer (in revision)*.

Sandén C, Lilljebjörn H, Orsmark-Pietras C, Henningsson R, Saba KH, Landberg N, Thorsson H, von Palffy S, Peña-Martínez P, Högberg C, Rissler M, Gisselsson D, Lazarevic V, Juliusson G, Ågerstam H, Fioretos T. **Clonal competition within complex evolutionary hierarchies shapes AML over time**. *Nature Communications. 2020;11(1):579.*

Abbreviations

ADC	Antibody-drug conjugate
AI	Artificial intelligence
ALL	Acute lymphoblastic leukemia
AML	Acute myeloid leukemia
ARCH	Age-related clonal hematopoiesis
ATAC-seq	Assay for transposase-accessible chromatin using sequencing
BiTE	Bispecific T cell engager
BCP-ALL	B cell precursor lineage acute lymphoblastic leukemia
BCR	B cell receptor
CAR	Chimeric antigen receptor
cDNA	Complementary DNA
CHIP	Clonal hematopoiesis of indeterminate potential
CITE-seq	Cellular indexing of transcriptomes and epitopes by sequencing
CLP	Common lymphoid progenitor
CML	Chronic myeloid leukemia
CMP	Common myeloid progenitor
CRS	Cytokine release syndrome
DC	Dendritic cell
DNMT	DNA methyltransferase
ELN	European LeukemiaNet
FDA	Food and Drug Administration
FACS	Fluorescence-activated cell sorting
GEM	Gel beads in emulsion

GMP	Granulocyte-macrophage progenitor
HeH	High hyperdiploid
HLA	Human leukocyte antigen
HSCT	Hematopoietic stem cell transplantation
HSC	Hematopoietic stem cell
ICC	International consensus classification
IgH	Immunoglobulin heavy chain
IgL	Light chain (Immunoglobulin)
KNN	K-nearest neighbors
LMPP	Lymphoid-primed multipotent progenitor
LSC	Leukemia stem cell
MDS	Myelodysplastic syndrome
MHC	Major histocompatibility complex
MPP	Multipotent progenitor
MR	Myelodysplasia-related
MSC	Mesenchymal stromal cell
NBM	Normal bone marrow
NOS	Not otherwise specified
NK	Natural killer
PI3K	Phosphoinositide 3-kinase
scRNA-seq	Single-cell RNA sequencing
TCGA	The Cancer Genome Atlas
TCR	T cell receptor
TKI	Tyrosine kinase inhibitor
UMAP	Uniform manifold approximation and projection
UMI	Unique molecular index
V(D)J	Variable, diversity, and joining
WGS	Whole genome sequencing
WHO	World Health Organization

Introduction

Cancer

Cancer is a malignancy caused by an uncontrolled growth of abnormal cells in any part of the human body. The human body consists of over 30 trillion cells, and new cells are produced when a pre-existing cell divides into two daughter cells. Before cell division starts, a dividing cell must duplicate its genome, and this process is strictly regulated by repair and multiple proofreading mechanisms.[1] However, the system is imperfect, and errors can be introduced into the genome of the daughter cells, carrying these errors on to the next generation of cells. Such errors, known as "mutations", within the genome can result in cells having fundamental growth or survival advantages, enabling them to outcompete normal cells. These abnormal cells can invade nearby tissues, spread to other parts of the body, and disrupt normal bodily functions, ultimately culminating in death.

The excessive cell expansion results in the formation of abnormal tissue, generally referred to as cancer, neoplasms, or tumors. Cancer was first documented about 3,500 years ago in ancient Egypt, in The Edwin Smith Surgical Papyrus.[2] In the 1910s, before the structure of the genome (the "DNA double helix")[3] was known, Theodor Boveri proposed the theory that cancer was caused by genetic changes.[4] This theory was not fully supported until the 1960s, when the first structural genetic aberration was discovered: the Philadelphia chromosome in chronic myeloid leukemia.[5] Today, tumors are recognized as being highly dependent on ten biological capabilities, known as the "hallmarks of cancer", one of which is genomic instability. [6–8] More than 200 different types of cancer are now characterized, each with a unique genetic makeup that affects different body systems. This diversity is reflected in the World Health Organization's (WHO) estimation of 20 million new cancer cases and 9.7 million cancer-related deaths in 2022.[9]

This thesis will focus on hematologic malignancies, which impact the blood system. Hematological malignancies serve as an exceptional model for cancer research, offering critical insights into the complex biology of cancer. Due to the accessibility of bone marrow and blood compared to solid organs, research in this field has significantly advanced our understanding of how cancer develops. These breakthroughs are directly influencing the development of innovative cancer therapies, solidifying hematological malignancies as an indispensable model to study cancer, improve diagnosis, and to identify new treatment alternatives.

Next generation sequencing

The Human Genome Project, launched in the 1990s and completed in 2003, successfully sequenced approximately 92% of the human genome with high accuracy, marking a major milestone and the beginning of a new era in the field of genetics.[10] However, the first truly complete, gapless, human genome was not reported until 2022.[11]

After the introduction of next-generation sequencing (NGS) technologies in the 2000s, enabling rapid and highly parallel sequencing of DNA and RNA, our understanding of the genetic landscape in cancer has dramatically increased.[12] These studies have enabled detailed mapping of mutational patterns and the effects of various mutations on global gene expression.[13] With this rapid technological advancement, The Cancer Genome Atlas (TCGA) has become the most comprehensive and largest repository of cancer genomic data. This repository holds genomic information from over 20,000 tumor and matched normal samples across 33 types of cancer and currently hosts more than 2.5 petabytes of publicly available sequencing data (https://cancergenome.nih.gov), providing an extensive resource for cancer research.[14] Today, next-generation sequencing encompasses several methods, each with unique advantages and applications, such as whole-genome sequencing (WGS), targeted DNA sequencing (e.g., whole-exome sequencing and gene panels), RNA sequencing, epigenomic sequencing, and chromatin immunoprecipitation sequencing. Importantly, many of these methods are now used daily in research and in clinical applications, spanning diagnostics, prognostics, and therapeutic development.

Bioinformatics

Advancements in sequencing technology have led to a decrease in costs, resulting in the rapid generation of immense volumes of data, that require specialized expertise for analysis. This transformative field of data analysis, bioinformatics, has become fundamental for modern cancer genomic research. Bioinformatics bridges biology and data science by combining computer science, mathematics, statistics, and biology to enable the interpretation of large-scale biological data. Sequence analysis is a standardized, multi-stage process that begins at the laboratory bench (**Figure 1**). Prior to sequencing, DNA or RNA molecules are extracted from millions of cells in a biological sample of interest. These molecules undergo library preparation, which typically involves fragmentation and addition of adapter sequences and molecular barcodes needed for sequencing. Billions of DNA fragments are loaded into the sequencing instrument and these fragments are sequenced simultaneously, generating an immense amount of data.[15]

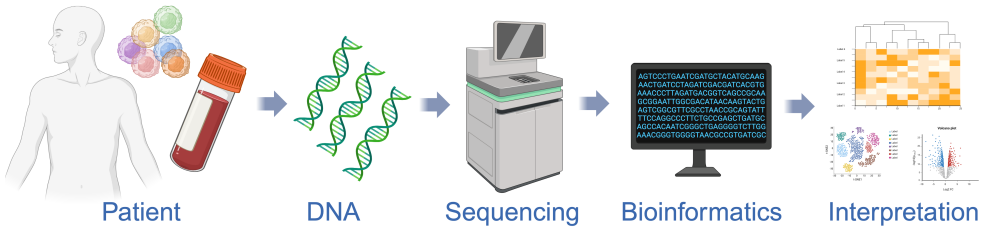

Figure 1 | Standard sequencing workflow.
A schematic illustration of the standard sequencing workflow, outlining the process from patient sample collection to bioinformatic analysis and interpretation.

To interpret these data, bioinformatics is applied. In short, this is the process of determining the precise order of nucleotides in DNA molecules, with the purpose of translating the genetic information into a readable format, allowing for the analysis of the DNA's structure and function. The first step involves preprocessing of the raw sequencing data that are converted into independent FASTQ files. FASTQ files contain sequencing reads, representing the nucleotide sequences of unique DNA molecules, tagged with a unique read name and quality scores that describe the probability that each nucleotide call is correct. Subsequently, sequence alignment is performed by aligning the reads in the FASTQ files to a human reference genome and mapping them to genes or transcripts. This results in binary alignment map files,[16] which can be further interpreted through integrative analysis (**Figure 1**).

Today, a plethora of data analysis tools exist to enable genetic data analysis, encompassing both standard preprocessing and integrative data analysis. Integrative analysis uses various analytical techniques, with key steps including data visualization, which graphically represents the data to facilitate biological interpretation.

Single-cell sequencing

Genetic data in healthy specimens and cancer have, to date, mainly been studied through bulk cell sequencing, where genetic alterations of individual cells are averaged across cell populations. While bulk analysis has offered valuable insights into genetic variability and identified several genetic drivers of tumor entities, it fails to capture the cellular heterogeneity of both normal and malignant cells within tumors. This heterogeneity likely plays a pivotal role in cancer pathogenesis and contributes to the highly individualized and often ineffective therapeutic responses observed in cancer treatment.

The emergence of powerful single-cell technologies has revolutionized cancer genomics, offering new ways to characterize cellular ecosystems by isolating individual cells before sequencing (**Figure 2**). The first full transcriptome sequencing study of a single cell was reported in 2009, involving a mouse embryonic cell.[17] In 2012, the first study describing single-cell RNA sequencing (scRNA-seq) of single cancer cells was published, analyzing 39 individual melanoma cancer cells.[18] These studies marked a milestone in genomics and lead to a broad uptake of single-cell technologies in biomedical research.

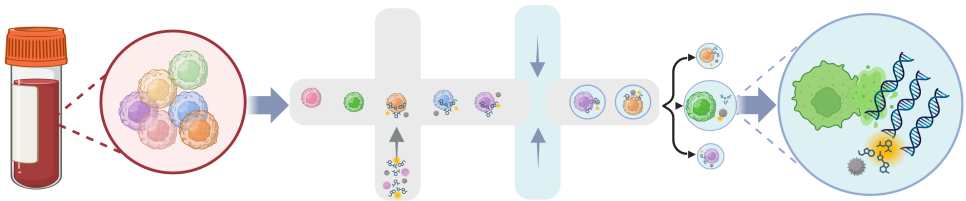

Figure 2 | Single-cell droplet isolation.
A schematic representation of single-cell isolation through droplet formation in a microfluidic device.

With continuous and remarkable advancements in single-cell technologies, two large-scale initiatives were subsequently launched in 2016 and 2018: the Human Cell Atlas (HCA) (https://www.humancellatlas.org/) and the Human Tumor Atlas Network (HTAN) (https://humantumoratlas.org/).[19,20] The HCA aims to map human cell types, including their molecular characteristics, across different tissues and organs in the healthy human body.[19] In contrast, the HTAN established a framework of single-cell atlases of tumor cells and their microenvironment across various cancers, utilizing single-cell and spatially resolved technologies alongside the collection of extensive clinical data.[20] This rapid generation of highly informative data have provided a foundation for our understanding of both healthy and cancerous cell development.

Many single-cell modalities have been developed enabling, transcriptomic, genetic, epigenetic, and proteomic studies of individual cells,[21] with some protocols also allowing multimodal measurement, for example combined readout of transcriptomic and epigenetic alterations in the same cell.[22] (see *Methodological approaches and single-cell technologies*). These advances present exciting avenues for a more comprehensive understanding of the cellular and molecular heterogeneity within a sample, allowing precise analyses across multiple dimensions within individual cells.

Normal Hematopoiesis

Daily, the human body produces 300 billion blood cells.[23] This formation of cellular blood components, known as the hematopoiesis, takes place in the bone marrow. All cells derive from the quiescent hematopoietic stem cells (HSCs) defined by self-renewal capacity and multipotency; giving rise to several lineage-restricted cells.[24,25] The definitive experimental evidence for the existence of self-renewing, multipotent HSCs capable of producing all blood cell types was provided by Till and McCulloch in 1961.[24] Since then, HSCs and the hematopoietic system have been among the most studied regenerative tissues. The essential functions of hematopoiesis, such as oxygen transport, coagulation, and immune response, are carried out by specialized, late-stage blood cells, many of which are short-lived and require continuous replenishment.[26] The high turnover of short-lived, mature blood cells demands a strictly regulated hematopoiesis, which is programmed to rapidly adapt to shifting conditions, such as infections and blood loss. This adaptability allows for the expansion and skewed production of blood cells.[27] If somehow this fine-tuned regulatory machinery of hematopoietic differentiation, proliferation, and survival is disturbed by genetic alterations, it can lead to severe hemopoietic malignancies, with fatal outcomes.

The hematopoietic hierarchy and blood lineages

The classical model of the hematopoietic system has a hierarchical structure, with repopulating HSCs residing at the apex of the hierarchy (**Figure 3**). Lineage development occurs through lineage-committed progenitor populations, such as multipotent progenitors (MPPs), common myeloid progenitors (CMPs), and lymphoid-primed multipotent progenitors (LMPPs), in a stepwise, ordered series of discrete branching steps.[28,29] As blood cells mature along the hierarchy, stemness and multipotency are lost. Progenitors differentiate into specialized cell lineages, which are broadly divided into two branches. The myeloid lineage includes granulocytes (basophils, eosinophils, neutrophils), monocytes/macrophages, dendritic cells (DCs), megakaryocytes/platelets, and erythrocytes. The lymphoid lineage comprises natural killer (NK) cells, T cells, B cells, and a subset of DCs (**Figure 3**).

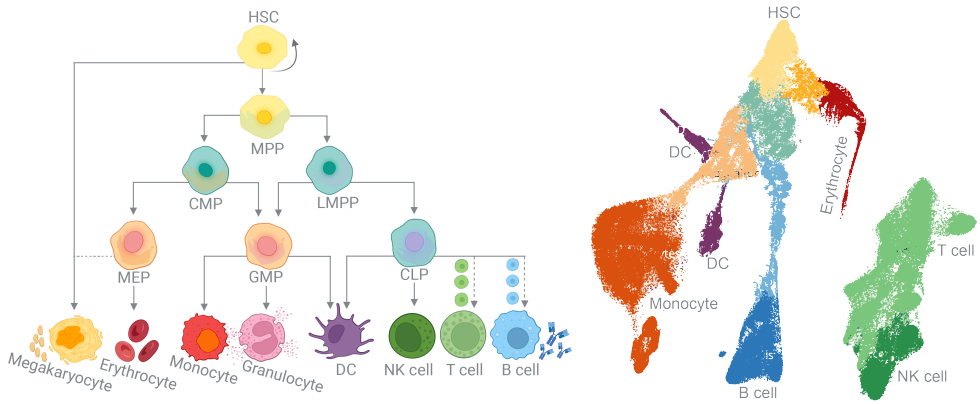

Figure 3 | The hematopoietic tree model.
The classical model of the hematopoietic system shows HSCs at the apex of the hematopoietic hierarchy, and as cells differentiate, multipotency is lost, and the cells become increasingly lineage-restricted (left). Both T cells and B cells involve a multi-step maturation process. A force-directed K-nearest neighbors graph of all hematopoietic cells from normal bone marrow samples, based on single-cell RNA sequencing data. Each cell is represented by a dot and colored according to its cell type (right). HSC; hematopoietic stem cells, MPP; multipotent progenitors, CMP; common myeloid progenitors, LMPP; lymphoid-primed multipotent progenitor, MEP; megakaryocyte-erythroid progenitors, GMP; granulocyte-monocyte progenitors, CLP; common lymphoid progenitors, and DC; dendritic cells.

The architecture of the hematopoietic system has been refined and redefined multiple times, particularly over the last few decades. The current understanding of the hematopoietic system primarily comes from the identification and isolation of progenitor populations using cell surface markers through fluorescence-activated cell sorting (FACS), colony-forming assays, and transplantation experiments. These studies initially relied on murine models, and later on humanized murine models.[30,31] As hematopoiesis is continually refined by additional studies, new modifications to the model have emerged. These include the later divergence of lymphoid and myeloid lineage fates in the differentiation hierarchy.[32,33] In addition, early megakaryocyte branching has been identified,[34] demonstrating that megakaryocyte cells can bypass the stepwise progenitor maturation process and arise directly from the multipotent HSC compartment (**Figure 3**). However, with the emergence of high-resolution single-cell technologies,[23,35–37] there has been a game-changing advancement in the understanding of hematopoietic cellular trajectories and a complete reassessment of the classical hierarchical model, which endorsed stepwise maturation. Today, hematopoiesis is suggested to follow a continuous process in which cells gradually mature and lose stemness, without distinct intermediate states, referred to as the continuum model (**Figure 4**).[38]

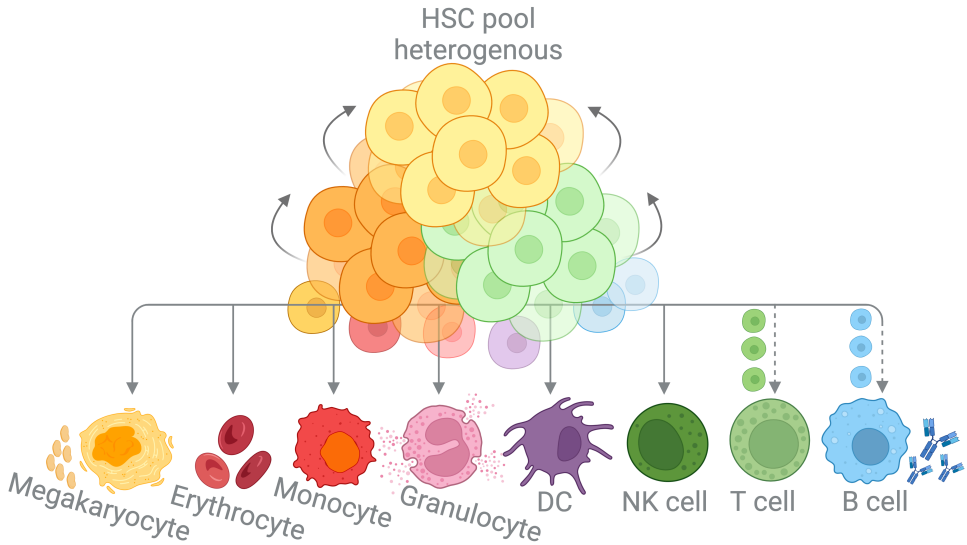

Figure 4 | The hematopoietic continuum model.
The continuum model of the hematopoietic system shows a heterogeneous HSC pool at the apex of the hematopoietic hierarchy, where cells do not pass through discrete progenitor states but instead differentiate from HSCs to mature cells in a continuous process. HSC; hematopoietic stem cells, and DC; dendritic cells.

Heterogenous HSC pool

The continuum model represents a diversity of possible routes of differentiation and emphasizes a functionally and molecularly heterogeneous HSC pool.[23,35–40] Instead of committed progenitor cells such as MPP, CMP, and LMPP, the HSC pool shows a repertoire of distinct HSC subsets, each with biases toward specific lineages and varying levels of self-renewal potential (**Figure 4**).[36,37] These biases reveal overlapping lineage potential, rather than a strict separation between myeloid and lymphoid fates, suggesting that HSC differentiation is a more flexible and plastic process. Under specific conditions, cells within the HSC pool can switch between lineages and have the capacity to transdifferentiate. As a result, various mature cell types can originate from more than a single differentiation trajectory.[38–41]

Specialized functional cells

While the heterogeneous HSC pool serves as the foundation of the hematopoietic system, ensuring continuous repopulation, specialized late-stage blood cells carry out the essential hematopoietic functions. The myeloid lineage plays crucial roles in homeostasis, tissue repair and innate immunity.[26] While erythrocytes and megakaryocytes are responsible for oxygen transport and coagulation, the

remaining myeloid cells, along with lymphoid-lineage NK cells, form part of the innate immune system. This system acts as the first line of defense, providing a rapid, nonspecific response against pathogens, foreign substances, and abnormal cells, such as cancerous cells. It is an intertwined system, with various cells working together, relying on physiological, endocytic, phagocytic, and inflammatory mechanisms.[42] In particular, monocytes play a role in orchestrating inflammation by releasing pro- and anti-inflammatory signals (e.g. IL-1β, TNF-α, IFN-γ, IL-10 and TGF- β). Some myeloid cells, such as monocytes and DCs, also function as antigen-presenting cells. Antigen-presenting cells reveal the antigen via major histocompatibility complex (MHC) molecules, encoded by human leukocyte antigen (*HLA*) genes, to T cells.[26] This process can activate the adaptive immune system, which includes T and B cells and provides a specific, targeted immune response and long-term immune memory.

The intricate maturation of T cells and B cells

Both T cells and B cells undergo a complex maturation process involving multistep receptor development through genetic rearrangements known as V(D)J recombination, which represents a hallmark of adaptive immunity.[26] This process generates a vast array of antigen receptors, T cell receptors (TCRs) and B cell receptors (BCRs), enabling these cells to recognize diverse pathogens.[26,43,44]

T cells originate in the bone marrow but fully mature in the thymus, where they undergo crucial stages of development. These include the double-negative stage, where T cells rearrange their TCR genes, followed by the double-positive stage, where they express both CD4 and CD8. They then differentiate into either $CD4^+$ or $CD8^+$ naïve T cells. Upon encountering an antigen, naïve T cells become activated, undergo clonal expansion, and differentiate into effector T cells ($CD8^+$ cytotoxic T cells or $CD4^+$ helper T cells) and memory T cells.[26,43]

Unlike T cells, the first critical stages of B cell maturation occur in the bone marrow. B cells produce antibodies, with immunoglobulin (Ig) genes serving as blueprints (**Figure 5**). During V(D)J recombination, these DNA segments are rearranged to encode the antigen-binding part of a functional membrane-bound BCR. Upon activation, B cells secrete the soluble form of the BCR, known as an antibody.[26,44] An antibody is composed of four polypeptide chains, two identical heavy chains (IgH) and two identical light chains (IgL). These are encoded from multiple gene regions, including the *IGH* locus, *IGK* locus, and *IGL* locus in humans, with the latter two encoding light chain kappa and light chain lambda (**Figure 5**). The *IGH* locus contains the gene segments V (Variable), D (Diversity), J (Joining), and C (Constant, which determines the antibody class: IgM, IgG, etc.), whereas the *IGL* locus lacks a D gene segment. Additional variability in the DNA sequences encoding IgH and IgL is generated through random nucleotide addition and trimming between the V-D, D-J, or V-J gene segments.[26,44]

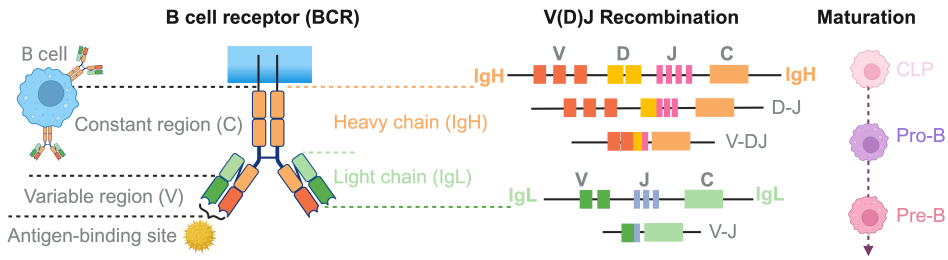

Figure 5 | V(D)J-recombination.
As B cells mature, they generate antibodies through a process called V(D)J recombination, in which immunoglobulin genes undergo DNA segment rearrangement to create diverse antigen receptors. CLP; common lymphoid progenitors.

Ig rearrangements occur gradually and are associated with discrete maturation stages during B cell differentiation. At the common lymphoid progenitor (CLP) stage, no rearrangement has occurred yet. However, in the early Pro-B cell stage, *IGH* rearrangement begins with D-J recombination, followed by V-DJ recombination in the late Pro-B cell stage. Pro-B cells then differentiate into Pre-B cells, where *IGK/IGL* rearrangement occurs between V and J segments. If successful, an early version of the BCR is expressed on the cell surface (**Figure 6**).[26,44]

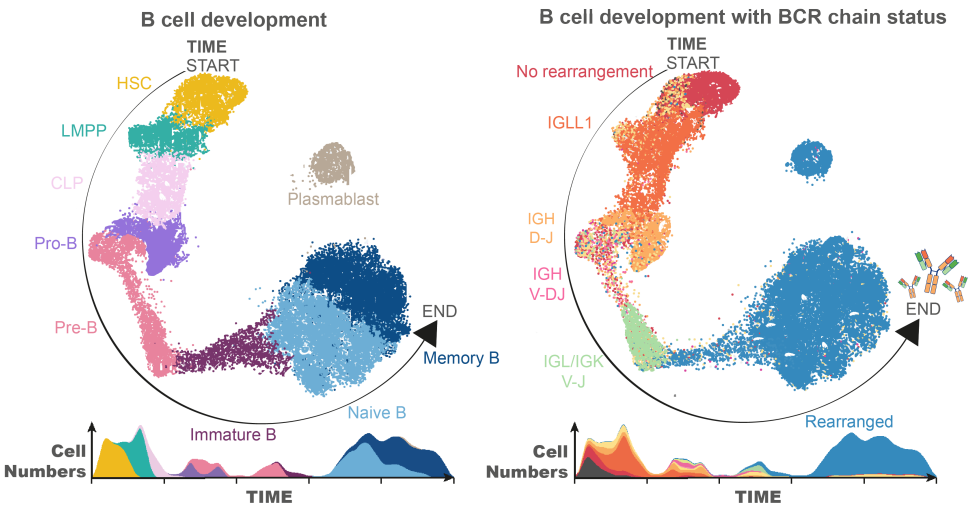

Figure 6 | V(D)J-recombination during B cell maturation..
A force-directed K-nearest neighbors graph of cells involved in B cell maturation based on single-cell RNA sequencing data, accompanied by a cell density plot showing the number of cells along the trajectory. Inferred time points are indicated along the B cell trajectory. Cells are colored by cell type (left) and B cell receptor chain status (right). HSC; hematopoietic stem cells, LMPP; lymphoid-primed multipotent progenitors, CLP; common lymphoid progenitors, BCR; B cell receptor, IGH; immunoglobulin heavy chain, and IGL; immunoglobulin light chain.

As pre-B cells transition into immature B cells, the membrane-bound BCR undergoes negative selection to ensure it does not bind to self-antigens. From this point, B cells leave the bone marrow as naïve B cells. Upon activation by antigen binding, they undergo clonal expansion and differentiate into either plasma cells, which directs large-scale antibody production, or memory B cells.[26,44]

Regulation of hematopoiesis

Together, extrinsic and intrinsic factors create a dynamic regulatory network that ensures proper hematopoietic differentiation by guiding HSCs and progenitor cells along specific pathways to produce specialized blood cells.[26] This process, known as lineage commitment,[45] occurs prior to the functional specialization of already committed cells. Extrinsic factors refer to external signals that influence hematopoietic cells, such as immune responses triggered by infection, physical forces in the bone marrow, or hypoxia. For instance, low oxygen (hypoxic) conditions influence erythropoiesis, the production of erythrocytes.[26] Extrinsic factors are typically supplied by the bone marrow microenvironment through various cell-to-cell interactions and the secretion of cytokines essential for lineage specificity.[46] Intrinsic factors refer to internal genetic, epigenetic, and molecular signals within cells that regulate their response to extrinsic cues and guide their commitment to a specific lineage or specialized function. These factors include transcription factors, epigenetic modifications, and intracellular signaling pathways.[26,45]

Bone marrow microenvironment

The bone marrow provides structural and biochemical lifelong support for hematopoietic progenitors.[46,47] These essential signals are largely supplied by specialized microenvironments known as niches (**Figure 7**). These niches form a multiparametric ecosystem composed of extracellular matrix components, multiple hematopoietic and non-hematopoietic populations, and their molecular signals. Specifically, the main components of the niche include structural extracellular matrix components, such as collagen and proteoglycans, as well as various cell types, including mesenchymal stromal cells (MSC), perivascular cells, osteoblasts, endothelial cells, and HSC-derived cells like megakaryocytes and immune cells. This matrix regulates HSC function through cell-cell interactions and secretion of factors like cytokines and growth factors, such as IL-3, IL-6, TGF-β, and CXCL12.[48,49] These non-autonomous signals are essential for the HSC pool, supporting niche retention, maintaining self-renewal, promoting quiescence, and restricting differentiation. In this context, CXCL12 is a critical chemokine primarily secreted by non-hematopoietic cells such as MSCs. CXCL12 binds to its receptor,

CXCR4, which is highly expressed on HSCs. The CXCL12-CXCR4 interaction anchors HSCs within the bone marrow and maintains them in a quiescent state. This well-studied interaction is relevant in hematological malignancies, where this signaling provides protective niches, allowing leukemic cells to evade treatments like chemotherapy.[50]

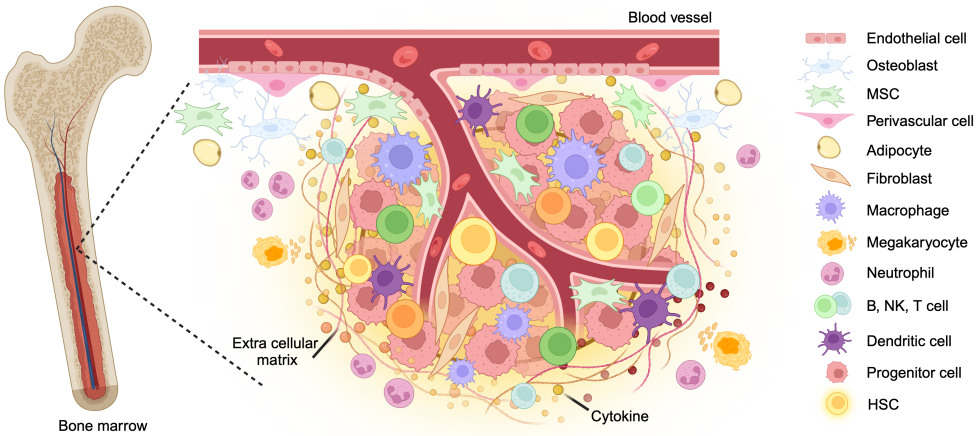

Figure 7 | Bone marrow microenvironment.
An illustration of the key components of the bone marrow niche, showcasing its complex, multiparametric ecosystem composed of extracellular matrix components, diverse hematopoietic and non-hematopoietic cell populations, and their molecular signals. HSC; hematopoietic stem cells, and MSC; mesenchymal stem cells.

Transcriptional lineage priming

Cell fate is governed by changes in gene expression, orchestrated by transcription factors that can be rapidly activated in response to external signals.[1] These transcription factors are crucial proteins that set the stage by activating or repressing gene expression programs. Master transcription factors achieve this by directly binding to specific DNA sequences, such as promoters or enhancers, acting as "commanders" that dictate cell identity. Lately, the heterogeneous HSC pool has been shown to display gene expression biases toward specific lineages,[38–41] a phenomenon proposed to result from multilineage priming. Here, transcription factors play a crucial role to ensure hematopoietic plasticity. This flexibility is achieved by fine-tuning minimal activation or inhibition of gene expression patterns for alternative lineages. During lineage commitment, one transcriptional program dominates, suppressing competing programs driven by transcription factors. One well-established cross-antagonistic interaction is that of PU.1 (encoded by *SPI1*) and GATA1.[51] When GATA1 is preferentially expressed, it results in the suppression of PU.1 and enforces differentiation toward megakaryocytic and erythroid cells, whereas PU.1 represses GATA1 and directs cells toward the

lympho-myeloid lineage. Transcription factors are universal regulators for all cells, coordinating many cellular trajectories, such as B-lymphoid lineage commitment (e.g., EBF1, ERG, IKZF1, and PAX5),[44,52] myeloid development (e.g., PU.1, C/EBPα, β, ε, and IRF8),[53] and T cell maturation (e.g., LEF1, GATA3, and TCF family factors).[43] Lately, many transcription factors have also been suggested to be mediators in T cell dysfunction through so-called epigenetic enforcement of exhaustion, such as NFAT, BATF, NR4A, and TOX-family genes.[54,55] This emphasizes the role of transcription factors in cell fate, whereas T cell dysfunction refers to a condition in which T cells fail to perform their normal functions effectively, such as recognizing and responding to cancerous cells.

Epigenetic regulation

Epigenetic modifications fine-tune hematopoiesis in response to environmental cues by controlling gene expression without altering the underlying DNA sequence. This is mainly achieved through DNA methylation, histone modification and chromatin remodeling.[56] These changes can influence cell reprogramming and the differentiation potential of HSCs, as demonstrated by clonal tracking of individual HSCs *in vivo*.[57] During lineage commitment, dynamic nature of chromatin structure regulates the accessibility of DNA sequences in the chromatin.[58,59] However, the temporal interactions between individual transcription factors and their chromatin targets during differentiation remain poorly understood.[60] Although many other key mechanisms influencing these chromatin dynamics have been described, such as the role of certain epigenetic modifiers, one example is the scaffold protein ASXL1,[61] which is heavily involved in regulating histone modifications. Other examples of epigenetic modifiers include enzymes such as TET2 and DNA methyltransferases (DNMTs), including DNMT1, DNMT3A, and DNMT3B.[62,63] These epigenetic modifiers play a role in maintaining and establishing DNA methylation patterns, which are crucial for epigenetic regulation. Furthermore, they have been linked to skewed lineage commitment. For instance, elevated DNMT expression favors differentiation toward the myeloid lineage. Mutations in the genes encoding ASXL1, DNMTs and TET2 are frequent early oncogenic events in hematological malignancies. These mutations can disrupt normal epigenetic regulation, leading to abnormal gene expression and, consequently, altered cell maturation.[61,64,65]

Signaling pathways

Signaling pathways play key roles in lineage commitment and are often influenced by both intrinsic and extrinsic signals. These pathways can be activated by extrinsic factors through interactions with receptors on the cell surface, leading to cascades of intracellular signaling events. For instance, Notch signaling is critical for T cell differentiation.[43] Moreover, common intracellular signaling molecules, such as

AKT, mTOR, and MAPK, are active in extensive signaling networks that can be subdivided into different groups.[66,67] For example, the phosphoinositide 3-kinase (PI3K)/AKT signaling pathway plays a critical role in various cellular processes. Persistent AKT signaling can promote C/EBPα-dependent mechanisms that enforce myeloid differentiation.[68] Additionally, the PI3K/AKT pathway is pivotal in the transition from early progenitor B cells to more differentiated B cells.[44,69,70] In B cells the PI3K/AKT pathway is activated as a downstream signaling cascade triggered by the pre-BCR complex on the cell surface and promotes cell survival by inhibiting apoptotic pathways. PI3K/AKT is one of many signaling pathways that interact with several others, including Notch, mTOR, and MAPK, all of which create an intertwined network capable of controlling cell survival, proliferation, and differentiation. When aberrantly activated by oncogenic events, these pathways can drive constitutive signaling, thereby facilitating the survival of cancerous cells. Hence, such dysregulation in malignancies represent vulnerabilities that can serve as basis for therapeutic interventions, as demonstrated in **Article I** and **Article II**.

Hematological malignancies

The WHO and the International Consensus Classification (ICC) have categorized malignancies of the hematopoietic system based on clinical aspects and distinct genetic alterations.[71–73] These classifications encompass over 50 different malignancies, reflecting advances in understanding these diseases. Malignancies originating in the bone marrow are referred to as leukemia. Leukemia can progress either slowly or rapidly, leading to impaired production of healthy blood cells in the bone marrow over years or just a few days, respectively. Based on disease progression, leukemia is further categorized into chronic and acute leukemia. This thesis focuses on the more rapidly and aggressively evolving leukemias, known as acute leukemias.

Acute Leukemia

Acute leukemias are highly heterogenous malignancies caused by recurrent mutations and chromosomal aberrations.[74,75] These genetic alterations disrupt normal hematopoiesis and result in an uncontrolled expansion of abnormal cells at the expense of normal blood cell formation (**Figure 8**). The expansion of abnormal cells and lack of functional blood cells results in symptoms such as fatigue, increased risk of infection, shortness of breath, and abnormal bruising and bleeding. If left untreated, acute leukemia progresses rapidly and may be fatal within weeks.[26] Acute leukemias are broadly divided into acute myeloid leukemia (AML) and acute

lymphoblastic leukemia (ALL), based on the lineage of the hematopoietic cell affected by acquired genetic alterations and how these alterations impair the maturation and differentiation of cells.[74,75] In AML, the myeloid lineage is affected, leading to the accumulation of immature myeloid progenitors. In contrast, ALL involves the expansion of lymphoid progenitor cells. In rare cases, leukemic cells express lineage markers of both myeloid and lymphoid cells; these leukemias are classified as mixed phenotype acute leukemia.[71]

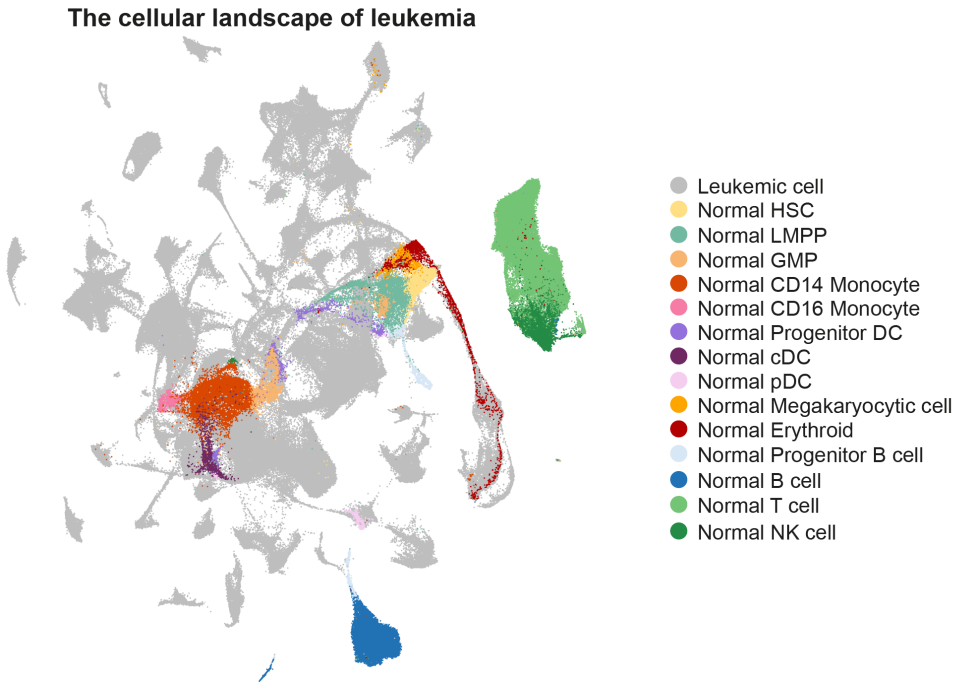

Figure 8 | The cellular landscape of acute leukemia.
UMAP visualization of single-cell RNA sequencing gene expression data representing cells from acute leukemia and normal bone marrow samples, representing a total of 569,993 cells. Every single cell represents a dot, and each cell is colored by cell type. HSC; hematopoietic stem cells, LMPP; lymphoid-primed multipotent progenitors, GMP; granulocyte-monocyte progenitors, cDC; Conventional dendritic dell and pDC; plasmacytoid dendritic cell.

AML is the most common form of acute leukemia, with a yearly incidence of approximately 4 per 100,000 inhabitants, whereas ALL has an incidence of 2 per 100,000 inhabitants.[76,77] Prior to the advent of next-generation sequencing, risk stratification in acute leukemias relied on methods such as morphology, conventional karyotyping, fluorescence in situ hybridization, immunophenotyping by FACS, and targeted molecular analyses. However, these techniques are limited in their ability to detect cryptic genetic rearrangements, structural DNA variations, and gene expression changes.[26] The outcome of acute leukemias is now understood

to be heavily influenced by the distinct types of genetic changes. Today, these genetic alterations serve as critical markers for the classification and risk stratification of acute leukemias, allowing treatment strategies to be tailored accordingly.[78]

Age at diagnosis is also a critical determinant of prognosis. Acute leukemias affect all age groups, but AML is more common in the elderly, with a median age at diagnosis of approximately 72 years.[79–81] In contrast, ALL is the most common childhood malignancy, accounting for approximately 30% of all childhood cancers, with a median age at diagnosis in children being approximately 5 years and in adults around 51 years at diagnosis.[82–85] The outcome for children is typically favorable, with a 5-year survival rate of around 75% in AML and 90% in ALL.[86–88] However, survival rates decline with age, dropping to about 50% in younger adults (<65 years) and below 10% in elderly patients (>65 years).[77,88,89] Nevertheless, the prognosis for patients with relapse is poor, regardless of age or type of leukemia.[77,86–89]

Leukemic transformation

During the last decade, scientific efforts have generated comprehensive insights into a large set of leukemia-associated mutations and their role in driving the disease. The number and type of mutations can differ drastically between different leukemia subtypes. In fact, the majority of mutations in a single leukemia are considered harmless and referred as "passenger mutations" as they are not expected to confer any growth advantage. Conversely, mutations that promote cancer development are termed "driver mutations".[13] The trajectory of cancer development is shaped by the accumulation of both driver and passenger mutations over time, with the patterns and timing of their acquisition varying both between and within different leukemia entities. Leukemia comprises a genetically diverse group of diseases, with its complexity arising from multiple somatic driver mutations that define distinct subgroups.[71–73] These mutations drive the disease pathogenesis and can therefore inform therapeutic strategies.

When a mutation occurs, it gives rise to a group of cells sharing the same genetic profile, known as a clone. Over time, new mutations can accumulate, leading to formation of subclones. A dominant clone arises when a mutation or a set of mutations confer a selective advantage.[90–92] In the context of cancer, this means that the mutation/mutations provide benefits such as faster growth and division, resistance to cell death (apoptosis), immune evasion, and/or resistance to treatments. The complexity of leukemia is partly due to the coexistence of multiple leukemia clones with distinct genetic profiles and their evolution over time.[90–94] A classical model of subclone progression follows a linear evolutionary trajectory, where the sequential acquisition of genetic alterations produces increasingly fit leukemia clones that outcompete their predecessors through selective sweeps (**Figure 9**).

However, high-throughput sequencing studies have revealed increased complexity in the clonal architecture of leukemia, which exhibits clonal evolution following a branching pattern, where a single founding clone can give rise to multiple genetically distinct subclones that evolve in parallel (**Figure 9**).[90–97] In this model, multiple subclones frequently coexist within the bulk leukemia population and actively compete for predominance. Subclones can play a crucial role in leukemia relapse. High-throughput sequencing studies have demonstrated that relapse often originates from minor subclones present at diagnosis or from rare ancestral clones that survive treatment.[90–98] These subclones can acquire additional mutations, providing a survival advantage that allows them to expand and dominate, leading to disease recurrence.

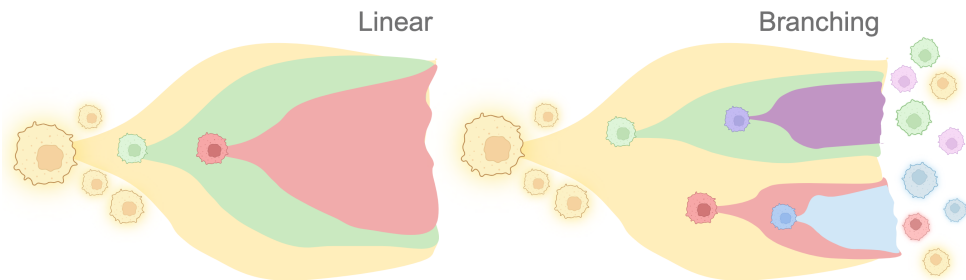

Figure 9 | Clonal evolution in leukemia.
Additional genetic alterations lead to the formation of distinct leukemia subclones, which evolve through either linear or branching paths. In linear evolution, dominant clones emerge sequentially via selective sweeps of fitter subclones. Branching evolution, however, results in multiple divergent leukemia clones.

Cancer progression can vary drastically, even within specific entities. In solid tumors and some myeloid malignancies, such as certain types of AML, mutations often accumulate gradually over time before a driver event occurs.[93] In contrast, ALL often develops following a single strong genetic event. As the disease progresses, however, new mutations may emerge, some of which can confer a survival advantage and lead to the formation of new subclones.[99,100]

The pre-leukemia model

In adult AML, clonal hematopoiesis is a common phenomenon characterized by the progressive expansion of clones with enhanced fitness over time.[101–103] This process often begins with an initial mutation in a HSC or a progenitor cell leading to the formation of pre-leukemia stem cells (pre-LSCs; **Figure 10**).[93] Temporal patterns of preleukemic mutation acquisition most commonly start with epigenetic modifiers (*DNMT3A*, *TET2* and *ASXL1*), potentially followed by splicing factors (*SF3B1* and *SRSF2*) and apoptosis regulators (*TP53* and *PPM1D*).[104] The pre-LSC

cells retain multilineage potential and can produce cells of all hematopoietic lineages, as demonstrated by the presence of mutations in non-leukemic lineages, such as T cells.[93] Pre-LSCs can remain clinically silent for years but lay the groundwork for leukemic transformation when additional cooperating mutations are acquired. These mutations drive the progression from clonal hematopoiesis to full-blown leukemia by altering the bone marrow microenvironment and promoting LSC dominance, which acquire unlimited self-renewal and proliferative capacity (**Figure 10**). At this stage, a block in differentiation results in the accumulation of leukemic blast that are produced by the LSC.[65,93] Secondary events or cooperating mutations required for leukemic transformation typically include driver mutations, such as *NPM1* or activating mutations in tyrosine kinase receptors like *FLT3*.[105]

A pre-leukemic stage is most common in AML, but can also exist in pediatric ALL. For example, 1-5% of healthy newborns are carrying an *ETV6::RUNX1* fusion gene.[106–108] This alone is not enough to cause leukemia. Secondary genetic events (such as *CDKN2A* deletions) are typically required to develop ALL.[106,109]

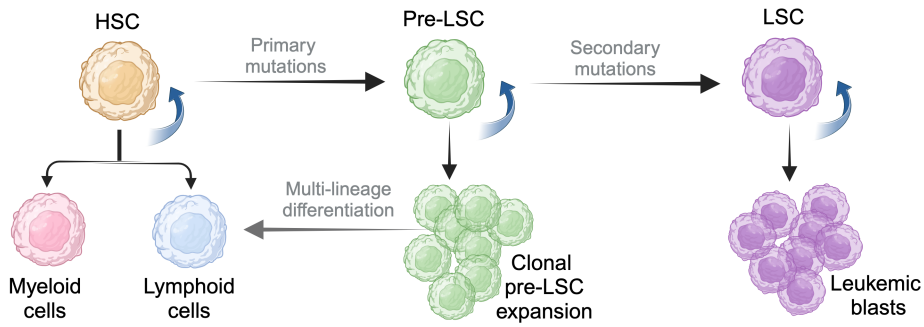

Figure 10 | The model of leukemic transformation.
Primary mutations in HSCs can transform them into pre-LSCs, which exhibit enhanced self-renewal and clonal expansion while retaining multilineage potential. The acquisition of secondary mutations further transforms pre-LSCs into LSCs, which drives the development of full-blown leukemia. HSC; hematopoietic stem cells and LSC; leukemia stem cells.

Clonal hematopoiesis is common in older adults and is referred to as age-related clonal hematopoiesis (ARCH).[101–103] ARCH has been identified in 50% of individuals over the age of 85.[110] If an ARCH associated mutation occurs in a candidate leukemia driver gene and the variant allele frequency exceeds 2%, the condition is referred to as clonal hematopoiesis of indeterminate potential (CHIP).[111] CHIP is suggested to act as a pre-malignant state. However, approximately 10% of healthy individuals over the age of 70 have a detectable CHIP.[112] In fact, when applying sensitive high-throughput sequencing techniques that allow for the detection of mutations below 2% variant allele frequency, CHIP clones can be

detected in nearly all elderly adults.[113] Despite this, only a small fraction of individuals with these mutations progress to AML, indicating that additional genetic or environmental factors are required for malignant transformation.

In comparison to solid tumors,[114] relatively few mutations are required for the development of acute leukemia, with an average of 13 coding mutations per patient in adult AML, as reported by the TCGA and Beat AML studies.[115–117] The influence of age on genomic instability is evident when comparing pediatric and adult AML. Pediatric AML typically harbors fewer mutations, averaging around 5 somatic coding mutations per patient,[117] and is more often driven by structural chromosomal alterations rather than the accumulation of point mutations seen in adults. Similar trends are observed in ALL, where the mutational burden is generally even lower than in AML.[100,118,119] The low mutational burden in childhood acute leukemia is partly due to the fact that genetic mutations accumulate with age, as evidenced by ARCH. Additionally, adults are exposed to environmental and lifestyle risk factors that further increase the likelihood of mutations.

Acute myeloid leukemia

AML was the first cancer to be fully characterized by WGS, with pioneering studies by TCGA leading to the discovery of multiple novel disease-driving alterations.[115,120] With continuous advances in molecular profiling, more AML subtypes are being genetically defined and redefined. The distribution of AML subtypes in adults is generally difficult to define, as it is largely influenced by age, even within adult and pediatric populations (**Table 1**).[115–117,121]

Table 1 | Summary of genetically defined subtypes in AML.
Distribution of genetically defined subtypes in AML according to WHO 5th edition in adult AML and pediatric AML.[73,122]

Defining genetic abnormalities	Epidemiology
AML with *NPM1* mutation	~33% in aAML, 6-8% in pAML
AML-MR (including cytogenetic abnormalities and somatic mutations)*,†	24-25% in aAML, less frequent in pAML
AML with *CEBPA* mutation	5-11% in aAML, ~5% in pAML
APL with *PML*::*RARA*	5-8% (more prevalent in younger patients, frequency declines with age)
AML with *CBFB*::*MYH11*	5-8% (more prevalent in younger patients, frequency declines with age)
AML with *KMT2A*-r	2-3% in aAML, ~20% in pAML
AML with *NUP98*-r	2-3% in aAML, 2-5% in pAML
AML with *MECOM*-r	1-2% in aAML, ~2% in pAML
AML with *RUNX1*::*RUNX1T1*	1-5%
Additional, rare (<2%) entities: AML with *DEK*::*NUP214*, *RBM15*::*MRTFA*, *BCR*::*ABL1*, *CBFA2T3*::*GLIS2* (more prevalent in pAML), *KAT6A*::*CREEBP, FUS*::*ERG*, *MNX*::*ETV6* (predominately in pAML) and *NPM1*::*MLF1*	
Additional somatic mutations not included as defining entities by WHO 5th edition	
AML with *TP53* mutations	10% aAML, < 1% pAML

aAML; adult acute myeloid leukemia, pAML; pediatric acute myeloid leukemia, APL; acute promyelocytic leukemia, MR; myelodysplasia-related, -r; rearranged, and WHO; world health organization.
*AML-MR cytogenetic abnormalities includes; complex karyotype (≥3 abnormalities), 5q deletion, Monosomy 7, 7q deletion, 11q deletion, 12p deletion, Monosomy 13, 13q deletion, 17p deletion, isochromosome 17q, and idic(X)(q13)
†AML-MR somatic mutations includes mutations in; *ASXL1, BCOR, EZH2, SF3B1, SRSF2, STAG2, U2AF1*, and *ZRSR2*.

The main subtypes in adult AML are defined by mutations in *NPM1,* myelodysplasia-related (MR) genes, *CEBPA* and *TP53* AML.[115,116] Pediatric AML has historically been less well-characterized compared to adult AML. However, recent advances have refined genomic classifications in this age group.[117,121,123] Pediatric AML shows a bias toward gene fusions, with *KMT2A* rearrangements in 20% being among the most prevalent driver events. These occur far more frequently in pediatric AML than in adult AML, where *KMT2A*-rearranged (*KMT2A*-r) AML is represented at a low frequency (~2-3%). Events defined by mutations are less

frequent in pediatric AML with exceptions such as *NPM1*, *CEBPA*, *FLT3* and the newly identified tandem duplications in *UBTF*.[124] For example, AMLs with mutated *TP53* are very rare, occurring in <1% of pediatric AML cases.[121] The high prevalence of AML with other rare recurring translocations' in pediatric cases, accounting for up to 16%,[121] is another distinction that emphasizes unique biology of pediatric AML compared to adult AML. Moreover, with current classification systems, AML cases that do not fit into any of the well-defined genetic, cytogenetic, or clinical categories are categorized as either "acute myeloid leukemia with other defined genetic alterations" or AML, not otherwise specified (NOS). This category functions as a residual group, accounting for approximately 15% of all AML cases.[116,121] This highlights the need for further refinement, and over time, with improved genomic and molecular characterization, the proportion of AML cases classified as NOS is likely to decrease.

Aside from important driver events that form the basis for categories in classification systems, other recurring mutations also significantly influence leukemic progression.[94,115,121] These can be further separated into groups that confer various biological functions, such as activated signaling gene mutations (e.g., *FLT3-ITD, FLT3-TKD, JAK2,* and *NOTCH*), which can also be grouped with RAS pathway gene mutations (e.g., *NRAS, KRAS,* and *PTPN11*). Other categories include DNA methylation-related gene mutations (e.g., *DNMT3A, TET2, IDH1,* and *IDH2),* chromatin-modifying gene mutations (e.g., *ASXL1* and *EZH2*), spliceosome-complex gene mutations (e.g., *SRSF2, SF3B1, U2AF1*, and *ZRSR2*), cohesin-complex gene mutations (e.g., *STAG2* and *RAD21*), cell cycle gene mutations (e.g., *CCND1, CCND2,* and *DNM2*), and transcriptional regulator gene mutations (e.g., *RUNX1, CEBPA, GATA2, IKZF1,* and *MYC*).

Risk stratification in AML

Genomic classification and clinical aspects have formed a critical basis for risk categories refined over the past decades. The first classification of AML, proposed by the French-American-British system, solely stratified cases based on blast morphology and the maturation stage of the cells.[125] Today, risk stratification relies on two genetic classification systems. While the WHO and ICC share many overlapping criteria,[71,73] notable discrepancies exist that may impact patient treatment regimens and outcomes. The ICC further subdivides certain groups within the WHO classification, as exemplified by the AML-MR group. While the WHO considers AML-MR a single entity, the ICC divides it into three subgroups: AML with MR cytogenetic abnormalities, AML with MR gene mutations (including *RUNX1*), and AML with mutated *TP53*. Other discrepancies in genetic subtypes apply to AML with *CEBPA* mutations, *KMT2A*-r, *MECOM*-r, and *NUP98*-r AML. Another difference lies in the diagnostic thresholds, with the WHO separating AML and Myelodysplastic Syndrome (MDS), while the ICC introduces MDS/AML as a transitional category. All these definitional differences can correlate with variations

in overall survival,[126–128] particularly for AML with mutated *TP53*, which is considered one of the subtypes with the poorest prognosis.[94] At present, reliance on two classification systems remains necessary. However, as research advances, a unified AML classification would be preferable to enhance consistency and comparability across diagnostic centers.

Today, the European LeukemiaNet (ELN) is an international consortium that integrates both the ICC and WHO classification systems and provides standardized risk stratification and treatment guidelines.[78] The ELN divides genomic classes into three risk categories: favorable risk, intermediate risk, and adverse risk, tailoring treatment regimens accordingly. However, the current classification systems and risk stratifications are predominantly based on evidence from adult AML. Given the distinct molecular characteristics of pediatric AML, there is ongoing debate about the need for an additional, adapted genetic framework specifically for children.[117,121] This discussion also extends to the development of age-tailored risk stratification strategies and alternative treatment approaches.[129–132]

Acute lymphoblastic leukemia

ALL is thought to arise from the B cell precursor lineage (BCP-ALL) or, less commonly, the T cell precursor lineage (T-ALL), and comprises over thirty genetically distinct subtypes of prognostic importance.[99,100,133,134] T-ALL is less common in children, representing approximately 12% of pediatric ALL cases, compared to around 25% in adults. T-ALL can be further subclassified into 10 subtypes,[100,135] which are not depicted in this thesis. The characteristics of BCP-ALL will be discussed in more detail.

In total, BCP-ALL cases are classified into approximately 25 different molecular groups,[99,100] recognized by both the WHO and ICC.[71,72] The WHO and ICC classifications are relatively well-aligned and present a more uniform terminology for BCP-ALL subtypes. These subtypes include those defined by single-point mutations, recurrent chromosomal gains and losses, and chromosomal rearrangements. These rearrangements often deregulate genes through the formation of chimeric fusions, many of which are frequently associated with hematopoietic transcription factors. Furthermore, these driver lesions can be broadly categorized into chromosomal abnormalities, transcription factor rearrangements, other transcription factor alterations, and kinase-driven alterations.[99,100,133,134] The distribution of these lesions varies significantly across different age groups (**Table 2**).[72,122]

Table 2 | Summary of genetically defined subtypes in BCP-ALL.
Distribution of genetically defined subtypes in BCP-ALL according to WHO 5th edition in adult ALL and pediatric ALL.[72,122]

Defining genetic abnormalities	Epidemiology
BCP-ALL with *KMT2A*-r	70-80% in infant ALL (<1y)
BCP-ALL with HeH	25-35% in pALL, 7-8% in aALL
BCP-ALL with *ETV6*::*RUNX1*	~25% in pALL (rare in infants and adults)
BCP-ALL with *BCR*::*ABL1*	Incidence increases with age (2-4% in <15 y, ~10% in 15-39 y, ~25% in 40-49y and 20-40% in >50 y)
BCP-ALL with *BCR*::*ABL1*-like	10-15% in pALL, 25-30% in adolescents and young adults, 20-25% in aALL
BCP-ALL with other defined genetic alterations (e.g. *DUX4*-r, *MEF2D*-r, *ZNF384*-r, *PAX5* alt, *PAX5* p.80R, *NUTM1*-r, *MYC*-r)	10-15% in pALL, 20-35% in aALL
BCP-ALL with *TCF3*::*PBX1*	~5% in pALL (rare in adults)
BCP-ALL with iAMP21	~2% in pALL (rare in adults)
Additional entities: BCP-ALL with hypodiploidy (<1% in pALL, >10% in aALL), *ETV6*::*RUNX1*-like (1-3% in pALL, uncommon in adults), *IGH*::*IL3* (<1%), *TCF3*::*HLF* (<1%)	

BCP-ALL; B cell precursor acute lymphoblastic leukemia, aALL; adult acute lymphoblastic leukemia, pALL; pediatric acute lymphoblastic leukemia, -r; rearranged, HeH; high hyperdiploidy, and iAMP21; isolated amplification of chromosome 21.

There could be multiple reasons for the opposite genetic patterns observed in children and adults diagnosed with ALL. One explanation may be that certain genetic lesions, such as *ETV6::RUNX1* and high hyperdiploidy (HeH), often arise during fetal development.[100] Conversely, *BCR::ABL1* and *BCR::ABL1*-like ALL are more prevalent in adults, likely due to an increasing risk over time that the fusion will arises in a cell vulnerable to the transforming event.

The presence of many "genetic-like" subtypes in ALL, such as *BCR::ABL1*-like and *ETV6::RUNX1*-like ALL reflects the complexity of genetic and molecular heterogeneity in the disease. These classifications are based on the fact that different genetic alterations can mimic the biology and clinical outcomes of well-defined subtypes without necessarily sharing the same primary genetic lesion. This

phenomenon highlights that leukemia is driven by networks of mutations rather than single genetic events, and that different mutations can lead to similar disease behavior. Similar to AML, there are still subtypes of BCP-ALL lacking defining lesions, collectively known as B-other ALL or BCP-ALL NOS.[71,72] These undefined subtypes represent a smaller proportion than in AML, accounting for approximately 10% of all BCP-ALL cases.[99,100]

Secondary genetic alterations are also essential in leukemogenesis and treatment response and vary both between and within BCP-ALL subtypes. Common targets include B cell pathway genes important for B-lymphoid lineage commitment (e.g., *EBF1, IKZF1, PAX5*, and *RAG1/2*), RAS pathway alterations (e.g., *KRAS, NRAS*, and *PTPN11*), transcription factors (e.g., *ERG, ETV6, LEF1*, and TOX), the JAK-STAT pathway (e.g., *JAK1, JAK2*, and *JAK3*), and epigenetic modifications (e.g., *CREBBP, KMT2D, SETD2,* and *TBL1XR1*).[100] Additionally, a handful of structural variations involve the non-coding genome, a field that remains largely unexplored but holds potential for future discoveries.

Risk stratification in ALL

Unlike AML, which has a well-defined classification and risk-stratification system like the ELN, ALL lacks a single, standardized classification to guide treatment decisions. Instead, several risk stratification systems are used in existing guidelines, typically with separate protocols for children and adults. One such example is the ALLTogether protocol (clinicaltrials.gov: NCT04307576), a collaborative initiative involving multiple European countries, which offers a standardized treatment approach for children and young adults. In this protocol patients are classified into four risk groups: standard risk, intermediate-low risk, intermediate-high risk, and high risk.

From a future perspective, today's genetic classifications will likely continue to evolve, redefining risk categories and further shrinking the AML NOS and B-other categories. Ultimately, this will lead to new risk category systems for more precise risk assessment and refined targeted treatment approaches.

Leukemia stem cell biology

A central objective in leukemia research is the characterization and selective targeting of LSCs, which play a crucial role in leukemia initiation, progression, and relapse. Gaining deeper insights into their biology and developing precise therapies against them remains the ultimate goal in leukemia research, as it holds the potential to achieve a lasting cure for these malignancies.

The cell of origin and the concept of leukemia stem cells

The "cell of origin" refers to the normal cell (e.g., a HSC or progenitor cell) in which the initial leukemogenic mutation arises.[136] When the full leukemic transformation has occurred this will give rise to LSCs that are responsible for the long-term maintenance of the disease and are believed to persist and cause relapse after treatment.[90–98]

Conceptually, LSCs are functionally defined by their ability to engraft and sustain leukemia in immunodeficient mice upon serial transplantation.[137,138] The first experimental evidence for the existence of LSCs in AML was provided by Bonnet, Dick, Lapidot, and colleagues, who demonstrated that not all leukemic cells have the functional capacity to generate leukemia in immunodeficient mice. The rare LSC population was identified within the same broad immunophenotypic compartment as normal HSCs ($CD34^+CD38^-$).[137,138] Further studies on LSCs in AML have demonstrated the presence of LSCs in the $CD34^-$ compartment, primarily in leukemias with *NPM1* mutations and *KMT2A* rearrangements.[139–141] The concept of LSC has proven challenging to fully implement experimentally, leading to the development of surrogate assays to approximate LSC capacity. Studying LSCs in immunodeficient mice presents additional challenges, as the degree of murine immune system compromise, the improvement of engraftment through the expression of human cytokines, and the genetic characteristics of the leukemia can all influence engraftment and stem cell properties.[142–144] Consequently, stem cell properties may be context-dependent, further complicating efforts to define LSCs consistently.

Transplantation studies in immunodeficient mice have provided evidence that LSCs are also present in ALL.[145,146] It has been reported that the cell of origin in ALL varies between HSCs and committed lymphoid progenitors.[147] Unlike AML LSCs, which are typically enriched in specific immature immunophenotypic populations (e.g., $CD34^+$ and $CD38^-$), BCP-ALL LSCs are found across a range of maturing immunophenotypic populations,[148] commonly within the $CD34^+$, $CD38^-$, and $CD19^+$ cell compartments. LSCs in ALL often exhibit characteristics similar to those of lymphoid progenitor cells and are generally more differentiated than LSCs in AML. Gene expression data from primary AML samples have, however, also revealed

LSC potential in various cellular subsets, including populations resembling healthy LMPPs and granulocyte-macrophage progenitors (GMPs), rather than HSCs.[149] Interestingly, normal HSC behavior changes with age, with HSCs in elderly being primed towards the myeloid lineage,[150] which may explain the more committed behavior of AML LSCs and the age-related discrepancy between ALL and AML. Together, these findings suggest that LSCs may either arise from committed progenitors that aberrantly acquire self-renewal capacity and multipotency or originate from HSCs that transitions into LSCs generating more differentiated and defective progenitors.

Phenotypic characteristics of leukemia stem cells

Despite HSCs and LSCs similarities, cell surface markers distinguish LSCs from HSCs, highlighting cell-intrinsic differences. Specifically, AML LSCs can exhibit higher levels of CD25, CD32, CD44, CD96, CD123, CLL-1, IL1RAP, and C3AR.[151–153] In **Article II**, C3AR was shown to be expressed both on the LSC and the progenitor cells in *NPM1*-mutated AML. In addition, the monocytic marker CD33 is also more highly expressed on AML cells compared to HSCs, although its expression is lower on LSCs compared to bulk cells.[152] In ALL, LSC markers vary depending on the leukemia subtype, with many being associated with either T cell or B cell phenotypes.[154,155] However, some subtypes deviate from this pattern and express additional markers. For example, *BCR::ABL1*-positive ALLs commonly express CD123,[156] which is a LSC marker also in AML.[152] In addition, *DUX4*-rearranged (*DUX4*-r) ALLs express CD371 (CLL-1),[157] which is also recognized as a LSC marker in AML,[158] and to be expressed on healthy monocytic cells.[159] Some of these unique LSC surface markers play a crucial role in clinical practice, including immunophenotypic diagnostics and minimal residual disease monitoring. However, these markers are also important as they may serve as targets for antibody-based therapies and cellular therapies.

Leukemia stem cell signatures and cell hierarchies

Identification of "stemness" signatures has showed notable prognostic value and superiority in predicting outcomes and treatment resistance in AML patients.[160–166] One of the most well-established signatures is LSC17, which consists of 17 genes highly expressed in both LSCs and healthy HSCs.[162,167] (**Figure 11**). Overall, LSC17 is clinically relevant across all ages, but its predictive power may be even more critical in older AML patients who have fewer treatment options.[161,168,169]

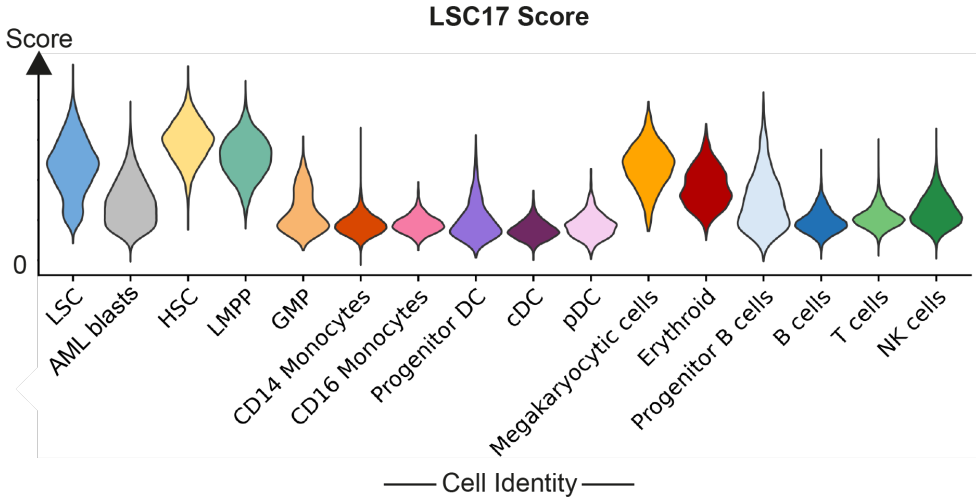

Figure 11 | Expression of LSC17 genes in different cell types.
Expression of the LSC17 gene set in distinct cell types from AML samples and healthy bone marrow samples, based on single-cell RNA sequencing data. LSC17 is a stemness signature composed of 17 genes that are highly expressed in stem cells and has demonstrated notable prognostic value. LSC; leukemia stem cell, AML; acute myeloid leukemia HSC; hematopoietic stem cells, LMPP; lymphoid-primed multipotent progenitors, GMP; granulocyte-monocyte progenitors, cDC; Conventional dendritic dell and pDC; plasmacytoid dendritic cell.

Lately, the prognostic value of the LSC17 signature has been suggested to be limited when used alongside conventional genetic risk stratification. Instead, a 47-gene LSC signature (LSC47) was found to be more effective and to improve risk prediction, particularly when considering molecular subtypes like *KMT2A*-rearrangements, *CEBPA* mutations, and *FLT3*-ITDs alterations.[161] While gene expression scores, such as the LSC17 and LSC47 gene signatures, and the more recent 3-Gene, APS, CODEG22 and pLSC6 signatures,[163–166] are valuable prognostic tools in AML, it has been becoming increasingly clear that a more detailed understanding of leukemia cell hierarchies may improve predictions of drug response and disease progression.[170,171,121]

LSCs that sustain leukemia result in the expansion of immature cells that exist in various maturation states,[170,171] as also emphasized in **Article I, Article III,** and **Article IV**. These maturation states influence their response to anti-cancer drugs.[171] For instance, certain drugs have been found to be more effective against AML with less mature cell states, while others work better against more mature forms. This variation underscores the necessity of considering cell maturation stages when developing treatment strategies. Thus, moving beyond traditional gene signatures like LSC17, or integrating LSC17 with hierarchy-based biomarkers, may enable more precise stratification of AML patients. In this context, a comprehensive mapping of all cellular states is essential for understanding the cellular hierarchies

of LSCs and the leukemic progenitor pool. Recently, more than 80 distinct subsets of hematopoietic stem and progenitor cells have been identified using machine learning-based refinement and single-cell sequencing techniques.[172] This represents the most detailed view of normal hematopoietic stem and progenitor cells to date. As a proof of concept, this study also aligned healthy HSC counterparts with LSCs and identified discrete LSC states by categorizing them into primitive and more mature LSCs. Collectively, these studies highlight the growing importance of understanding LSC heterogeneity and integrating hierarchical frameworks with molecular signatures to predict treatment response.

Leukemia stem cell plasticity

Lately, lineage plasticity has been described in more detail, especially in the context of treatment or relapse. For example, venetoclax-resistant AML can, in some instances, arise from a monocytic LSC rather than a more primitive LSC driving the disease initially.[173,174] Additionally, single-cell studies of pediatric AML samples have revealed that transcriptional networks can shift from myeloid to lymphoid programs upon relapse.[175] In pediatric ALL, single-cell multiomics revealed increased plasticity and the emergence of resistant populations in *KMT2A*-r leukemia.[176] This plasticity was correlated with age, where leukemic cells from patients under six months old exhibited significantly higher lineage plasticity during treatment compared to those from older patients. These two latter studies also described lineage switching during treatment between AML and ALL,[175,176] respectively, highlighting the dynamic nature of leukemia. Lineage switching cells were also observed in **Article I** in one diagnostic ALL sample, even before treatment.

Altogether, LSC and leukemic progenitor populations exhibit remarkable plasticity, encompassing both phenotypic and functional diversity, much like the heterogeneous HSC pool in the continuum model of normal hematopoiesis.[38] Similar to the heterogenous HSC pool, LSC plasticity is partly regulated by transcriptional programs and epigenetic deregulation, both of which play crucial roles in disease evolution and the establishment of intratumoral heterogeneity.[177] Understanding these transcriptional programs and epigenetic alterations may offer therapeutic avenues to target the disrupted processes.

The leukemic bone marrow niche

Similar to HSC regulation, the LSC pool is shaped by a dynamic interplay between cell-intrinsic mechanisms and extrinsic signals from the microenvironment. The bone marrow niches serve as a sanctuary for LSCs and malignant cells, supporting their clonal evolution and reinforcing both drug resistance and leukemia relapse.[178]

This complex regulatory network, which encompasses the adhesion between leukemic cells and their microenvironment as well as signal transduction, has gained increasing attention with the advent of single-cell techniques. First, single-cell studies have shown that LSCs in AML and malignant cells in ALL activate genes involved in signaling pathways typically associated with non-hematopoietic cell lineages, such as *VEGF* and IL-5 in AML,[179] and *MCAM* and *PDGFRA* in ALL,[180] which are involved in angiogenesis and tissue interactions.[181–183] This aberrant activation of lineage-inappropriate signaling by LSCs emphasizes their high plasticity and may contribute to their growth and persistence within the bone marrow microenvironment. Secondly, other single-cell studies have focused on defining the immune microenvironment. For example, a recent study identified a nonclassical monocytic cell population in the ALL microenvironment that supports leukemic clone expansion, further highlighting the complex interactions at play.[184] This myeloid dependency in ALL has also been observed in T-ALL, where leukemia-associated myeloid cells are required for leukemia cell survival. Notably, depletion of these myeloid cells significantly reduces leukemia progression.[185]

In addition, inflammation has emerged as a crucial player in shaping the microenvironment. In AML, samples from patients in complete remission show a higher presence of inflammatory M1 macrophages.[186] In contrast, high-inflammation AML patients exhibit an enriched dysfunctional B cell subtype, alongside an increase in $CD8^{+}$*GZMK*$^{+}$ and regulatory T cells.[187] These findings suggest that the inflammatory state remodels the immune microenvironment in leukemia, potentially influencing disease progression and patient prognosis. Additionally, immune cell dysfunction is increasingly recognized in both AML and ALL,[180,187–190] when examining the leukemic microenvironment, which may represent a novel immune escape mechanism. In **Article I**, a distinct $CD4^{+}$ T cell subset displaying transcriptional signatures of dysfunction was identified within the leukemic niche in ALL cases.

Another aspect of microenvironment remodeling is the impact of aging.[191,192] In children, the bone marrow niche supports rapid cell growth, which may create a favorable environment for specific leukemia subtypes. However, in adults, the immune system and bone marrow function decline with age, potentially allowing the expansion of aggressive leukemic clones. Whether this shift contributes to the generally better outcomes observed in pediatric leukemia patients remains to be elucidated. Taken together, investigating the mechanisms of microenvironment remodeling in leukemia and understanding how LSCs and other leukemic cells exploit and reshape their niche may help identify essential pathways that drive disease progression, drug resistance, and relapse. Targeting these pathways could provide new therapeutic strategies to improve treatment response and patient outcomes.

Therapies of Acute Leukemia

Despite decades of groundbreaking advancements and relentless efforts in leukemia research, intensive chemotherapy remains the backbone of treatment. Chemotherapy drugs target all rapidly dividing cells, a hallmark of leukemia cells. However, these drugs lack specificity, and severely impact also healthy tissues, which leads to considerable morbidity and long-term adverse effects. Despite the aggressive nature of chemotherapy, residual leukemia cells that are not actively dividing will not be eliminated, which is believed to contribute to the high relapse rates in leukemia.[90–98] Residual leukemic cells adapt to survive the cytotoxic assault through various molecular mechanisms, leading to resistant and potentially more aggressive disease, where therapeutic options become increasingly limited, often leaving patients with no viable treatment options.

Despite the challenges associated with chemotherapy, ALL is considered a success story in pediatric oncology, with cure rates exceeding 90%.[74,88] This success has been achieved through the optimization of dosing and scheduling of the same chemotherapeutic agents that have been in use for the past five decades. For AML, modest improvements in survival have primarily been attributed to the optimization of curative therapies over the years, including intensive chemotherapy and allogeneic hematopoietic stem cell transplantation (HSCT),[193] in which the patient's hematopoietic system is replaced through the transplantation of HSC collected from healthy donors. The decision to pursue HSCT is influenced by factors such as the patient's age, overall health (performance status), and specific risk factors.[78] According to the ELN treatment guidelines, standard chemotherapy may be sufficient for AML patients with favorable risk profiles. However, for those with adverse-risk profile, HSCT is recommended. Nevertheless, these advancements in treatment are largely confined to younger patients and those without high risk profile, while older and less fit patients continue to exhibit a poor prognosis.[77]

Although chemotherapy remains the primary therapeutic approach today, alternative options to traditional chemotherapy have become a key focus in recent years. Since the year 2000, a growing number of targeted therapies, including both molecular and immunotherapies, have become available. Some targeted therapies can be used as single agents but the trend in leukemia treatment is shifting toward combination therapies rather than monotherapies. As a result, chemotherapy is often used in conjunction with other therapies, such as targeted drugs, immunotherapies, and HSCT, to enhance survival rates and improve patients' quality of life.

Molecular targeted therapies

By the mid-1990s, the success story of targeted therapies in leukemia began with the success of all-trans retinoic acid in *PML::RARA*-positive AML.[194,195] Shortly thereafter, the tyrosine kinase inhibitor (TKI) imatinib was approved,[196,197] revolutionizing the treatment of chronic myeloid leukemia (CML) by benefiting the entire patient population. These breakthroughs established the concept of targeted therapies in hematological malignancies. Unlike CML, which is characterized by the *BCR::ABL1* rearrangement as a universal genetic lesion,[197] the genetic complexity of acute leukemias has made the discovery of molecular targeted therapies for AML and ALL more challenging. As a result, existing therapies are often limited to a small subset of patients. For example, *PML::RARA*-positive AML, which previously had a dismal prognosis with most patients dying within weeks before the advent of all-trans retinoic acid, experienced another major breakthrough with the discovery that arsenic trioxide degrades the *PML::RARA* fusion protein.[194,195] The all-trans retinoic acid + arsenic trioxide combination therapy is now the standard of care for *PML::RARA*-positive AML, offering cure rates greater than 90%.[198] This combination therapy has now replaced chemotherapy for most *PML::RARA* patients, making it the first and only leukemia subtype curable without traditional chemotherapy.

In ALL, TKI therapy remain the only available molecular targeted therapies for both pediatric and adult patients. However, these therapies are limited to the *BCR::ABL1* positive subtype and the success is not as high as in CML,[199,200] with less durable remissions and a greater reliance on combination therapies.

At least 12 targeted therapeutic agents have been Food and Drug Administration (FDA) approved for the treatment of AML.[201] These drugs are now being incorporated into induction, consolidation, and relapsed/refractory (R/R) settings. For instance, FLT3 inhibitors (e.g., Midostaurin and Quizartinib) have been integrated into first-line therapy for *FLT3*-mutated AML, while IDH1/IDH2 inhibitors (e.g., Ivosidenib and Enasidenib) are used in the R/R setting for *IDH1*- and *IDH2*-mutated AML. Additionally, BCL-2 inhibitor combination therapy,[202] where Venetoclax is used in combination with other active agents, has proven especially effective with improved survival rates in elderly AML patients not eligible for intensive chemotherapy. The most recently approved targeted agent in AML was the menin inhibitor Revumenib,[203] which was FDA-approved in 2024 for use in the R/R setting, for pediatric and adult AML with mutated *NPM1* or *KMT2A* rearrangements. These targeted therapy approvals have primarily benefited adult AML patients, whereas pediatric AML has historically lacked effective targeted therapies.[204] Therefore, the recent approvals of Midostaurin and Revumenib, intended also for treatment of pediatric patients, have been especially welcome. However, ongoing clinical trials and research are likely to further expand the

toolbox of targeted therapies, providing a hope that safe and effective treatment options for pediatric and adult acute leukemia can become available in the future.

Immunotherapies

Contrary to traditional chemotherapy, immunotherapies aim to eradicate the leukemia cells using the immune system while sparing healthy tissue. The origins of immunotherapy can be traced back over 50 years to the introduction of HSCT,[205] where immune cells from a healthy donor were observed to help eliminate leukemia cells. The ability to evade the immune system is described as one of the hallmarks of cancer.[6–8] In this regard, leukemia cells employ various strategies to evade immune surveillance. They may, for example, downregulate antigen-presenting molecules such as MHC class I and II, alter the cytokine milieu by inducing an inflammatory state, increase immunosuppressive cell types, or entice T and NK cells into dysfunctional states.[206] Thus, immunotherapies for leukemia represent strategies to counter these evasion mechanisms. The available immunotherapies can broadly be categorized into antibody-based and cellular immunotherapies (**Figure 12**).

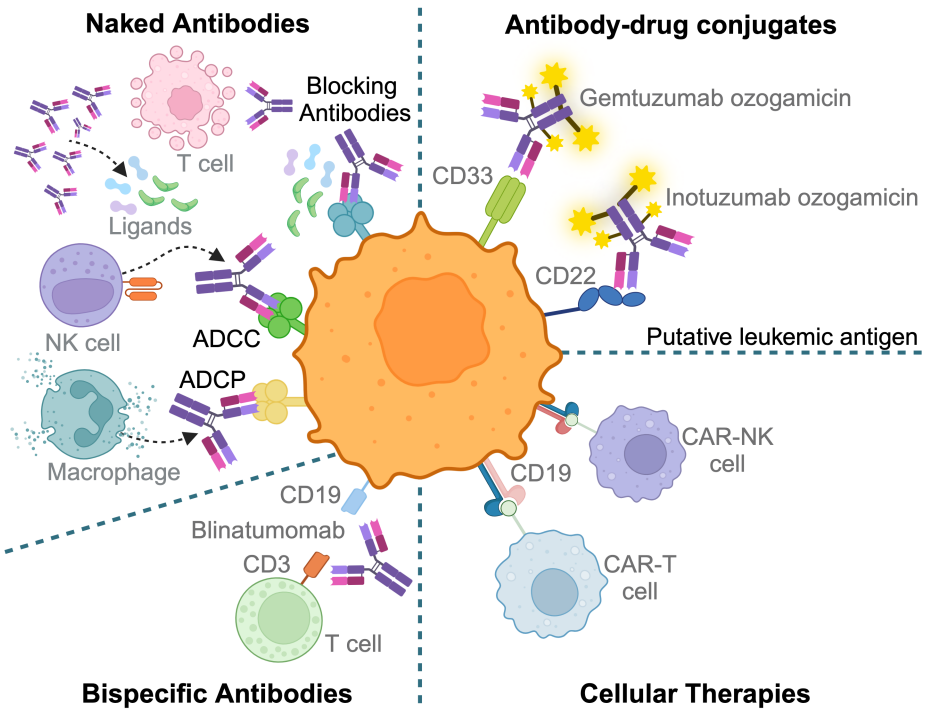

Figure 12 | Immunotherapies in acute leukemia.
Representative immunotherapeutic strategies are shown, including blocking antibodies, antibody-drug conjugates, bispecific antibodies and cellular therapies. ADCC; antibody-dependent cellular cytotoxicity, ADCP; antibody-dependent cellular phagocytosis, CAR; chimeric antigen receptor.

Antibody-based Immunotherapies

Targeting LSCs through their expression of cell surface markers absent on normal cells offers an attractive treatment strategy. Monoclonal antibodies can be used for the specific targeting of leukemic cells. Upon binding to specific cell surface markers on target cells, antibodies can inhibit or enhance essential signaling and opsonize cells for immune-mediated killing. Additionally, when targeting LSCs, antibodies can mobilize them from their bone marrow niches.[207] There are several ways of engineering antibodies to increase their therapeutic effect, and these can be broadly categorized based on the mechanism of action into antibodies engaging effector cells, blocking antibodies, bispecific antibodies, and antibody-drug conjugates, which are widely used in hematologic malignancies.

Antibodies engaging effector cells exert their effect by recruiting immune cells such as NK cells, macrophages, and neutrophils to eliminate target cells via mechanisms such as antibody-dependent cellular cytotoxicity (ADCC) and antibody-dependent cellular phagocytosis (ADCP).[208,209] In ADCC, fragment crystallizable γ receptor (FcγR)-expressing effector cells, such as NK cells, recognize and kill antibody-coated target cells, while in ADCP, macrophages and neutrophils phagocytose opsonized cells. Examples of therapeutic antibodies using these mechanisms include Rituximab (anti-CD20) and Daratumumab (anti-CD38).[210,211] Based on discoveries in our laboratory,[212,213] a Phase 1B/2A clinical trial (clinicaltrials.gov: NCT06548230) was recently initiated to evaluate ADCC-enhanced IL1RAP-targeting antibodies in patients with high-risk MDS and AML.

Blocking antibodies bind to specific receptors or ligands to block their function. For example, checkpoint inhibitors are a subclass of blocking antibodies that block inhibitory immune checkpoint proteins like PD-1 and CTLA-4,[214] thereby hindering one strategy for immune evasion by the cancer cells. This discovery, which was awarded the Nobel Prize in 2018, has proven most effective against solid cancers. However, there are no FDA approvals of checkpoint inhibitors for leukemia so far.[215]

Bispecific antibodies work by simultaneously binding two different targets, such as a cancer cell and an immune cell (e.g., T or NK cells), consequently redirecting the immune cells to attack the cancer. Bispecific T cell engagers (BiTEs) are a subclass of bispecific antibodies,[216] with Blinatumomab being the first FDA-approved BiTE therapy,[217] currently used for relapsed or refractory CD19-positive ALL.

Antibody-drug conjugates (ADCs) are antibodies linked to drugs that are utilized in cancer therapy to deliver chemotherapy directly to tumor cells, minimizing toxicity to healthy cells. For AML, Gemtuzumab ozogamicin (Anti-CD33 ADC) and Inotuzumab ozogamicin (Anti-CD22 ADC) are used for relapsed/refractory AML and ALL, respectively.[218,219]

Cell-based Immunotherapies

So far, the most successful form of adoptive cell therapy involves T cells, which utilizes their cytotoxic mechanisms by which T cells identify and eliminate tumor cells. This is accomplished through various strategies, including the genetic modification and expansion of a patient's autologous T cells (**Figure 13**).[220] One such approach is chimeric antigen receptor (CAR) T cell therapy, where the endogenous TCR is engineered with an artificial receptor composed of three parts: an intracellular signaling domain from a TCR that provides activation signals upon binding to its target, a transmembrane domain that anchors the receptor in the T cell membrane, and an extracellular antigen-recognition domain, typically derived from an antibody, enabling the recognition of tumor antigens (**Figure 13**).[220] Unlike normal T cells, which require antigen presentation via MHC, CAR-T cells, with their engineered antigen-recognition domain linked to a functional intracellular signaling domain, can recognize surface antigens directly on leukemic cells and elicit a cytotoxic immune response.

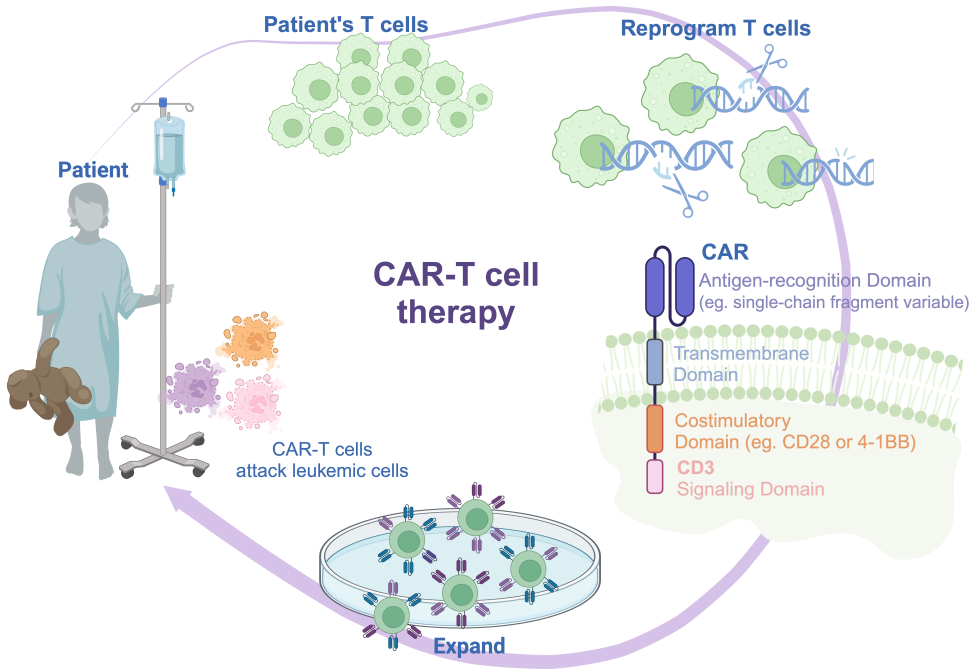

Figure 13 | CAR-T cell therapy.
A schematic illustration of CAR-T cell therapy. In this therapeutic approach, a patient's T cells are isolated, genetically modified to express CAR for better cancer recognition and killing, and then reinfused into the patient. CAR; chimeric antigen receptor.

Unlike AML, which lacks tumor-associated antigens consistently expressed across all patients and subtypes, ALL cells express CD19 as a nearly universal target, covering most ALL subtypes. As a result, CD19-directed CAR-T cell therapy has emerged as a groundbreaking benchmark in targeted therapies for ALL.[221,222] However, current FDA-approved CAR-T cell therapies for ALL are primarily reserved for relapsed or refractory cases. The complex manufacturing process, high cost, potential side effects such as cytokine release syndrome (CRS) and neurotoxicity, further limit accessibility of CAR-T cell therapy. In addition, some patients fail to respond due to CD19 antigen loss.[223] Nevertheless, as the therapy continues to evolve, its use may expand through combination treatments or earlier integration into the treatment process. Unfortunately, CAR-T cell therapies for AML remain in research stages, as progenitor and myeloid markers such as CD123, KIT, CD33, and FLT3 have not demonstrated the same strong clinical utility as CD19-targeted CAR-T cells.[224–229] This is likely due to the prolonged ablation of healthy myeloid lineages that result from eliminating cells expressing the selected markers, a challenge that must be addressed. However, applying a dual-target approach to target AML specific surface antigen combinations may help overcome this obstacle and is currently being explored for CAR-T cell therapies.[230–232]

Other emerging cellular therapies for acute leukemias that are not yet FDA-approved include TCR therapy and NK cell therapies.[233,234] TCR therapies are being explored for acute leukemias but remain in an experimental stage. Unlike CAR-T cells, TCR T cells are genetically modified with an intact TCR of choice, enabling them to recognize intracellular antigens, thereby expanding the range of potential targets.[234] Currently investigated targets include TdT for ALL and FLT3 for AML.[235,236] Engineered NK cell-based approaches have the potential to offer safer alternatives without some of the challenges associated with T cell therapies, such as CRS and neurotoxicity.[233,237–239] Allogeneic NK cells are typically derived from peripheral blood and cord blood, though induced pluripotent stem cell-derived NK cells are also emerging as an option. Additionally, unlike autologous CAR-T cells, CAR-NK cells offer off-the-shelf availability. However, CAR-NK cells remain in clinical development, and their full therapeutic potential will depend on further research and validation.

Methodological approaches and single-cell technologies

Single-cell transcriptomic technologies

Single-cell transcriptomics has become one of the most widely used techniques in recent years, driven by rapid technological advancements. Specifically, improvements in library preparation, high-throughput sequencing platforms, and bioinformatics have made scRNA-seq methods more scalable, efficient, and accessible.

Technical platforms

A wide range of technical platforms is available, each offering unique advantages depending on the specific application, such as Smart-seq2,[240] Smart-seq3,[241] Single-Cell Tagged Reverse Transcription sequencing (STRT-seq),[242] Cytometry Sequencing (Cyto-seq),[243] Indexing Droplets sequencing (inDrop),[244] Droplet Sequencing (Drop-seq),[245] 10x Genomics,[246] Cell Expression by Linear amplification Sequencing 2 (CEL-seq2),[247] Quartz Sequencing (Quartz-seq),[248] Massively Parallel RNA Single-cell sequencing (MARS-seq),[249] Sequential Well sequencing (Seq-Well),[250] multiplexed dual annealing sequencing (RamDA-seq),[251] and full-length alternative transcriptome Sequencing (FLASH-seq).[252]

ScRNA-seq methods can generally be divided into two main types based on whether they generate full-length or partial RNA molecules, with a trade-off between transcriptome quality and the number of cells sequenced. Full-length methods capture the entire RNA transcript, from the 5' end to the 3' end, whereas partial-length methods capture only a portion of the RNA transcript, typically targeting specific regions like the 3' end or the 5' end. Full-length methods, such as Smart-seq2 and Smart-seq3, can capture up to 65-75% of expressed genes in a cell.[240,241] However, some advanced full-length methods, like RamDA-seq[251] and FLASH-seq,[252] can achieve up to 80-85% coverage of the transcriptome. Full transcriptome profiling allows for the study of isoform diversity, alternative splicing, and genetic changes. However, full-length transcriptional profiling techniques typically capture a smaller number of single cells (1-1,500 cells).[240,241] In contrast to full-length

techniques, 3' or 5' end-biased methods like 10x Genomics or Drop-seq typically have lower accuracy and detect about 10-20% of the transcriptome.[245,246] However, partial-length RNA techniques allow the isolation of around 100-100,000 single cells. Moreover, 10x Genomics continues to enhance its high-throughput capabilities by optimizing single-cell isolation techniques like GEM-X[253] and refining multiplexing methods such as cell hashing[254] and CellPlex[255]. In addition to differences in transcriptome quality and the number of cells sequenced, partial-length methods are currently a more cost-effective alternative.[256]

In summary, full-length RNA-seq methods provide a deeper understanding at the transcriptional level by offering the highest sensitivity and the most comprehensive coverage of genes per cell. However, partial-length techniques enable large-scale gene expression studies with higher throughput and are more cost-effective alternatives.

Experimental designs

Methodologically, single-cell sequencing involves several steps, including the isolation of individual cells by various techniques, library preparation, high-throughput sequencing, and subsequent bioinformatic analyses. The process of single-cell isolation, *i.e.* capturing individual cells, varies between platforms and is typically performed through dilution into PCR plates, chips containing thousands of nanowells, or microfluidic devices.[240–252]

In this thesis, the technique used for transcriptomic profiling of single cells was the 10x Genomics Chromium platform, which utilizes a microfluidic device (**Figure 14**). Single-cell isolation is achieved by encapsulating cells in nanoliter-scale microdroplets, referred to as gel beads in emulsion (GEMs).[257] Cells, along with gel beads and reagents, are loaded onto a chip and processed in the Chromium Single Cell Controller (10x Genomics, USA). The gel beads play a critical role by providing oligonucleotides required for cell identification. Once the cells are encapsulated in GEMs, cell lysis occurs, releasing mRNA molecules. These mRNA molecules are then tagged with a unique cell barcode (linking the transcript to its cell of origin) and unique molecular indices (UMIs; distinguishing individual RNA molecules) during the reverse transcriptase step. This constitutes the initial reaction of library preparation. Once this step is completed, the GEMs are dissolved, and the resulting complementary DNA (cDNA) is released. Library preparation proceeds with several steps, including cDNA amplification, fragmentation, end-repair, A-tailing, and index adaptor ligation. Finally, the prepared libraries are then subjected to high-throughput sequencing, typically using Illumina sequencing technology (**Figure 14**).

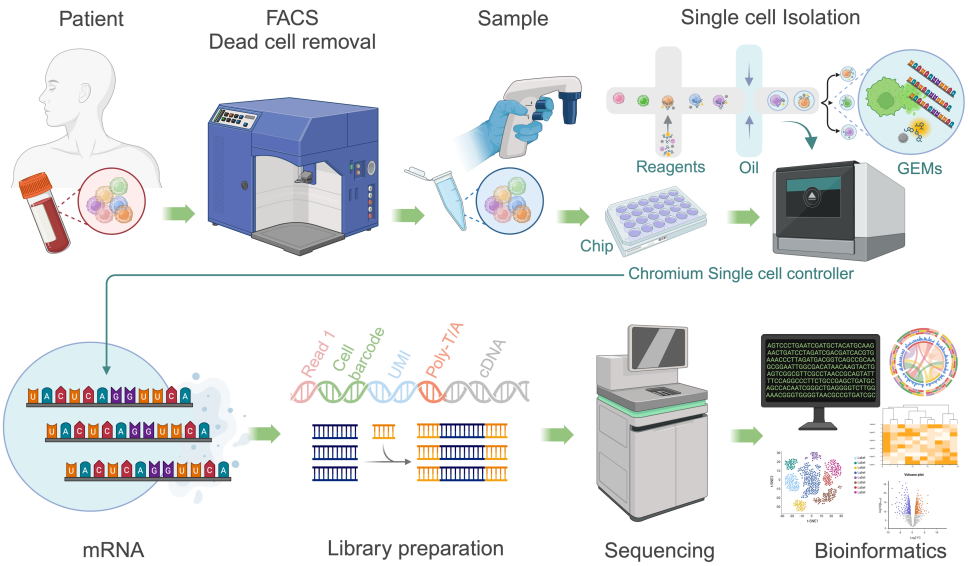

Figure 14 | Workflow of single-cell RNA-sequencing.
A schematic illustration of the single-cell RNA sequencing workflow. The process begins with the collection of patient samples, followed by an optional FACS step used for sorting specific cells or removing dead cells to improve sample quality. Next, single-cell isolation is performed, followed by library preparation before sequencing. Finally, the data undergoes bioinformatic analysis. FACS; Fluorescence-Activated Cell Sorting, GEM; Gel Bead-in-Emulsion, and mRNA; messenger RNA.

Single-cell targeted transcriptomics

In **Article III**, a novel mutation tracking method, scRNAmut-seq, was developed to identify which cells constituted a leukemic clone in each sample. ScRNAmut-seq was performed using leftover full-length cDNA from the original scRNA-seq libraries (**Figure 15**). Mutation-specific PCR primers were designed to target mutations identified through whole exome sequencing for each patient. Evaluation was based on transcript coverage from both scRNA-seq and bulk RNA-seq data, as well as the mutation's distance from the gene's 3′-end. Prior to performing scRNAmut-seq, the full-length cDNA material was repaired, followed by targeted PCR amplification of the selected mutations and a second PCR using primers that added sequencing adapters.

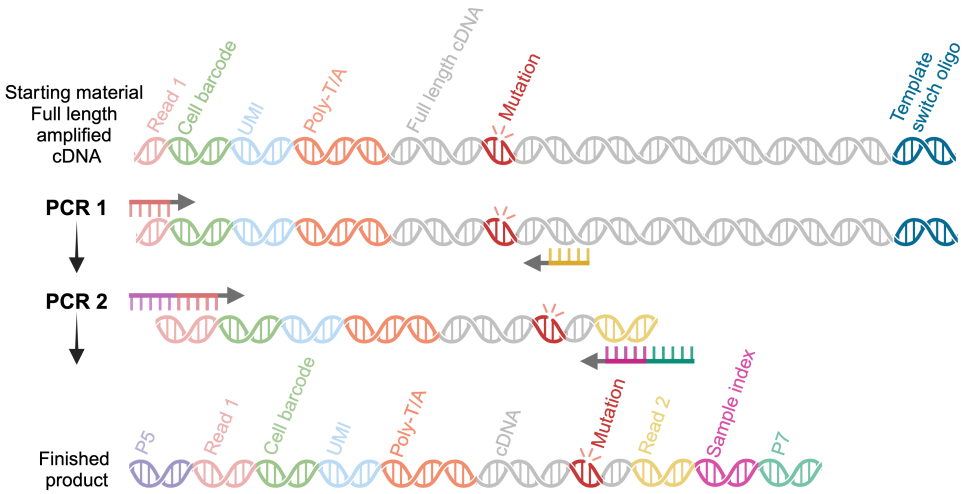

Figure 15 | Overview of mutation calling PCR strategy.
Single-cell mutation calling, developed in Article III, was performed on single-cell RNA sequencing material using a two-step PCR amplification protocol. cDNA; complementary DNA, and UMI; unique molecular index.

Single-cell V(D)J sequencing

Other transcriptomic single-cell modalities include the characterization of the repertoire of BCRs and TCRs, which are crucial components of adaptive immunity. This is achieved by amplifying the V(D)J regions of the antigen receptors (BCR or TCR), a process known as V(D)J-sequencing, BCR-sequencing, or TCR-sequencing. These regions are located within the first ~500 nucleotides of the 5′ end of the transcripts.[258] Thus, most technologies that capture BCR and TCR sequences are performed in combination with 5′ scRNA-seq, amplifying both the V(D)J sequences and the transcripts. One such method is commercially available through 10x Genomics' single-cell immune profiling technology.[259] This technology is utilized in **Article I**.

In short-read scRNA-seq, after the first step of reverse-transcribing mRNA into cDNA, full-length V(D)J segments are enriched from the amplified cDNA via PCR amplification using primers specific to either the BCR or TCR constant regions of the immune receptor genes. Amplifying the V(D)J regions is necessary because the BCR/TCR regions are present at low levels in individual cells and require robust amplification to be effectively sequenced. These data provide an in-depth understanding of the immune repertoire by examining the diversity of the immune system, accounting for receptor types generated by V(D)J recombination, the clonality of specific immune responses by tracking the expansion of receptor sequences, and BCR/TCR pairing and interaction with antigens.

Single-cell transcriptomic data analysis

Today, data processing follows a well-established standard workflow, with initial steps resembling certain aspects of traditional bulk data analysis. These data are typically processed through specialized bioinformatics pipelines.

Raw data processing

Following short-read sequencing on Illumina instruments, data processing begins with the conversion of sequencing output, where BCL files are transformed into FASTQ files. This is followed by demultiplexing, where sequencing reads are assigned to their respective cells. Next, the sequencing reads are aligned to a reference genome, mapping them to genes or transcripts. Lastly, quantification is performed, where the number of reads corresponding to each gene is counted to quantify gene expression for each individual cell. The result is a gene-by-cell count matrix, also referred to as a single-cell count matrix, where rows represent genes, columns represent cells, and the values indicate the number of detected transcripts. A higher number of detected transcripts reflects higher expression of the specific gene. These steps form the preprocessing phase, which is then followed by downstream analysis.[260,261]

Downstream analysis

Recent advancements in experimental technology have driven large-scale innovation in computational methods, resulting in the availability of over 1,400 tools for scRNA-seq data analysis,[262] with popular frameworks for downstream analysis including Bioconductor[263] and Seurat[264] for the R environment, as well as Scanpy[265] for Python.

Downstream analysis of the single-cell count matrix commonly involves quality control steps to ensure data reliability (**Figure 16**). During this stage, low-quality genes, such as those expressed in only a few cells, are excluded. Similarly, low-quality cells are filtered out; these include cells with low read counts or a higher percentage of mitochondrial reads, which often indicate cell death. Quality control can also include the removal of doublets, achieved by identifying and excluding droplets that contained more than one cell.[260] Following quality control, a normalization step is performed to account for differences in sequencing depth across cells and to stabilize variance across genes. This is then followed by dimensionality reduction, a process involving several steps. Here, the first step involves the selection of the most informative genes, referred to as highly variable gene selection. Next, principal component analysis (PCA) is performed to reduce the dimensionality of the data. This is commonly followed by nonlinear methods

such as t-distributed stochastic neighbor embedding (t-SNE),[266] uniform manifold approximation and projection (UMAP),[267] or k-nearest neighbors (KNN) force layout graphs,[268] which are used to visualize the data in two or three dimensions. On the dimensionally reduced data, clustering methods are applied to group cells based on their similarities in gene expression patterns. Techniques such as graph-based clustering, hierarchical clustering, and k-means clustering are often used to uncover distinct cell states or types represented in the data. Alternatively, instead of mathematical clustering, cell type annotation can be employed. This involves assigning groups of cells to specific cell types based on gene expression or reference datasets.[260]

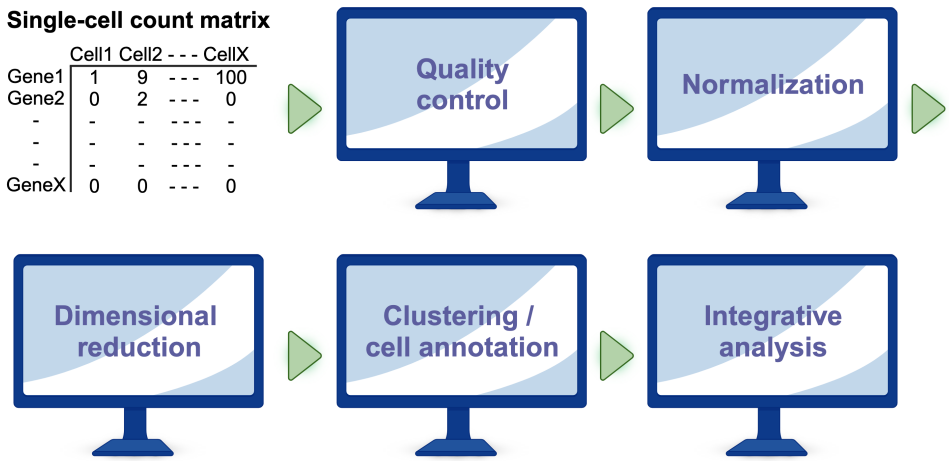

Figure 16 | Overview of single-cell RNA sequencing bioinformatic analysis.
After raw sequencing, the resulting single-cell count matrix is further processed with downstream analysis. The downstream analysis of the single-cell count matrix commonly involves five key steps: quality control, normalization, dimensionality reduction, mathematical clustering and/or cell type annotation, followed by various specific analyses tailored to the research question, referred to as integrative analysis.

After these analyses, depending on the desired application, various types of computational approaches can be performed (integrative analysis; **Figure 16**). For example, differential expression analysis can be applied to identify genes that differentiate between cell states, while pathway enrichment analysis highlights biological pathways that are overrepresented under specific cell conditions. Additionally, trajectory analysis, also known as pseudotime analysis, is used to infer developmental progression or differentiation patterns within the data.[260] For each step mentioned in this description, there are countless software tools and repositories available,[262] which are continuously being developed and improved to meet the growing demands of single-cell data analysis.

Data challenges and future perspectives

Computational analysis of single-cell transcriptomic data is known to be challenging, as the data are usually noisier compared to standard bulk sequencing. Technical noise in single-cell sequencing arises primarily due to the low concentration of starting material obtained from an individual cell.[256,261] Dropout events, where transcript molecules fail to be captured from individual cells, contribute significantly to this noise. Additionally, the low RNA concentration necessitates extensive amplification of genetic material, which can introduce polymerase-induced errors and allelic dropout. This technical noise results in data sparsity, as scRNA-seq datasets often contain near-zero or zero counts, complicating downstream analyses. Furthermore, technical variability and batch effects present additional challenges. Batch effects, caused by differences in library preparation, sequencing depth, and platform-specific biases, introduce technical variability that hinders comparisons across studies. Even with similar study designs, small technical differences, such as sampling individuals at different time points, can lead to batch effects. Proper handling of these batch effects is crucial for ensuring improved comparability and accurate biological interpretations.[260] Partial-length sequencing methods are associated with additional challenges. In addition to limited transcript coverage and low detection of isoforms and non-coding RNAs, quantification biases also pose a challenge. These biases include favoring gene ends or shorter transcripts, reduced sensitivity for lowly expressed genes, and an overrepresentation of highly expressed genes.[256]

An additional challenge is the high dimensionality of the dataset. Even though some single-cell data suffer from sparsity and limited transcript coverage, they still hold immense amounts of information. With tens of thousands of genes measured per cell and datasets containing millions of cells, scRNA-seq data result in high-dimensional matrices (e.g., millions of cells × 30,000 genes).[260] Due to rapid technological advancements, increasingly larger datasets are being produced,[21] further challenging computational analysis and visualization. Although a growing number of computational tools and algorithms are being developed to facilitate the analysis of single-cell sequencing data,[262] the field of single-cell sequencing is still relatively young, rapidly evolving, and becoming more cost-effective. This calls for the development of improved computational tools. In addition, there will be a continuous need for substantial computational resources and efficient algorithms to handle the complexity and size of large-scale single-cell datasets.

Other single-cell modalities

Single-cell proteomic technologies

A number of single-cell proteomics techniques have been developed to understand the proteomic makeup of individual cells. These techniques are highly dependent on protein concentration, which is very low in single cells. Unlike DNA or RNA, proteins cannot be amplified, presenting unique challenges for single-cell proteomics. Single-cell proteomics approaches mainly include antibody-based methods, mass spectrometry based methods, and imaging based methods capable of detecting proteins while providing spatial context.[269]

FACS is one of the first antibody-based methods and a widely used single-cell technique in immunology, cancer research, and stem cell studies.[270] This method labels cells with fluorescent antibodies, enabling laser detection of fluorescence and light scattering to analyze size and granularity for precise sorting. All antibody-based methods typically target a limited number of predefined proteins. Other antibody-based methods have gained popularity due to their ability to integrate protein and transcriptomic data. These methods use oligo-conjugated antibodies to target specific cell surface proteins, which are sequenced alongside transcriptomic data. Cellular indexing of transcriptomes and Epitopes by Sequencing (CITE-seq) was the first antibody-based method developed to merge transcriptome and proteome analysis, marking a milestone in the field.[271] Subsequently, 10x Genomics adapted CITE-seq for integration with their transcriptomic workflows (**Figure 17**), known as Antibody-Derived Tag Sequencing (ADT-seq).[272] Initially limited to a few markers, these methods have expanded, with companies like BioLegend now offering antibody cocktails that target over 100 proteins.[273] In **Article I** and **Article III**, the ADT-seq technology was used to investigate the protein expression of approximately 20 specific cell surface markers.

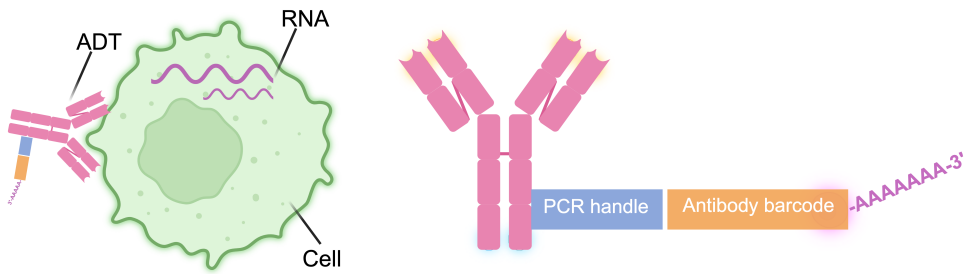

Figure 17 | Antibody-Derived Tag Sequencing.
ADTs are antibody clones labeled with unique barcodes, which are attached to a capture sequences (eg. poly(A) sequence) and a PCR handle. These barcodes are amplified during subsequent library preparation. The antibodies bind to specific cell surface proteins, and the sequenced ADT counts represent the expression levels of these specific proteins. ADT; Antibody-Derived Tag.

Mass spectrometry is the gold standard for large-scale proteomic analysis at the bulk level. In the context of single-cell proteomics, single-cell mass spectrometry has shown success, particularly with the method Cytometry by Time-of-Flight (CyTOF), which combines MS with flow cytometry.[274] Recently, single-cell mass spectrometry as a standalone tool has also demonstrated promising results, with increased protein throughput achieved through automated sample preparation methods and highly sensitive, multiplexed instruments, enabling the quantification of 1,000 to 1,500 proteins per cell.[275–277] Unlike antibody-based methods such as CyTOF, ADT-seq, and CITE-seq, single-cell mass spectrometry does not require consideration of antibody specificity or cross-reactivity, which can become problematic as the number of targeted proteins increases. Single-cell Imaging Mass Cytometry (IMC) combines high-plex proteomic analysis with spatially resolved tissue imaging. However, this technique operates at a subcellular level. While there is ongoing work to optimize experimental designs and improve sensitivity at the single-cell level, IMC holds significant potential for advancing spatial proteomics and tissue analysis in the future.[278]

Single-cell DNA technologies

Today, transcriptomic methods remain the primary single-cell tools for defining cell states due to scalability, accessibility and ability to capture the dynamic nature of gene expression. However, single-cell genomic analysis, specifically single-cell DNA sequencing, is indispensable for understanding genetic mutations, structural variations, and the fundamental framework that regulates RNA expression. Single-cell DNA sequencing includes whole-genome amplification to amplify DNA from a single cell, followed by WGS to uncover genetic information.[256,279–281] Several whole-genome amplification methods have been developed, with early examples including PCR-based methods such as degenerate oligonucleotide-primed PCR (DOP-PCR),[282] isothermal multiple displacement amplification (MDA),[279] and multiple annealing and looping-based amplification cycles (MALBAC).[280] Single-cell DNA sequencing has traditionally been limited by low cellular throughput. However, in recent years, modest increases in throughput have been achieved by optimization of protocols, employing microfluidic devices, and single-cell, combinatorial indexed sequencing and digital library preparation.[283–285]

Single-cell epigenetic technologies

The field of single-cell epigenomics is rapidly evolving, driven by continuous advancements in technology. Several single-cell epigenetic methods are now available for studying various aspects of the epigenome, such as DNA methylation, histone modifications, and chromatin accessibility. For example, single-cell reduced

representation bisulfite sequencing (scRRBS-seq), single-cell Assay for Transposase-Accessible Chromatin using sequencing (scATAC-seq), and Single-cell Cleavage Under Targets and Tagmentation sequencing (CUT&Tag-seq) are a few of these applications.[286–289] ScRRBS-seq isolates DNA from individual cells, enriches for regions of interest such as promoters (CpG-rich regions) through restriction enzyme digestion, and treats the DNA with bisulfite to distinguish between methylated and unmethylated positions.[287] ScATAC-seq profiles chromatin accessibility in individual cells by identifying open chromatin regions, which are areas where DNA is accessible to transcription factors and other regulatory proteins.[286] The scATAC-seq workflow involves nuclear isolation, followed by the addition of a hyperactive mutant Tn5 transposase (**Figure 18**). This enzyme recognizes open chromatin regions, cleaves the DNA, and tags the resulting fragments. Next, individual nuclei are captured using microfluidic technology, followed by library preparation and sequencing. The number of sequencing reads for a region reflects the degree of chromatin accessibility in individual cells. This process generates a matrix similar to a single-cell count matrix, with nuclei instead of cells and peaks instead of genes. Notably, this type of data not only identifies accessible DNA but also maps regions of transcription factor binding sites. In this thesis, scATAC-seq was employed to examine the epigenetic state of childhood BCP-ALL (**Article I**).

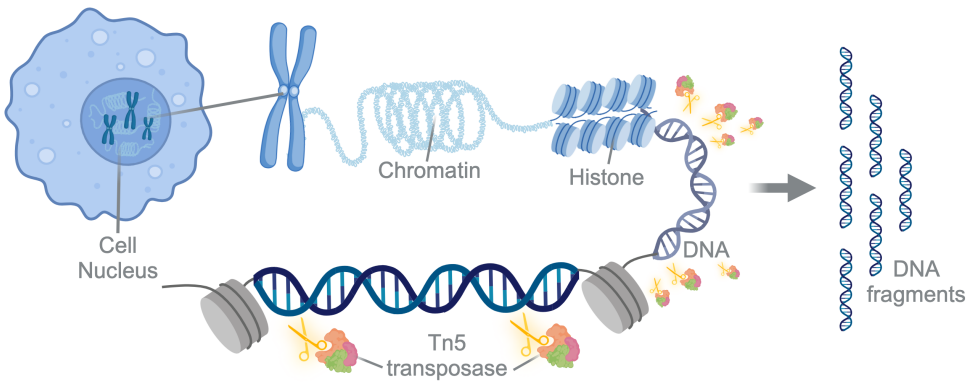

Figure 18 | Assay for Transposase-Accessible Chromatin using sequencing.
To investigate chromatin state dynamics at the single-cell level, scATAC-seq measures genome-wide chromatin accessibility by isolating individual nuclei. The open chromatin regions are tagged through Tn5 transposition, generating DNA fragments that are then sequenced.

ScCUT&Tag-seq enables the analysis of histone modifications and has recently gained popularity for its ability to examine tens of thousands of single cells simultaneously.[288,289] This method is conceptually similar to scATAC-seq; however, scCUT&Tag-seq employs an antibody-directed approach, where a specific antibody binds to a histone modification or protein of interest. As a result, scCUT&Tag is

more targeted and specific for studying histone modifications or protein-DNA interactions, whereas scATAC-seq provides a broader overview of chromatin accessibility across the genome.

Mutational detection at a single-cell level

In cancer research, understanding the initiation, development, and maintenance of cancer pathogenesis requires insights into the mutational makeup of cells. Techniques such as single-cell DNA-seq, RNA-seq, and ATAC-seq, as previously mentioned, provide valuable tools for studying these mutational patterns. However, much remains to be explored in this area. While full-length methods are not yet cost-effective,[256] partial-length and targeted single-cell sequencing have emerged as popular alternatives for studying mutational patterns at the individual cell level. The Mission Bio Tapestri platform is a leading partial-length single-cell multi-omics platform based on a microfluidic droplet-based system, enabling the identification of genetic abnormalities, including single nucleotide variations, insertions and deletions, copy number variations, and translocations.[290] In addition to genetic analysis, this platform simultaneously performs protein analysis using oligo-tagged antibodies.[291] Moreover, targeted single-cell sequencing methods, such as TARGET-seq,[292] have been developed. These approaches, while sacrificing genome-wide mutational patterns, are limited to predefined regions using primers designed to capture specific areas of interest. This enables deeper sequencing for those regions, increasing sensitivity and precision. While these methods focus on clinically relevant regions or specific mutations in well-defined diseases, they also reduce computational resource demands and enhance cost-efficiency.

Multiomic single-cell modalities

Single-cell sequencing technologies, when used as standalone modalities, have already revolutionized our understanding of the cellular landscape of acute leukemia. However, combining multiple modalities within the same cell offers even greater potential by integrating and mapping interconnected networks.[21] This approach is known as single-cell multiomics. For instance, combining V(D)J sequencing, proteomics via ADT sequencing, and transcriptome analysis within the same cell exemplifies this methodology, as applied in **Article I**. Other multiomics modalities enable the integration of transcriptomic and epigenetic data in the same cell,[22] as well as genetic and transcriptomic information.[293–295] These representative technologies are just a few examples of the many multi-modality approaches currently in use.

A future perspective for single-cell sequencing

Several technologies are currently being optimized, along with a few emerging ones that, while not yet widely used, have the potential to provide valuable insights in the future. One example is single-cell CRISPR-based tools,[296–298] which leverage the CRISPR-Cas9 system to manipulate genes at the individual cell level, enhancing our understanding of gene function. A popular platform, Perturb-seq,[297] combines CRISPR perturbations with scRNA-seq to assess gene expression changes following edits. However, challenges remain, particularly in delivery efficiency and off-target effects. Another example is spatially-resolved single-cell techniques, which combine genetic data with spatial context to preserve a cell's location within tissue.[278] Moreover, with advances in single-cell multiomic technologies, the combination of transcriptomics, proteomics, epigenomics, and metabolomics with spatial and temporal context in individual cells holds great promise for mapping the full spectrum of molecular changes that drive biological systems.

The question remains whether single-cell sequencing will become the gold standard in the long term and fully replace bulk methods. While bulk sequencing remains far cheaper, which is particularly desirable for large-scale studies, it is still practical for routine clinical diagnostics, population studies, and longitudinal monitoring. Additionally, bulk sequencing excels in detecting low-frequency mutations by aggregating signals across all cells, which improves sensitivity in some cases. For the foreseeable future, both approaches will likely coexist and complement one another. Although single-cell methods are likely to become the backbone of research and precision medicine, particularly as costs decrease and scalability improves, both approaches will likely coexist for the foreseeable future, ultimately advancing cancer diagnostics, treatment approaches, and improved patient outcomes.

The present study

Background

Acute leukemia is a heterogeneous group of malignancies driven by genetic alterations that disrupt normal hematopoiesis, resulting in an accumulation of immature cells in the bone marrow and peripheral blood. Traditionally classified into AML and ALL based on their lineage, these diseases arise from a diverse set of genetic alterations arising in primitive cells, altering their self-renewal, growth, differentiation and survival properties.

While genomic and transcriptomic profiling has provided important insights into leukemia classification and risk stratification, it has become increasingly clear that the disease is not solely dictated by mutational status. Instead, leukemia exists within a dynamic cellular ecosystem in which both malignant and normal hematopoietic cells interact, shaping disease evolution and treatment response.

Despite advances in targeted therapies, such as the development of TKIs, FLT3, IDH1/2 and BCL2-inhibitors, the backbone of acute leukemia treatment has remained largely dependent on intensive chemotherapy and HSCT. A major challenge in AML and ALL therapy is the persistence of LSC, a subpopulation of malignant cells capable of self-renewal and sustaining the leukemic clone. These LSCs often evade current therapies and provide a reservoir for disease relapse. Therefore, identifying LSC-specific vulnerabilities and understanding how cellular heterogeneity contributes to disease initiation and progression is essential for improving patient outcomes.

Recent advances in single-cell sequencing technologies have made it possible to dissect acute leukemia at a level of granularity that was previously unattainable. Advances in single-cell technologies, coupled with advanced bioinformatic analysis, now allows the identification of distinct subpopulations of normal and leukemic cells, detailed characterization of their molecular characteristics, maturation patterns, and the discovery of new therapeutic vulnerabilities.

Aims of the study

The overall aim of this thesis has been to study the cellular and molecular characteristics of acute leukemia using single-cell sequencing technologies to improve disease understanding, refine diagnostic approaches, and identify novel therapeutic targets. Specifically, the aims of this thesis were to:

1) Examine the genetic, epigenetic, transcriptomic and cell surface characteristics of childhood BCP-ALLs using multimodal single-cell sequencing (**Article I**).

2) Delineate the cellular state space of defined subtypes of adult AML using scRNA-seq and develop methods allowing broad mutational calling within expressed genes (**Article II** and **III**).

3) Investigate potential age-related differences and comprehensively characterize the cellular maturation patterns, immune cell landscape, and leukemic stem cell dynamics in adult and pediatric AML using single-cell analysis. (**Article IV**)

Experimental outline of the study

A central method used in the present study has been a diverse set of single-cell sequencing technologies, enabling the isolation and analysis of single cells from bulk samples. These techniques form the basis of the screening strategies utilized in this thesis and have been applied to investigate a wide range of acute leukemias in both children and adults. The experimental design for each article in the present study is summarized in **Figure A**.

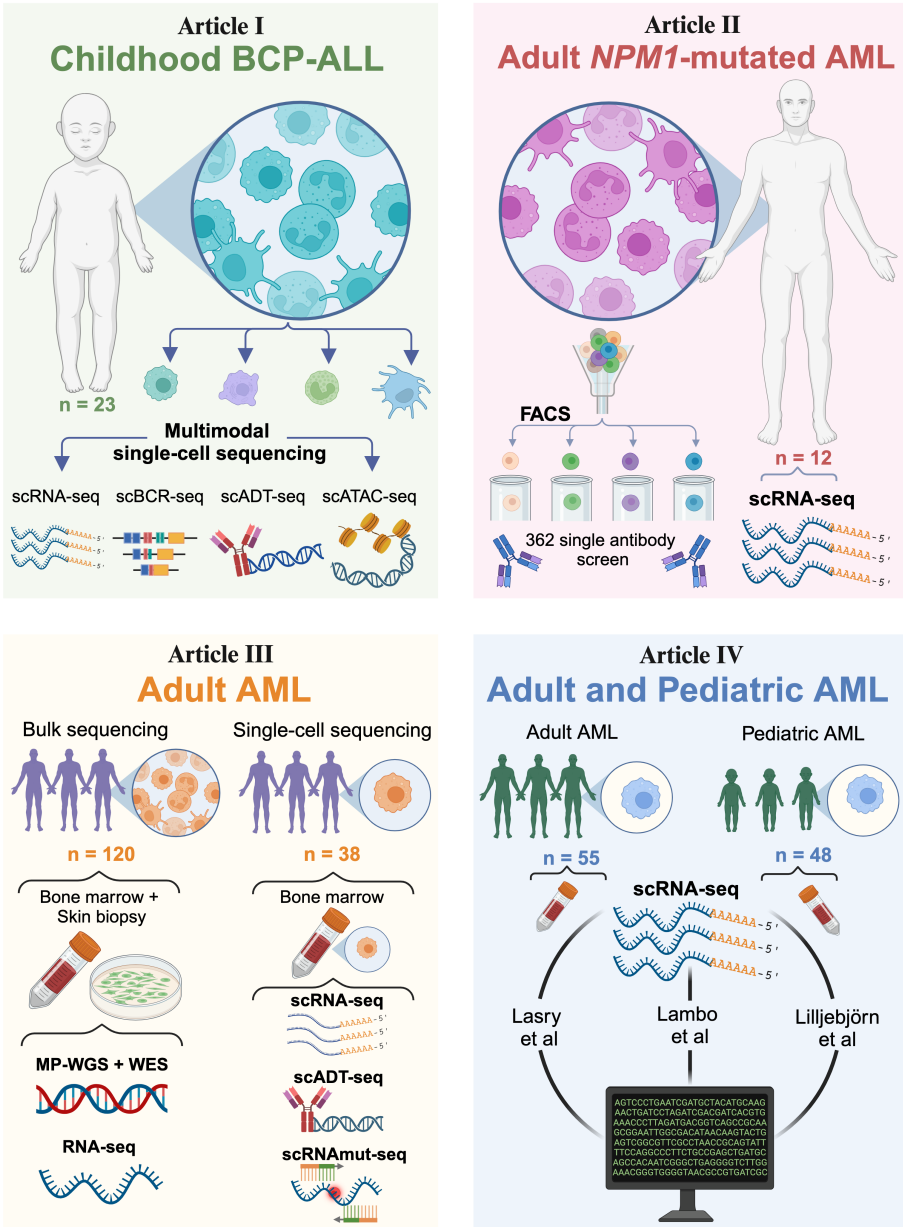

Figure A | Overview of the experimental design in the present study.
Schematic illustrations of the experimental design are presented for Article I (upper left), Article II (upper right), Article III (lower left), and Article IV (lower right). This summary provides details on age groups, the number of patients, and the techniques used to analyze the cellular and molecular characteristics within each patient sample. scRNA-seq; single-cell RNA sequencing, scBCR-seq; single-cell B cell receptor sequencing, scADT-seq; single-cell antibody-derived tag sequencing, scATAC-seq; single-cell assay for transposase-accessible chromatin sequencing, FACS; fluorescence-activated cell sorting, MP-WGS; mate-pair whole-genome sequencing, WES; whole-exome sequencing, and scRNAmut-seq; single-cell RNA mutation sequencing.

Summary of articles

Article I

Single-cell genomics details the maturation block in BCP-ALL and identifies therapeutic vulnerabilities in *DUX4*-r cases

To detail the maturation patterns of the malignant cells in BCP-ALL, we used a multimodal single-cell sequencing approach including transcriptional profiling through scRNA-seq and scBCR-seq, epigenetic profiling via scATAC-seq, and immunophenotyping using scADT-seq. In total, 23 childhood BCP-ALL cases were included representing the following genetic subtypes: *BCR::ABL1*-positive, *ETV6::RUNX1*-positive, HeH, and the recently discovered *DUX4*-r ALL subtype. ScRNA- and scATAC-seq data revealed expanded aberrant cell populations in each patient sample, which were confirmed as the blast cell populations of each case through immunophenotyping.

Cell maturation patterns were elucidated by developing new algorithms, enabling projection of leukemic cells onto the normal B cell differentiation axis and by studying BCR-chain status. In addition, a scoring algorithm was developed to highlight gene expression differences between the leukemic cells and corresponding normal B cells. This combined approach revealed diversity in maturation patterns among different BCP-ALL subtypes. *BCR::ABL1*-positive, *ETV6::RUNX1*-positive, and HeH ALL primarily resembled normal pro-B cells, as indicated by a distinct peak along the trajectory's timeline corresponding to the pro-B cell developmental stage (**Figure AI-1**). Although a distinct pro-B cell profile was observed, subtype-specific patterns were also identified. In *BCR::ABL1*-positive ALL, cells with an identified immature profile were distributed across CLP and pro-B cell stages. HeH ALL displayed a broader profile, extending from pro-B cells to immature B cells. However, the majority of *DUX4*-r ALL cases uniquely exhibited a high number of cells with transcriptional signatures resembling more mature B cells, visualized by a distinct second peak along the timeline (**Figure AI-1**).

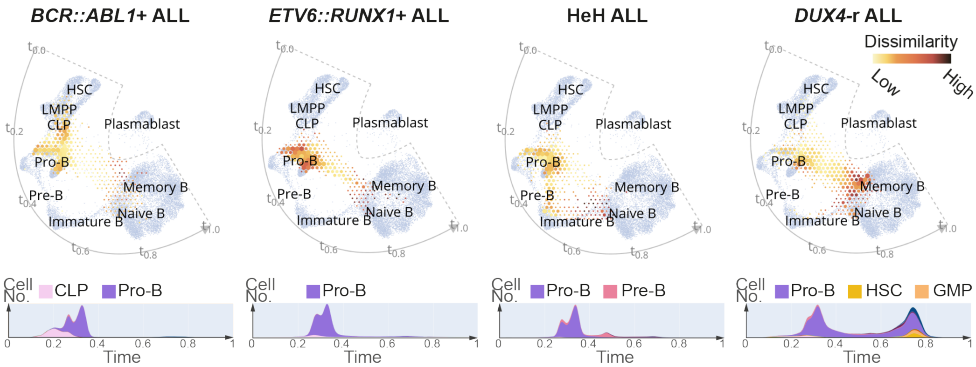

Figure AI-1 | Projection of blast cells onto the normal B cell maturation trajectory.
Single-cell projections of blast cells from ALL patients onto the normal B cell trajectories. Gene expression dissimilarity relative to the corresponding cells in the normal B cell trajectory is represented by a color gradient, with red indicating the highest level of dissimilarity. Cell density plots below the projection illustrate the number of cells along the trajectory and are color-coded by cell type.

Analyzing the co-occurrence of expressed BCR chains allowed us to infer the clonal structure of leukemia. In each case, the leukemia was generally dominated by a major clone that expressed at least one rearranged *IGH* chain. Interestingly, the maturation patterns in BCP-ALL appeared to be related to the cell of origin. Leukemias with a pro-B cell phenotype were derived from cells with *IGH* rearrangements, as seen in the majority of *ETV6::RUNX1*-positive and HeH ALL cases. The clonal trees inferred for *BCR::ABL1*-positive ALL cases indicated that the cell of origin of these leukemias lacked *IGH* rearrangements, in contrast to the other cases studied, suggesting a cell of origin before B cell commitment (**Figure AI-2**).

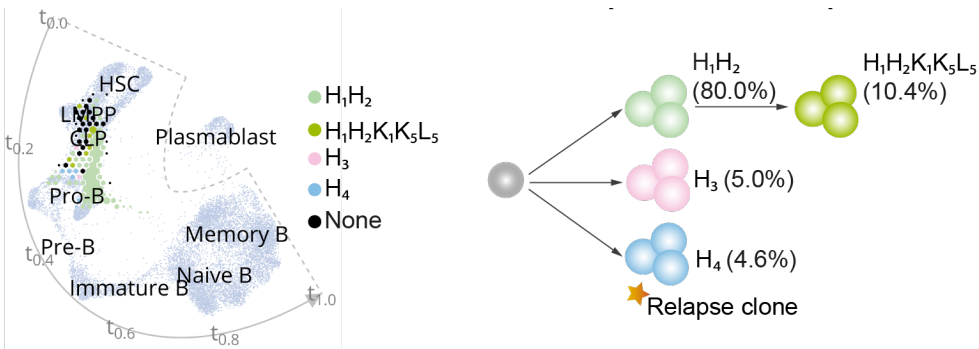

Figure AI-2 | Clonal tree of one *BCR::ABL1*-positive ALL case.
Single-cell projections of an individual *BCR::ABL1*-positive ALL case onto the normal B cell trajectory with cells color-coded according to expressed clonal IG rearrangements (left). Clonal trees were constructed based on the co-occurrence and relative frequency of expressed IG rearrangements (clone size indicated within parentheses). The number of rearranged IGH alleles and their co-occurrence suggest that the cell of origin is an immature cell without a preexisting rearranged *IGH* allele. *The minor clone H_4 (4.6%) survived treatment and became the dominant clone at relapse. H; Heavy chain, and L; Light chain

Leukemias containing blast cells with a more mature appearance were likely derived from cells with both *IGH* and *IGK/L* rearrangements, as observed in approximately half of the *DUX4*-r ALL cases. In addition, four *DUX4*-r ALL cases exhibited a productive chain, i.e., an actual functional chain, suggesting a more advanced B cell developmental stage than other subtypes. Taken together, *BCR::ABL1*-positive, *ETV6::RUNX1*-positive, and HeH ALL cases preferentially displayed gene expression profiles and a BCR chain status consistent with the pro-B cell stage. In contrast, *DUX4*-r ALL exhibited substantial heterogeneity in both maturation patterns and BCR status (**Figure AI-3**).

With the particular focus on *DUX4*-r ALL, profiling of non-leukemic cells unexpectedly identified a characteristic $CD4^+$ T cell population that displayed transcriptional signs of dysfunction. However, these cells did not exhibit DUX4 activity, suggesting that they were altered by external factors rather than originating from *DUX4*-r cells (**Figure AI-3**).

Elucidating the gene expression signatures of the leukemic *DUX4*-r blasts revealed signs of multilineage priming toward non-hematopoietic, myeloid, and T cell lineages. An additional feature of multilineage priming was highlighted, where greater cellular heterogeneity correlated with the recently described *DUX4*-a/*DUX4*-b subdivision.[100] *DUX4*-b cases were found to particularly express *CEBPA* and *FLT3*, which have previously been described as overexpressed in patients who undergo a monocytic switch during treatment. Notably, we identified one *DUX4*-b case with monocytic cells expressing a clonal BCR chain, consistent with a B cell-like leukemic cell undergoing a monocytic switch. To investigate the potential of *DUX4*-r ALL to undergo a monocytic switch, enforced expression of *CEBPA* in the *DUX4*-rearranged cell line NALM6 was shown to be sufficient to prime the cells toward a monocytic immunophenotype, as revealed by upregulation of the monocytic markers CD14, CD33, and CD371, while concurrently downregulating the B cell marker CD19. This multilineage priming suggests that *DUX4*-r blast cells, which express the embryonic transcription factor DUX4, activate pathways involved in non-lymphoid lineages, potentially deregulating the programs involved in lymphoid development. However, *DUX4* expression in BCP-ALL is typically regulated by the *IGH* locus, active only in B cells, which may limit the plasticity of *DUX4*-r blast cells (**Figure AI-3**).

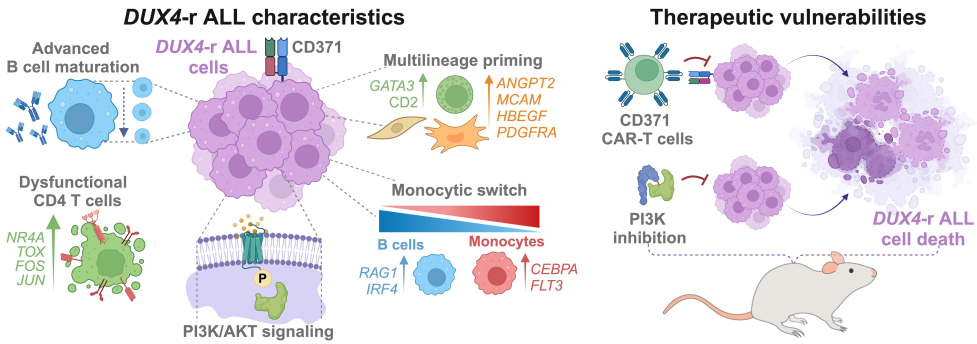

Figure AI-3 | Overview of *DUX4*-r ALL characteristics and therapeutic vulnerabilities. A summary of therapeutic vulnerabilities identified in *DUX4*-r ALL using multimodal single-cell sequencing and validation of therapeutic effects in pre-clinical murine models.

Given the lineage infidelity in *DUX4*-r cases and their tendency to transition to a myeloid phenotype, effectively targeting the entire blast population would necessitate a marker present on cells both in the B cell and myeloid blast compartments. Notably, *CD371* was the only marker identified as being expressed on both lymphoid and monocytic blast cells, offering a promising therapeutic target. In addition to high multilineage priming, activation of PI3K/AKT signaling was found to be significantly upregulated in *DUX4*-r ALL blast cells. The upregulation specifically involved the PI3K subunits encoded by *PIK3CA* (p110α) and *PIK3R1* (p85α), which are known to be targetable by existing FDA-approved PI3K inhibitors. Collectively, this study highlighted two main therapeutic vulnerabilities in *DUX4*-r ALL, which were further tested first *ex vivo* and then *in vivo.* PI3K/AKT signaling activation in *DUX4*-r cells was found to be susceptible to PI3K inhibitors, effectively depleting *DUX4*-r ALL cells *in vivo*. Secondly, the generation of CD371-targeting second-generation 4-1BB-based (CAR) T cells targeting the aberrant myeloid marker CD371 showed effective elimination of *DUX4*-r ALL cells *in vivo* (**Figure AI-3**).

In conclusion, detailed single-cell analysis provides crucial insights into BCP-ALL biology and highlights therapeutic vulnerabilities in the *DUX4*-r ALL, offering promising novel treatment strategies for this subtype.

Article II

The complement receptor C3AR constitutes a novel therapeutic target in *NPM1*-mutated AML

Uncovering AML-specific surface proteins paves the way for precision-targeted therapies. In this study, a FACS-based screen directed at 362 cell surface markers was used to identify novel markers uniquely expressed in *NPM1*-mutated AML.

From this screen, the complement receptor C3AR was identified as being specifically expressed in *NPM1*-mutated AML. C3AR is a G protein-coupled receptor that functions within the complement system and plays a crucial role in immune regulation, inflammation, and cellular signaling.[299–301] To study the cellular landscape of *NPM1*-mutated AML and the expression of *C3AR*, scRNA-seq was performed on bone marrow samples from patients, along with bone marrow samples from healthy donors. Mutation calling of the single-cell reads covering the *NPM1* mutation site allowed us to identify the leukemic cells. These leukemic cell populations exhibited notably high *C3AR* expression compared to residual normal hematopoiesis (**Figure AII-1**). Flow cytometry and scRNA-seq of normal hematopoietic stem and progenitor cells demonstrated that C3AR was not expressed in these primitive populations, providing an opportunity to develop antibody-based therapies specifically targeting this cell population, while sparing normal progenitor cells.

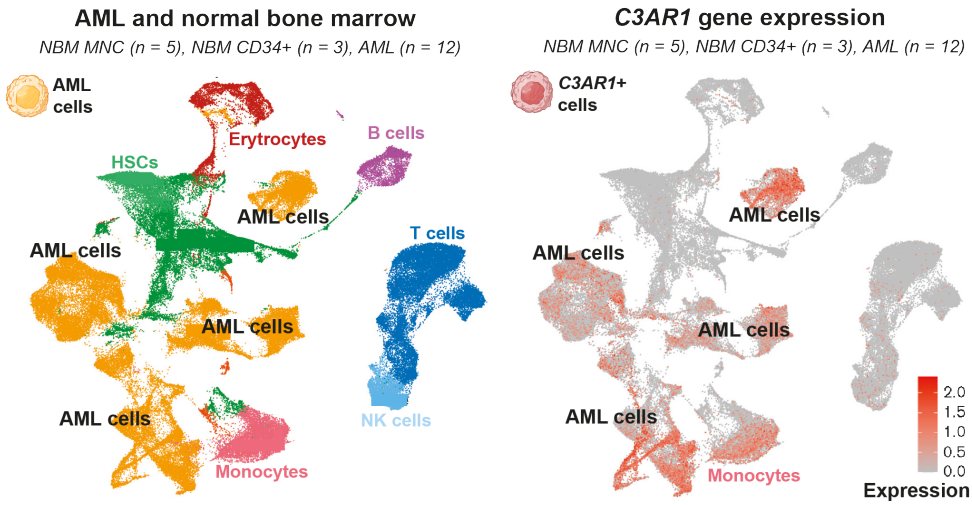

Figure AII-1 | *C3AR1* expression in *NPM1*-mutated AML blasts.
ScRNA-seq confirms *C3AR1* expression in *NPM1*-mutated AML blasts and normal monocytes. UMAP visualization shows cell types (left) and *C3AR1* expression, with red indicating high expression (right). NBM; normal bone marrow, and MNC; mononuclear cells.

The leukemia-initiating potential of C3AR-expressing cells was further confirmed through transplantation into immunodeficient mice. Notably, co-expression of C3AR and GPR56 identified a population with enhanced leukemia-initiating capacity, thereby defining LSCs in *NPM1*-mutated AML (**Figure AII-2**). To study a possible functional role of C3AR expression on *NPM1*-mutated cells, we stimulated C3AR-expressing cells with its ligand, C3a. This stimulation was found to enhance AML cell survival and specifically activate the MAPK pathway via

ERK1/2 signaling, highlighting a critical signaling axis in this AML subtype. Additionally, *NPM1*-mutated AML cells were shown to produce complement factor D (CFD), a crucial enzyme in the alternative complement pathway.[302] Notably, CFD was the most upregulated gene in *NPM1*-mutated AML cells compared to normal bone marrow cells, as revealed by scRNA-seq. These findings imply that *NPM1*-mutated AML cells exploit the complement system to their advantage, promoting survival and proliferation, while evading immune responses, through C3AR signaling (**Figure AII-2**). Targeting this pathway suggests a promising therapeutic strategy for treating *NPM1*-mutated AML.

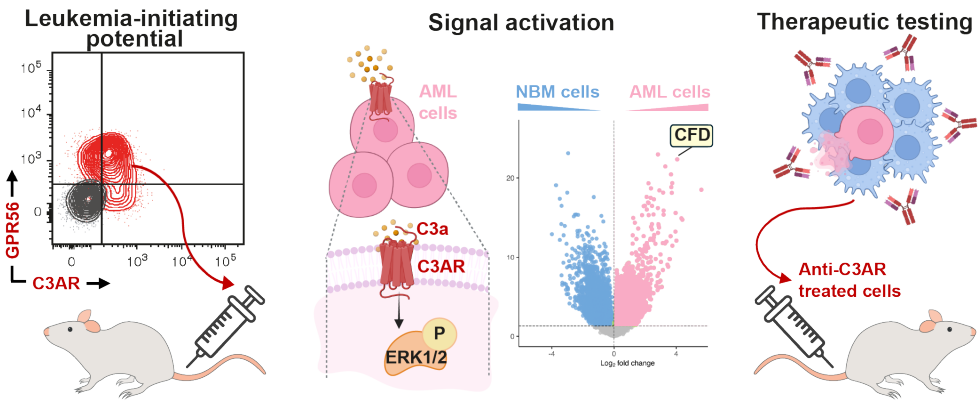

Figure AII-2 | The functional role and therapeutic potential of C3AR signaling.
C3AR and GPR56 identify a population with enhanced leukemia-initiating capacity (left). C3AR activates the MAPK pathway via ERK1/2 signaling. *CFD* is upregulated in *NPM1*-mutated AML cells (center). C3AR antibody treatment of *NPM1*-mutated AML cells, followed by transplantation into immunodeficient mice shows therapeutic efficacy (right).

The therapeutic potential of targeting C3AR was tested using antibodies directed against the receptor. These antibodies effectively elicited NK cell-mediated killing of primary *NPM1*-mutated AML cells *ex vivo* while sparing normal HSCs. Furthermore, when residual primary AML cells treated with C3AR antibodies were transplanted into immunodeficient mice, they engrafted at lower levels than control-treated cells, underscoring the therapeutic efficacy of C3AR antibodies. In summary, these findings highlight C3AR as a therapeutic target in *NPM1*-mutated AML, the most common form of AML, offering the potential to develop antibody-based therapies to target the LSC compartment in this subtype of AML.

Article III

The cellular state space of AML unveils novel *NPM1* subtypes with distinct clinical outcomes and immune evasion properties

To identify novel subtypes of AML, a comprehensive integrative sequencing analysis was performed on 120 consecutive matched tumor-normal AML cases using whole exome sequencing, mate-pair whole genome sequencing, and RNA sequencing. The transcriptional profiles from the bulk RNA gene expression data did not cluster based on mutational patterns. They were instead primarily driven by cellular signatures, potentially masking the intrinsic features of the AML blast cells in each subtype. To delineate these cellular signatures, scRNA-seq was performed on samples from 38 AML cases from the original cohort, along with normal bone marrow samples. While *RUNX1::RUNX1T1* and *CBFB::MYH11* AML showed uniform cellular patterns, other subtypes exhibited cellular heterogeneity, even within the same genomic group, highlighting biological differences beyond the current genomic classifications.

A novel mutation tracking method (scRNAmut-seq) was developed to determine which cells constituted the leukemic clone in each sample. All AML samples contained a distinct cluster of immature cells that were clearly distinguishable from immature cells in CD34-enriched normal bone marrow samples. Data from scRNAmut-seq analysis highlighted that close to 100% of these immature cells in each AML sample harbored AML mutations. Hierarchical clustering of the gene expression profiles of immature cell populations across samples provided a clearer representation of genetic subgroups compared to bulk profiles, grouping patients with shared genetic alterations into distinct clusters (**Figure AIII-1**). However, AML-MR and *TP53*-mutated AML were grouped together, highlighting shared molecular features between these subtypes. Intriguingly, the *NPM1*-mutated samples formed two distinct groups within the *NPM1* group, with distinct expression profiles. These two groups were designated "*NPM1* class I" and "*NPM1* class II".

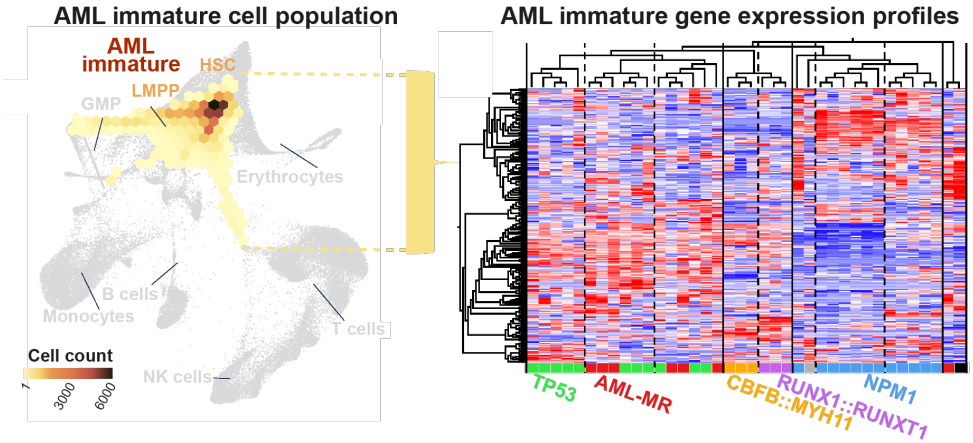

Figure AIII-1 | Identification of the most immature cells in AML.
Single-cell projection of "AML immature" cells onto normal bone marrow highlights their similarity to immature normal bone marrow cells (left). Hierarchical clustering of average gene expression profiles from AML immature cells delineates the genetic subgroups better than bulk sequencing (right).

The newly identified *NPM1* classes were associated with differing cell distributions: *NPM1* class I predominantly consisted of immature AML cells, while *NPM1* class II included both immature and differentiated AML cells. By generating a gene list based on the differential expression in immature AML cells, *NPM1* class I and *NPM1* class II could be clearly distinguished in bulk gene expression data from our own dataset and external ones such as Beat AML and TCGA (**Figure AIII-2**). However, a third group of the samples in the external datasets could not be assigned to either of the two *NPM1*-mutated subtypes, most likely due to low content of immature AML cells.

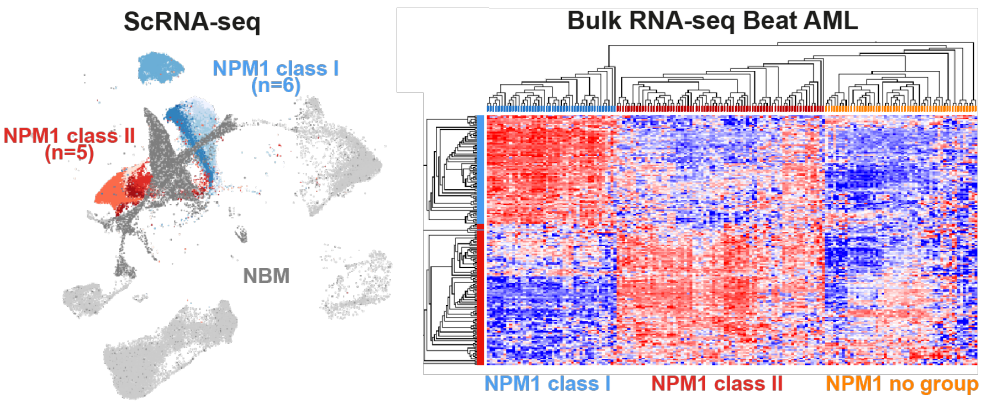

Figure AIII-2 | Identification of *NPM1* class I and *NPM1* class II.
UMAP representation of cells from normal bone marrow samples (gray) and AML immature cells from individual *NPM1*-mutated samples, with cases from *NPM1* class I indicated in blue and *NPM1* class II cases indicated in red (left). Hierarchical clustering of bulk gene expression data from 174 *NPM1*-mutated AML samples from the Beat-AML cohort shows *NPM1* class I genes (indicated in blue) and *NPM1* class II genes (indicated in red) (right).

Focusing on differences between *NPM1* class I and *NPM1* class II, mutations in *IDH2*, *TET2*, and *SRSF2* were significantly associated with *NPM1* class I, whereas mutations in *FLT3*, *DNMT3A*, *NRAS*, *PTPN11,* and *WT1* were more common in *NPM1* class II (**Figure AIII-3**). Notably, the *NPM1* classification also correlated with clinical outcomes. While *NPM1* class II was associated with dismal survival after hematopoietic stem cell transplantation, *NPM1* class I, demonstrated significantly improved survival following hematopoietic stem cell transplantation. Additionally, *NPM1* class I showed downregulation of genes encoding MHC Class II molecules as a potential immune evasion mechanism, whereas *NPM1* class II was instead associated with increased resistance to allogeneic T cells. Thus, this difference in immune evasion mechanisms between the two *NPM1*-mutated subtypes could potentially explain the survival difference following hematopoietic stem cell transplantation.

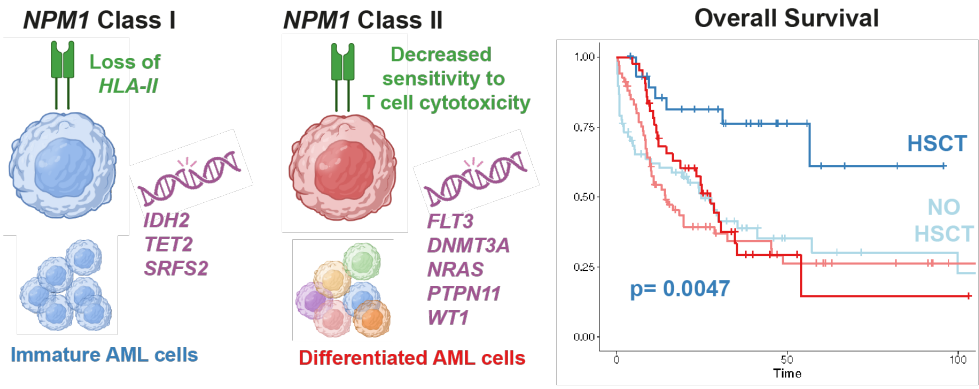

Figure AIII-3 | Summary of characteristics in *NPM1*-mutated AML subdivision.
Differences in immune evasion mechanisms, mutational status, and cellular composition were observed between the two different *NPM1*-mutated AML subtypes. These *NPM1* classes also correlated with clinical outcomes, with *NPM1* class II cases not benefiting from HSCT (red), while NPM1 class I showed significantly improved survival (blue).

In summary, this study demonstrates that bulk AML gene expression profiles are driven by a diverse set of cellular signatures. This was confirmed by scRNA-seq, which also revealed an unexpected degree of cellular heterogeneity that extends beyond current genomic classifications. Notably, *NPM1*-mutated AML could be stratified into two novel, clinically relevant subclasses, each associated with distinct immune evasion mechanisms and differences in survival following hematopoietic stem cell transplantation. These findings provide novel insights into the cellular landscape of AML, define new diagnostic entities, and highlight potential therapeutic intervention points.

Article IV

Mapping the cellular state space of pediatric and adult AML unveils complex differentiation patterns and defines the subtype-specific maturation of the leukemia stem cell compartment

Although age is a known prognostic factor in AML, the similarities and differences in the single cell landscapes of adult and pediatric AML remain poorly investigated. To address this, we used scRNA-seq data from 35 cases of adult AML (**Article III**) and collected data from two recently published studies including 20 adult and 48 pediatric AML cases, respectively.[175,187] After manual data curation, stringent quality control, and batch correction, the final dataset included 55 adult and 48 pediatric AML cases, along with 18 healthy donor bone marrow samples.

The cellular state space of both adult AML and pediatric AML was found to be extensively remodeled compared to normal bone marrow samples. Each AML sample was associated with a distinct cluster of abnormal cells that constituted the bulk of the AML samples, distinguishing them from all cell types observed in normal bone marrow. The abnormal cells were classified into ten distinct cell types, spanning a maturation axis from primitive to myeloid cells, along with two lymphoid cell types (**Figure AIV-1**). When classifying the abnormal AML cells into distinct cell types along the normal differentiation axis, no major differences were observed between adult and pediatric AML. Instead, these features were found to be associated with the genetic subtype. Most subtypes exhibited an enrichment of cells with primitive characteristics, resembling normal HSCs, LMPPs, and immature GMPs. However, two notable exceptions were observed: *KMT2A*-r AML displayed a notable enrichment of cells resembling differentiated GMPs, whereas AML-MR lacked cells resembling GMPs cells and displayed lymphoid transcriptional features (**Figure AIV-1**). In addition, AML-MR had a significant increase in apparently normal CLPs, suggesting either that LSC or other leukemia-related immature cells retain the capacity to produce lymphoid cells.

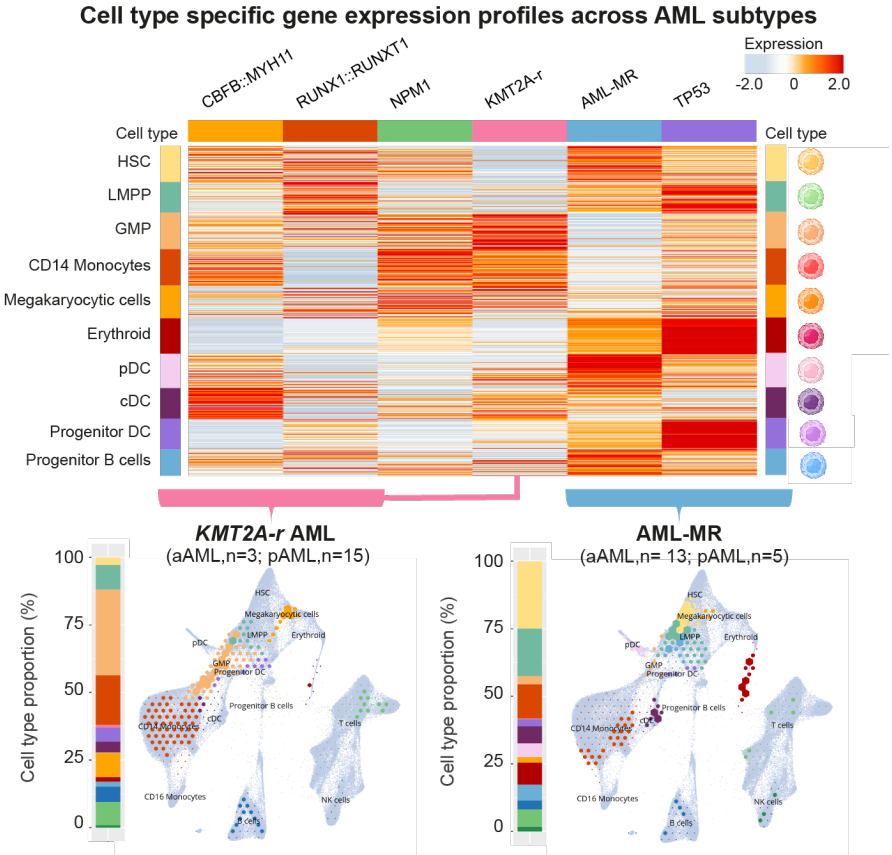

Figure AIV-1 | The maturation characteristics of AML.
Heatmap showing cell type-specific gene expression across various AML genetic subtypes, with red indicating high expression (upper). Cell composition is visualized by bar plots, and maturation characteristics are displayed in single-cell projections of AML cells onto a normal bone marrow reference for *KMT2A*-r AML (lower left) and AML-MR (lower right).

Subtypes such as *CBFB::MYH11* and *NPM1*-mutated AML, in addition to primitive cells, also displayed enlarged monocytic compartments, whereas *TP53*-mutated AML showed a notably higher proportion of erythroid cells. Although the maturation characteristics of AML were primarily determined by genetic subtypes, differences within each subtype were also observed. Consistent with our findings in **Article III**, *NPM1*-mutated AML exhibited varying degrees of primitive cells and monocytes. Similar variations were seen in AML-MR and *TP53*-mutated AML. These differences within genetic subtypes of AML suggest biological differences that extend beyond current genomic classifications.

By applying a stemness score, we could identify putative LSCs and determine their maturation (**Figure AIV-2**). This revealed substantial variability in LSC maturation

that was associated with genetic subtype but not with age, although a non-significant trend was observed for pediatric LSCs having a more differentiated profile. *TP53*-mutated LSCs exhibited the most immature profile among all subtypes.

Notably, most LSC populations resembled the LMPP compartment rather than the stem cell compartment, except for *KMT2A*-r AML, where the LSCs were positioned within the GMP compartment. Despite the absence of major differences in maturation characteristics of both LSCs and the more differentiated AML cells between adults and children, gene expression differences were observed within specific cell types when comparing the two age groups (**Figure AIV-2**). This suggests that molecular programs in these cell types vary across ages. Notably, pediatric AML LSCs exhibited increased inflammatory and chemokine signaling, along with the activation of non-hematopoietic characteristics such as epithelial-mesenchymal transition (EMT) and angiogenesis. These findings suggest that LSCs in adult and pediatric AML may differ in how they interact with the bone marrow microenvironment. In line with this, the cell surface marker MSLN (mesothelin), which is thought to play a role in cell adhesion within the extracellular matrix, was identified as specifically upregulated in pediatric patients. This potential interaction may constitute a candidate for immunotherapy in children diagnosed with AML.

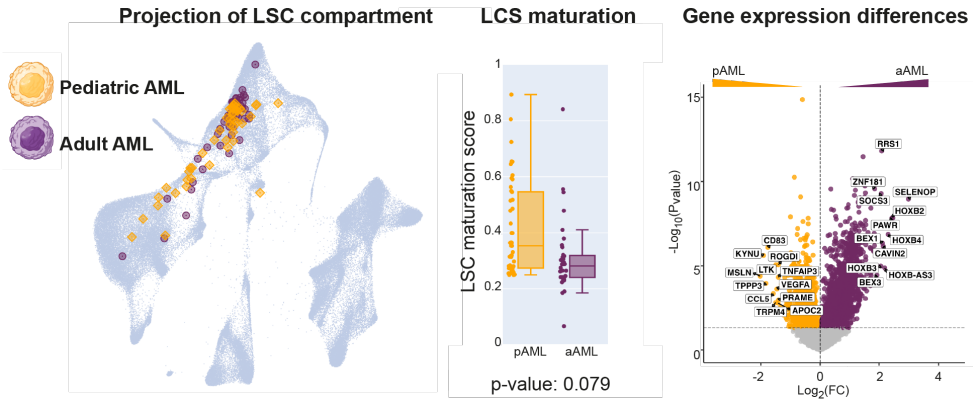

Figure AIV-2 | Identification of putative LSC.
Putative quiescent LSC populations were projected onto the normal bone marrow reference map (left), with each population colored by age group. The LSC maturation score across patients showed no significant difference between age groups when controlling for subtype (center). However, a significant gene expression difference was observed between LSCs in aAML and pAML (right).

Collectively, this study provides valuable insights into cellular hierarchies and LSC features across ages, highlighting biological characteristics that are not fully captured for in today's genomic classification system. These findings may contribute to a more refined classification and treatment approach for both adult and pediatric AML.

Conclusions

The main conclusions from the present study can be summarized as follows:

Article I | Single-cell genomics details the maturation block in BCP-ALL and identifies therapeutic vulnerabilities in *DUX4*-r cases

- **BCP-ALL subtypes exhibit distinct maturation patterns:** *BCR::ABL1*+, *ETV6::RUNX1*+, and HeH subtypes are primarily arrested at the pro-B cell stage, whereas *DUX4*-r ALL cells exhibit transcriptional signatures resembling mature B cells and display strong evidence of multilineage priming.
- ***DUX4*-r ALL is highly susceptible to targeted therapies:** *DUX4*-r ALL cells demonstrate marked sensitivity to PI3K inhibitors and high susceptibility to CD371 CAR-T cell cytotoxicity *in vivo*.

Article II | The complement receptor C3AR constitutes a novel therapeutic target in *NPM1*-mutated AML

- **C3AR$^+$/GPR56$^+$ cells define the LSC population in *NPM1*-mutated AML:** C3AR is selectively expressed on the entire population of *NPM1*-mutated AML cells and, when co-expressed with GPR56, defines the LSC population within this subtype. Functionally, C3AR plays a role in the complement system and can activate ERK1/2, while also contributing to the survival of AML cells.
- **C3AR constitutes a novel therapeutic target in *NPM1*-mutated AML:** Targeting C3AR with antibodies effectively eliminates AML cells while sparing normal progenitor cells, making it a highly attractive therapeutic strategy. In contrast, GPR56 is also expressed on normal HSCs, which limits its therapeutic potential.

Article III | The cellular state space of AML unveils novel *NPM1* subtypes with distinct clinical outcomes and immune evasion properties

- **The bulk transcriptional profiles of AML are mainly driven by a diverse set of cellular signatures:** Bulk AML gene expression profiles are influenced by cellular composition, which mask the intrinsic leukemic features. ScRNA-seq uncovers marked cellular heterogeneity, with immature AML cells exhibiting subtype-specific transcriptional patterns that extend beyond current genomic classifications.

- **Identification of novel *NPM1*-mutated AML subclasses:** *NPM1*-mutated AML can be classified into two novel, clinically relevant, subtypes (*NPM1* class I and *NPM1* class II) that exhibit distinct immune evasion mechanisms and different survival outcomes after stem cell transplantation.

Article IV | Mapping the cellular state space of pediatric and adult AML unveils complex differentiation patterns and defines the subtype-specific maturation of the leukemia stem cell compartment

- **AML maturation patterns are determined by genetic subtype rather than age:** Analysis of over 500,000 single cells from 103 AML cases identified ten distinct AML cell types along the maturation axis. The cellular composition and maturation characteristics of AML are mainly determined by genetic subtype and not age. However, age-related transcriptional differences were observed within the same cell types.
- **LSC maturity is highly variable in pediatric and adult AML:** LSC maturity as defined by stemness scores shows high variability but correlates with genetic subtype. Pediatric and adult LSCs show differences in their transcriptional programs, revealing potential therapeutic vulnerabilities.

General discussion and future perspectives

With the aim to improve disease understanding, refine diagnostic approaches, and identify novel therapeutic targets in acute leukemia, the present study applied multimodal single-cell sequencing technologies to dissect the cellular and molecular complexity of BCP-ALL and AML. Leveraging these techniques and developing novel bioinformatic approaches, the present study was able to infer the cellular composition, maturation patterns, and aberrant transcriptional programs of different cell populations. In addition, a specific focus was on studying the LSC compartment to delineate aberrant transcriptional programs and to identify new therapeutic targets.

By studying the cellular composition and maturation patterns in BCP-ALL and AML (**Articles I, III,** and **IV**) using well-annotated reference data sets, known marker genes, and force-layout plots, in which the leukemic cells were projected onto either the normal B cell differentiation axis or the single-cell space of normal bone marrow samples, we were able to delineate the cellular composition and maturation patterns in great detail. This approach allowed for a detailed delineation of differentiation patterns. For example, in BCP-ALL, *DUX4*-r cases exhibited three main differentiation patterns and showed signs of multilineage priming. Similarly, in AML, the most common genetic subtypes demonstrated similarities in

differentiation patterns but also displayed marked heterogeneity (**Article III and IV**). This approach revealed patterns that associated with the genetic subtype but also findings extending beyond current genomic classification schemes. Within *NPM1*-mutated AML, single-cell analysis identified two novel subclasses (*NPM1* class I and II), each characterized by distinct gene expression signature within their immature cell compartments (**Article III**). Using defined gene expression signatures of these populations, it was possible to assess their prognostic impact in large bulk RNA-sequencing data sets, revealing differences in outcomes following stem cell transplantation. Functionally, *NPM1* class I exhibited downregulation of MHC class II molecules as a possible immune evasion strategy, whereas *NPM1* class II displayed increased resistance to T cell-mediated cytotoxicity. These findings may explain the observed differences following stem cell transplantation. Thus, the novel *NPM1*-mutated subtypes may become important in the future clinical diagnostic work-up of AML.

Despite advancements in the therapeutic development of BCP-ALL and AML, chemotherapy remains the backbone of current treatment and is associated with severe side effects, particularly in growing children. This urgently calls for more targeted therapies that are both more effective against the leukemic cells and associated with fewer side effects. In this context, single-cell analysis allows the study of aberrant transcriptional programs in any of the defined cell populations, facilitating the identification of new therapeutic targets. In this thesis several new targets were identified: For example, in *DUX4*-r BCP-ALL, dysregulation of the PI3K/AKT pathway was identified and marked sensitivity to PI3K inhibitors, both *ex vivo* and *in vivo*, was demonstrated (**Article I**). Additionally, the cell surface marker CD371 (CLL-1), a well-known marker of AML stem cells,[158] was found to mark both lymphoid and monocytic blast cells in *DUX4*-r ALL. These findings were used to engineer CAR-T cells directed against CD371 and to demonstrate proof-of-concept of a strong anti-leukemic activity *in vivo* using patient-derived xenograft models of *DUX4*-r ALL (**Article I**). Finally, in **Article II**, using an arrayed flow cytometry-based screen of 362 cell surface markers, combined with single-cell analysis, C3AR was identified as being specifically expressed on both bulk AML cells and the LSCs population in *NPM1*-mutated AML. Antibodies directed against C3AR efficiently elicited NK cell-mediated killing of primary AML cells *ex vivo*, highlighting C3AR as a promising therapeutic target in *NPM1*-mutated AML.

A central focus of leukemia research remains the characterization and targeting of LSCs, which are pivotal in leukemia initiation, progression, and relapse. Understanding their nature and targeting them specifically remains the ultimate goal in leukemia research as this may result in a potential cure of these malignancies. In **Articles IV**, single-cell analysis was used to study their molecular properties. As LSCs are operationally defined by transplantation into immunodeficient mice, it was only possible to study this population indirectly using different bioinformatic approaches. A stemness score was developed to identify the LSCs, and their

maturation was inferred by projection onto normal bone marrow cells. While substantial heterogeneity was observed, the maturation of LSCs was mainly associated with the genetic subtype. Currently, work is on-going within the research group to further characterize the LSC compartments in the different subtypes of AML. While LSCs are defined by their ability to self-renew and sustain the bulk of more differentiated leukemic cells, their precise cellular origin also remains a subject of ongoing investigation. In this thesis, single-cell analysis provided direct insights into the cellular origin of LSCs in BCP-ALL by resolving *IGH* and *IGL* chain rearrangements that appear to be fixed at the time of leukemia development. The findings suggest that LSCs can arise from different progenitor stages depending on the genetic subtype. For example, in BCP-ALL, *BCR*::*ABL1*-positive cases appeared to originate from an early hematopoietic stage that lacked both *IGH* and *IGL* rearrangements, whereas *DUX4*-r ALL cases were more differentiated, where *IGH* and *IGL* chain rearrangements appeared to have been present in the cell of origin. In AML, instead, the maturation level of LSCs was studied, providing a more indirect indication of the maturation level also of the cell of origin. There, most LSCs displayed an LMPP-like phenotype, with some subtypes, such as *KMT2A*-r AML, showing LSCs at later differentiation stages, including the GMP compartment. However, in the latter instances, it is for example possible that the cell of origin is a more committed progenitor that through the oncogenic hits regained stem cell-like properties.

Although age is a known prognostic factor in leukemia, the similarities and differences in the single-cell landscapes of adult and pediatric leukemia remain poorly studied. The frequency of defining genetic alterations differs greatly between children and adults, which may imply differences in disease initiation, progression, and treatment response. Additionally, targeted therapy approvals have historically benefited adult AML patients, while pediatric AML has lacked effective targeted treatment options. Comparing these age groups can reveal shared or unique therapeutic vulnerabilities and identifying age-related variations may help develop age-specific therapeutic strategies. In **Article IV**, we aimed to investigate age-related characteristics by analyzing 103 samples spanning both adult and pediatric AML cases. This analysis revealed no marked differences in cellular composition or maturation characteristics between adult and pediatric AML, instead these features mainly correlated with the genetic subtype and not age. However, age-related transcriptional differences were observed within the same cell types. For example, pediatric LSC populations exhibited higher levels of bone marrow remodeling and inflammatory signaling compared to adult AML. These findings suggest that molecular programs within specific cell types vary across ages, potentially influencing disease progression and treatment response. Although work is still ongoing, preliminary findings point to potential therapeutic intervention opportunities for both adult and pediatric patients.

The rapid development of single-cell technologies continues to revolutionize leukemia research by enabling a more precise characterization of cellular heterogeneity, disease evolution, and therapeutic vulnerabilities. While current scRNA-seq techniques provide high-resolution transcriptomic data, emerging technologies are pushing the boundaries further. Single-cell multiomics approaches, which integrate transcriptomic, epigenomic, and proteomic data in the same cell, will allow for a more comprehensive understanding of leukemia at multiple layers. In parallel, advances in spatial transcriptomics will enable researchers to study leukemic cells within their native bone marrow microenvironment, uncovering critical interactions between LSCs, immune cells, and stromal components that drive disease progression. As these technologies become more refined, they will likely be integrated into clinical workflows, allowing for real-time monitoring of leukemia evolution and treatment response. However, before clinical implementation, cost for single-cell analysis need to significantly decrease and more robust workflows need to be implemented to allow clinical grade diagnostics.

Moreover, artificial intelligence (AI) and machine learning-driven data analysis are expected to play an increasingly pivotal role in interpreting the vast datasets generated by single-cell technologies; in this thesis alone, more than 800.000 single cells were analyzed. AI-powered algorithms have the potential to identify hidden patterns in large-scale patient datasets, predict treatment response, and uncover novel therapeutic targets that may have been overlooked using traditional bioinformatic approaches. In addition, AI-driven drug discovery platforms are already accelerating the identification of candidate compounds, potentially reducing the time required to bring new treatments to clinical trials. As AI continues to evolve, it is likely to further streamline biomarker discovery, optimize personalized treatment strategies, and enhance clinical decision-making in leukemia care.

In conclusion, this thesis has demonstrated the power of single-cell sequencing technologies in dissecting the cellular and molecular landscape of acute leukemia, leading to the identification of novel leukemia subtypes, deeper insights into LSC biology, and the discovery of potential therapeutic targets. Undoubtedly, leukemia research is experiencing remarkable progress. Looking ahead, continued advancements will expand the therapeutic arsenal in the fight against leukemia, offering new hope for patients and their families affected by these diseases.

AI tools and software for illustrations

In preparing this thesis, OpenAI's ChatGPT (GPT-4) was utilized to enhance the clarity, grammar, and style of the scientific writing. All AI-generated suggestions in scientific writing were critically reviewed and revised by the author to ensure that the final version fully preserved the original scientific content and integrity, while enhancing only the linguistic quality.

Figures were created using BioRender (https://biorender.com), a tool for professional scientific illustrations.

Populärvetenskaplig sammanfattning

Varje sekund pågår en livsviktig kamp i din kropp, en kamp för att hålla dig vid liv. Miljarder blodceller arbetar hårt, slits ut och dör, bara för att omedelbart ersättas av nya. Varje dag producerar människokroppen ofattbara 300 miljarder blodceller, en intensiv och välorkestrerad process som aldrig får stanna av. Hematopoesen, det vill säga blodbildningen, är ett finjusterat maskineri som snabbt anpassar sig till kroppens behov och är redo att reagera på allt från infektioner till plötslig blodförlust. Alla blodceller har sitt ursprung från en extraordinär stamcell, den hematopoetiska stamcellen. Denna kraftfulla cell besitter en förmåga att självförnya sig och mogna ut till en mångfald av blodceller. Dessa nybildade celler lämnar snabbt benmärgen och ger sig ut i blodomloppet för att påbörja sina livsviktiga uppdrag. Det finns en mängd olika typer av blodceller, var och en med sin unika roll för att säkerställa att kroppen fungerar som den ska. B-celler är exempelvis viktiga soldater i immunförsvaret, medan myeloiska celler fungerar som kroppens städare genom att bryta ner skadliga mikroorganismer och celler. Röda blodkroppar ser i sin tur till att syret når alla kroppens delar.

Blodcellernas DNA bär på den genetiska koden som bestämmer dess funktion. Varje gång en cell delar sig kopieras den genetiska koden och förs vidare till nästa generation av celler. Med tanke på att kroppen producerar hisnande 300 miljarder blodceller varje dag är det kanske inte så förvånande att fel ibland smyger sig in vid kopieringen, vilket kan resultera i att skador i gener (mutationer) uppstår. De flesta mutationer är harmlösa, men om de sker i viktiga gener kan detta leda till att cellerna programmeras på fel sätt. Det kan leda till att de slutar fungera normalt och i värsta fall kan det leda till cancer och när cancer uppstår i blodet kallas det för leukemi.

Att få en diagnos som blodcancer är utmanade, men vad händer egentligen i kroppen? Det hela börjar när en hematopoetisk cell i benmärgen drabbas av mutationer och börjar bete sig som en upprorsmakare. Den förvandlas till en leukemisk stamcell, som sedan skapar en armé av onormala blodceller. De normala blodcellerna som normalt skyddar kroppen och utför livsviktiga funktioner, förlorar sin ursprungliga funktion. Samtidigt delar sig dessa rebeller okontrollerat och tränger undan de friska blodcellerna, vilket försvårar kroppens förmåga att upprätthålla balans och hälsa. För att förstå cancerförloppet är det viktigt att studera enskilda celler och deras samspel med omgivande celler.

Det finns många olika varianter av leukemi, men vad avgör vilken typ som uppstår? Svaret ligger med stor sannolikhet i vilka mutation som inträffar och i vilken celltyp. Leukemi är en komplex sjukdom, där olika genetiska förändringar styr hur den utvecklas. Beroende på vilken celltyp som drabbats delas leukemi traditionellt sett in i två grupper, akut myeloisk leukemi (AML) och akut lymfatisk leukemi (ALL).

Idag används genetiska avvikelser (mutationer) för att klassificera sjukdomen och bedöma risknivå, vilket hjälper läkaren att välja rätt behandling. Hos vuxna är AML den vanligaste formen av akut leukemi. Hos barn är det istället ALL som dominerar och denna sjukdom utgör den vanligaste cancerformen bland barn. Här kommer en twist: även om det är ovanligt, kan ALL i vissa fall övergå till AML. Det innebär att sjuka celler kan ändra skepnad, från B-celler till myeloiska celler, eller tvärtom. Föreställ dig att blodceller är som frukter i en fruktkorg. B-celler kan liknas vid äpplen, men under selektionstryck, som vid leukemi under behandling, kan ett äpple plötsligt förvandlas till ett päron. Dessa oväntade förändringar visar att cancercellerna har en egen överlevnadsstrategi och kan byta taktik när det behövs. Leukemi är långt mer komplext än vad många tror. Det handlar inte bara om mutationer, utan om ett dynamiskt cellulärt ekosystem där både friska och sjuka blodceller påverkar varandra. Detta ekosystem styr sjukdomens utveckling och kan även påverka hur den svarar på behandling.

Ålder vid diagnos är en viktig prognostisk faktor vid leukemi. Även om överlevnaden är relativt god hos barn, sjunker den avsevärt med stigande ålder. Hos patienter över 65 år är överlevnaden endast cirka 10 procent. Trots årtionden av banbrytande forskning och stora framsteg inom leukemibehandling är intensiv cellgiftsterapi fortfarande den primära behandlingsmetoden. Syftet med cellgifterna är att eliminera cancerceller, men behandlingen är mycket krävande och förknippad med svåra biverkningar. Biverkningarna kan vara både kort- och långsiktiga. Exempel på akuta biverkningar är håravfall och nedsatt immunförsvar, medan långsiktiga effekter kan inkludera tillväxtrubbningar, koncentrationssvårigheter och infertilitet. Dessa biverkningar är särskilt förödande för växande barn, men behandlingen är också så aggressiv att den inte alltid kan användas på äldre patienter, vars kroppar helt enkelt inte tål behandlingen. Tyvärr är återfall vanligt eftersom cellgifterna inte alltid lyckas eliminera alla cancerceller. Leukemiska stamceller, som ligger till grund för sjukdomen, är svåra att döda. De överlevande cancercellerna kan dessutom utveckla resistens mot cellgifterna, vilket gör leukemin ännu svårare att behandla och kraftigt försämrar överlevnadsprognosen. En central del av leukemiforskningen idag är att förstå och rikta in sig på de leukemiska stamceller. Att kartlägga deras egenskaper och utveckla sätt att specifikt angripa dem är det ultimata målet eftersom detta kan vara nyckeln till ett möjligt botemedel för sjukdomen.

Att studera genetiska förändringar vid leukemi är avgörande för att kunna ställa en korrekt diagnos och ge rätt behandling. Fram tills nu har det bara varit möjligt att analysera den sammanslagna genetiska koden från en blandning av många olika celler. Som att analysera en smoothie, där alls frukter är mixade och sammanslagna, när målet egentligen är att studera varje enskild frukt. Nu har forskningen fått en kraftfull ny verktygslåda med en nyutvecklade teknik kallad singelcellsekvensering. Med singelcellsekvensering går det att isolera och studera enskilda celler från ett prov bestående av tusentals celler. Plötsligt kan de enskilda frukterna i smoothien plockas ut och granskas var för sig. Tekniken gör det möjligt att analysera akut leukemi på en detaljnivå som tidigare varit utom räckhåll.

I denna avhandling användes singelcelltekniken för att studera celler i benmärgsprover från både vuxna och barn som insjuknat i leukemi. Totalt isolerades och karakteriserades omkring en miljon celler, en process som kräver både teknisk noggrannhet och avancerad databearbetning, så kallad bioinformatik. Genom att kartlägga cellernas genuttrycksmönster kunde olika cellpopulationer identifierats och deras mognadsgrad fastställas. Detta har lett till upptäckten av tydliga genetiska mönster, specifika för olika patientgrupper, och möjliggjort en mer exakt klassificering av leukemier som kan ligga till grund för framtida förbättrad diagnostik. Pusslet slutar inte där. Bland de upptäckta cellpopulationerna har även omogna celler undersökts, vilka sannolikt utgör de leukemiska stamcellerna, de celler som driver sjukdomen, gör den svårbehandlad och orsakar återfall efter given behandling. Undersökningarna i denna avhandling som baseras på fyra vetenskapliga arbeten (**Artiklarna I-IV**) har förutom nya biologiska insikter också avslöjat nya angreppspunkter för mer träffsäkra behandlingar.

Sammanfattningsvis har denna avhandling identifierat nya leukemityper, gett fördjupad kunskap om leukemiska stamceller och lett till upptäckten av potentiella behandlingar för både vuxna och barn. Med den fortsatta tekniska utvecklingen och fördjupade biologiska insikter ser framtiden ljus ut. Förhoppningsvis kommer framtidens behandlingar att bli mer träffsäkra än någonsin och ge nytt hopp till patienter och deras familjer som drabbas av akuta leukemier.

Acknowledgement

How the same type of cell can be so similar, yet so life-changingly different. How a single blood test result can turn your entire world upside down, making the everyday life you once knew shift in a matter of seconds. Childhood ALL, the most common cancer in children, was once a death sentence. However, with the introduction of chemotherapy in the 1960s, along with optimized dosing and scheduling, it has today become a success story in pediatric oncology, with cure rates now exceeding 90%. Or take something not quite as serious, rheumatism, which I was diagnosed with at the age of four. I was treated with chemotherapy, and today, I can call myself healthy, something that would not have been possible just a few decades ago. All of this is thanks to research and hospital staff, who are absolute wizards at their jobs. To **everyone working in research and healthcare**: in my eyes, you make our society a better place.

There are many experiences and so many people that have inspired me to do what I do today. No one gets far alone, and real progress is made together. Working in this field would not have been possible without the one and only **Thoas Fioretos**. You took me under your wings and showed me how to make a real difference in the world by doing outrageously good research. You are the master of juggling a thousand things at once, keeping them all in motion, and scoring straight into the top corner. Do I even need to mention all the companies, board positions, prestigious grants, and now, your seat in the Royal Swedish Academy of Sciences? It is not without a little pride that I call myself your PhD student. What impresses me the most is your ability to bring an outstanding team together, not only incredibly skilled at what they do but also genuinely wonderful human beings.

Henrik, they say that after a PhD, you know your field better than your supervisors… but I must admit, you have been a tough nut to crack. With you by my side, everything has felt scientifically possible. Thank you for always being there, especially in the toughest moments. You have both a heart and a brain of gold.

Christina, no one is as skilled as you when it comes to genetic abnormalities, and I am sure you could rattle off everything in your sleep. You have taught me the power of believing in myself and speaking out, even when it is not always comfortable. Having a supervisor who is so good at their job is truly a blessing.

Rasmus, you clever genius. My work would never have come to life without your mad bioinformatic skills. Sometimes, I wish I could step into your brain for a day and experience that incredible mathematical intelligence of yours. Behind all that brilliance, there is a kindness and a sense of humor. I always want to share a cup of amazing Uruguayan tea with you, preferably on La Rambla. **Calle,** when I grow up, I want to be just like you. Successful both in your career and personal life. I wish we could watch NBA games live every weekend; too bad we cannot just hop on a flight to New Orleans that often. **Marianne**, you have been my best cheerleader throughout this time. Our group would not have been a happy group without you. Every research group should have a Marianne, exceptionally skilled at what you do, a true doer, and above all, always with a smile on your face. My greatest sorrow during my PhD is when I became a bioinformatician and had to leave my lab bench next to you. We made a great lab team, even when we overdid it with the alcohol and "poisoned" our single cells. **Helena Å**, your laugh always echoes during our lab meetings, and sharing the fascination with *DUX4*-r ALL has been a privilege. **Pablo**, you saved me during the toughest period in revisions, and I will always be grateful for your incredible expertise. **Mia**, you are a true joy-bringer, though new to the group, you have already become indispensable. **Johan**, genomic meetings have been more enjoyable since you arrived, and sharing our fascination for food and drinks has been a pleasure. As for my younger PhD siblings in the Fioretos science family, **Vendela** and **Petr**, let's sit at a Danish sky bar and toast many more times! Being the PhD big sister is such a pleasure, not that you need my help or advice, as your brilliant minds will take you far beyond your goals. **Vendela**, your humor and big heart have lifted my spirits many times. **Petr**, you are like a fish in water programming at night and working clinically during the day - my only tip is to hold on tight to your laptop. As for my older PhD siblings, **Sofia**, you boosted my confidence in science, and it is a privilege to be second author on your paper. Our COVID-19 walks were a highlight of the home office days. **Niklas**, academia runs deep in your bones: Stanford and soon PI! You have already made an incredibly cool journey, and an even more exciting one is ahead of you.

Best division, **Clinical Genetics**, I am proud to be part of this team. It is amazing how such a group of people can create such a great workplace. The laughter and cheerful shouts make it a joy to go to work every day. A work environment like this is not only desirable but crucial for creating incredible science, just like we do! Thanks to everyone who was here during this time: *Felix, Nils, Anette, Axel, Jakob, Ludvig, Tina, Valeriia, Marcus, Josephine, Katrin, Maria, Nils, Ram, Somadri, Anna, Louise, Helena S, Mattias, Ton, Kajsa, Charlotte, Efe, Ellie, Larissa, Minjun, Bertil, Andrea, Stina, Rebeqa, Karolin, Jenny N, Karim, Linda M, Valeria, Figge, Jakob, Jan, Saskia, David, Bahar, Carro, Jenny K, Linda HM, Kaname, Natalie, Subhayan* and of course, the "goding" *Alexandra*! Also, a special mention to Thoas' group, mentioned earlier. A big extra thank you to the founding father of the Department, **Felix**, and your eminent colleague, **Nils**, for showing me what real

curiosity and passion for science look like. **Anette**, your management of administrative tasks is mind-blowing and thank you for always untangling my messy requisitions with ease. **Axel**, for being an incredibly skilled researcher and role model, going from PhD to a well-known PI. It's been an honor to collaborate with you and your amazing first PhD student, **Ludvig**. Let's keep this spirit going! **Marcus**, for your clever remarks, tips, and guidance, which not only made going to work more enjoyable but also inspired me. And thank you for being able to drink my strong, very black, chewable coffee without complaining! **Karim**, for being the indispensible member of the "party pixies." Sharing an office with you has made my PhD twice as fun. **Bahar**, for being my solid rock, keeping me grounded when the winds were the strongest. You are mastering both the clinic and research with brilliance and grace. **Kajsa, Anna,** and **Karolin**, for being role models, and I am going to say it - badass bosses. The Fika mafia: **Tina, Carro, Jenny N, Linda M, Andrea, Marianne, Helena S, Stina, Linda HM and Mia** the ones who truly run the lab with mad skills. You are all amazing, and it is been a privilege to be surrounded by such strong women I admire so much.

Noelia, Natta, Valeria, and **Louise**. You have truly been the best part of my PhD trip. From colleagues to my favorite humans, and that I get to call you my best friends: "best friends with some seriously impressive brainpower". You all are already amazing, with limitless potential, and I know you are unstoppable. You will push boundaries that others cannot even touch. **Noelia**, for being the caring one. **Natta**, for being the sneaky, fun one. **Valeria**, for being incredibly determined, loving, and always providing us with wonderful food. **Louise**, where do I even start? I am not whole without you. Few have made me laugh as much as you. Let's drink pink champagne from steins for the rest of our lives.

Sisterhood, so many incredible women have shaped my journey, and I just want to take a moment to emphasize the power of sisterhood. It is more important than ever that we, as women, celebrate and support one another. With that said, my first and deepest thanks go to my great-grandmother **Hanna** and my grandmother **Ingegärd** - two remarkable women whom I never had the chance to meet, yet who left an undeniable mark on my path. Both were. Fierce believers in education. Hanna trained as a nurse as part of Sweden's very first class of nursing students but never got to practice her profession because she became a mother. My grandmother Ingegärd also became a nurse, but unlike her older brothers, she had to pay her own tuition. Because of you, a seed of determination was planted, a belief in knowledge, in education, and in the power of perseverance. That seed grew strong in Åhus, where my heart will always find its home.

My first female role models, **Charlotte** and **Ylva**. My gymnastics coaches and pioneers of GF Akrobaten that taught me the power of "I can, I want to, I dare to." To my high school crew: **Cecilia, Nina, Matilda, Nicole**, and **Amanda**. Can you

believe we have stuck together for over 15 years? Without you, I would not be the person I am today. **Amanda**, you have been and continue to be one of the most pivotal human beings in my life. We found each other at a time when I needed to turn the page, and from that moment on, it has been *us*. You have shown me that life is beautiful but not to be taken too seriously, and most importantly, that full-fat milk is truly the best. Then came mini-you, our little sunshine, **Hedvig**. Becoming her godmother has made life even more precious. **Margit**, you walked into my life with a smile so warm that I knew instantly, I never wanted to let you go. **Johanna**, what started as a "kidnapping" turned into lockdown and somehow ended with you becoming family. You are everything and more; my ultimate party queen, yet also my rock, the one I can call in the middle of the night when life goes sideways. **Lojz,** because you and I are the definition of sweat and tears, though we both know wine is our liquid of choice. We took on *En Svensk Klassiker* together, pushed through when we were completely broken, and still managed to squeeze out that last bit of energy to achieve something we thought was impossible. If that is not friendship, I do not know what is.

Then we have the guys, starting with André's boys. André seem to collect brothers from every corner of the world, and luckily for me, I got them as a bonus. A special thanks to **Korpengudar** for taking responsibility for André and his slightly alarming golf addiction. **Hampus, Marcus, Ola, Shan, Gustav**, and **Johan**, you have been there for half my life. From wild house parties to running lap after lap at IFK Lund's athletics track, from sharing a sweaty tent at Emmaboda to sipping Mai Tais in Thailand, we have been through it all.

To my beloved Whisky Group *Uisce beatha*: **Johan, Felicia, Marcus, Johanna, Fia**, **Nils** and **André**. Thank you for every crazy adventure, for all the whiskey tastings, midsummer celebrations, Christmas dinners, and New Year's Eve parties. I do it all over again with you… well, maybe not the part where we got blindfolded and left in the woods to find our way home the fastest. The real question is: when are we finally buying that whiskey barrel?

Now, we come to the bioglogical family.

To the **Thorssons**. **Farmor** and **Farfar**, you never had the opportunity to pursue higher education, yet you were both incredibly wise. **Farfar**, you always preached that if you work with something you love, you will live a happy life. I carry that lesson with me every day. **Henrik, Frida, Elin, Hans**, and **Ulrika**, thank you for a wonderful childhood filled with celebrations, gingerbread baking, and joyful chaos, and for a future in Kivik. Until then, I will keep humming *"Här kommer familjen Thorsson"* - is it too late to apply for Lilla Melodifestivalen?

To the **Westergrens. Morfar**, it was you and me against the world, and we always won against those terrifying motorcycles. **Elisabeth, Håkan**, and **Erik**, thank you

for all the unforgettable summer memories. **Fanny**, you are more than a cousin; you are a sister. We are equally crazy, and your creative mind and fierce determination, convince me that you can handle anything life throws your way. **Jakob**, for being the family's most academic one yet also the cool one with tattoos. I have always been fascinated by how you make every discussion both socially relevant and entertaining. Professor, uncle, and godfather **Albert**, your thirst for life, childlike spirit, and just the right amount of stubbornness are the greatest legacy I could ever wish for. **Barbro**, thank you for making the family whole.

To the **Schoug**: **Jöns, Pia,** and **Fanny**, you are the extra family everyone dreams of. Thank you for welcoming me into your warm-hearted and wonderfully playful family. You have shown me how to add a golden touch to life's simplest moments. Life is a celebration, and if not… *sous la vous*! (C'est la vie)

Mamma and **Pappa**, thank you for filling me with confidence from day one and making everything feel possible. Even though I have gone from *Do not tell me what to do* to *Please tell me what to do*, from *Leave me alone* to *Never leave my side*, from *You do not understand* to *I do not understand*. One thing has remained unchanged: my heart will always belong to the two who created it. **Mamma**, the first woman dean in the history of Lund University, you taught me the most important lesson of life: to do everything with heart and give my all. **Pappa**, you are the ultimate source of inspiration for fixing, planning, and the very foundation of the project manager in me. "Min pappa sa att..." I can still say that with full conviction and be absolutely sure it is right. You are possibly the most knowledgeable man on the planet, but we both know that is simply because you have the biggest head on the planet.

Speaking of big heads, **Johan** *Ljushuve* **Thorsson**. Not only do I have the best friend in the world, but I also get to share your brilliant genes. These are not just any genes: they have now landed you an amazing PhD position at The Karolinska Institute! But hey, let's not get caught up in rankings… Let's just remember which university has been around the longest. Though I will never openly admit it, we both know who has the best rhythm. So, please let's keep dancing through life together!

Min älskling, **André**, thank you for teaching me the art of being fully present; for making me pause, look up, and truly live, rather than getting lost in plans and routines. Livet med dig, *c'est la kiffance.* You are the most outstanding human being, the most brilliant minded, and the only person who has ever truly figured me out, something I'm still learning to do myself. If I am to continue this journey of self-discovery, there is no one I would rather do it with than you.

References

1. Alberts B, Heald R, Johnson AD, et al. Molecular Biology of the Cell. W. W. Norton; 2022.
2. Smith E. *Edwin Smith Surgical Papyrus*. Translated by James P. Breasted. The University of Chicago Press; 1930.
3. Watson JD, Crick FHC. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*. 1953;171(4356):737–738.
4. Boveri T. Zur frage der entstehung maligner tumoren. Fischer; 1914.
5. Nowell, P.C. and Hungerford, D.A. (1960) A minute chromosome in human chronic granulocytic leukemia. *Science* 132, 1497–1497.
6. Hanahan D, Weinberg RA. The Hallmarks of Cancer. *Cell*. 2000;100(1):57–70.
7. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell*. 2011;144(5):646–674.
8. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov*. 2022;12(1):31–46.
9. Ferlay J, Ervik M, Lam F, et al. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. (2022). Available at: https://gco.iarc.who.int/today.
10. International Human Genome Sequencing Consortium, Whitehead Institute for Biomedical Research, Center for Genome Research:, Lander ES, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860–921.
11. Nurk S, Koren S, Rhie A, et al. The complete sequence of a human genome. *Science*. 2022;376(6588):44–53.
12. Mardis ER. Next-Generation DNA Sequencing Methods. *Annu Rev Genom Hum Genet*. 2008;9(1):387–402.
13. Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer Genome Landscapes. *Science*. 2013;339(6127):1546–1558.
14. The Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013;45(10):1113–1120.
15. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The Next-Generation Sequencing Revolution and Its Impact on Genomics. *Cell*. 2013;155(1):27–38.
16. Pevsner, Jonathan. *Bioinformatics and Functional Genomics*. 3rd ed., Wiley-Blackwell; 2015.

17. Tang F, Barbacioru C, Wang Y, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*. 2009;6(5):377–382.
18. Ramsköld D, Luo S, Wang Y-C, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol*. 2012;30(8):777–782.
19. Rozenblatt-Rosen O, Stubbington MJT, Regev A, Teichmann SA. The Human Cell Atlas: from vision to reality. *Nature*. 2017;550(7677):451–453.
20. Rozenblatt-Rosen O, Regev A, Oberdoerffer P, et al. The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution. *Cell*. 2020;181(2):236–249.
21. Baysoy A, Bai Z, Satija R, Fan R. The technological landscape and applications of single-cell multi-omics. *Nat Rev Mol Cell Biol*. 2023;24(10):695–713.
22. Chen S, Lake BB, Zhang K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nat Biotechnol*. 2019;37(12):1452–1457.
23. Notta F, Zandi S, Takayama N, et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*. 2016;351(6269):aab2116.
24. Till JE, McCulloch EA. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat Res*. 1961;14:213–222.
25. Till JE, McCulloch EA. Hemopoietic stem cell differentiation. *Biochim Biophys Acta Rev Cancer*. 1980;605(4):431–459.
26. Hoffman R, Benz EJ, Silberstein LE, et al. *Hematology: Basic Principles and Practice*. Elsevier Inc. 2017.
27. Takizawa H, Boettcher S, Manz MG. Demand-adapted regulation of early hematopoiesis in infection and inflammation. *Blood*. 2012;119(13):2991–3002.
28. Haas S, Trumpp A, Milsom MD. Causes and Consequences of Hematopoietic Stem Cell Heterogeneity. *Cell Stem Cell*. 2018;22(5):627–638.
29. Laurenti E, Göttgens B. From haematopoietic stem cells to complex differentiation landscapes. *Nature*. 2018;553(7689):418–426.
30. Spangrude GJ, Heimfeld S, Weissman IL. Purification and Characterization of Mouse Hematopoietic Stem Cells. *Science*. 1988;241(4861):58–62.
31. Osawa M, Hanada K, Hamada H, Nakauchi H. Long-Term Lymphohematopoietic Reconstitution by a Single CD34-Low/Negative Hematopoietic Stem Cell. *Science*. 1996;273(5272):242–245.
32. Kondo M, Weissman IL, Akashi K. Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. *Cell*. 1997;91(5):661–672.
33. Akashi K, Traver D, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature*. 2000;404(6774):193–197.
34. Sanjuan-Pla A, Macaulay IC, Jensen CT, et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*. 2013;502(7470):232–236.

35. Yamamoto R, Morita Y, Ooehara J, et al. Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells. *Cell*. 2013;154(5):1112–1126.
36. Cabezas-Wallscheid N, Klimmeck D, Hansson J, et al. Identification of Regulatory Networks in HSCs and Their Immediate Progeny via Integrated Proteome, Transcriptome, and DNA Methylome Analysis. *Cell Stem Cell*. 2014;15(4):507–522.
37. Pietras EM, Reynaud D, Kang Y-A, et al. Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell*. 2015;17(1):35–46.
38. Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol*. 2017;19(4):271–281.
39. Tusi BK, Wolock SL, Weinreb C, et al. Population snapshots predict early haematopoietic and erythroid hierarchies. *Nature*. 2018;555(7694):54–60.
40. Carrelha J, Meng Y, Kettyle LM, et al. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*. 2018;554(7690):106–111.
41. Rodriguez-Fraticelli AE, Wolock SL, Weinreb CS, et al. Clonal analysis of lineage fate in native haematopoiesis. *Nature*. 2018;553(7687):212–216.
42. Diamond MS, Kanneganti T-D. Innate immunity: the first line of defense against SARS-CoV-2. *Nat Immunol*. 2022;23(2):165–176.
43. Rothenberg EV, Moore JE, Yui MA. Launching the T-cell-lineage developmental programme. *Nat Rev Immunol*. 2008;8(1):9–21.
44. Nemazee D. Mechanisms of central tolerance for B cells. *Nat Rev Immunol*. 2017;17(5):281–294.
45. Nimmo RA, May GE, Enver T. Primed and ready: understanding lineage commitment through single cell analysis. *Trends Cell Biol*. 2015;25(8):459–467.
46. Hofmann J, Kokkaliaris KD. Bone marrow niches for hematopoietic stem cells: life span dynamics and adaptation to acute stress. *Blood*. 2024;144(1):21–34.
47. Hadland B, Varnum-Finney B, Dozono S, et al. Engineering a niche supporting hematopoietic stem cell development using integrated single-cell transcriptomics. *Nat Commun*. 2022;13(1):1584.
48. Espín-Palazón R, Stachura DL, Campbell CA, et al. Proinflammatory Signaling Regulates Hematopoietic Stem Cell Emergence. *Cell*. 2014;159(5):1070–1085.
49. Crosse EI, Gordon-Keylock S, Rybtsov S, et al. Multi-layered Spatial Transcriptomics Identify Secretory Factors Promoting Human Hematopoietic Stem Cell Development. *Cell Stem Cell*. 2020;27(5):822-839.e8.
50. Ramakrishnan R, Peña-Martínez P, Agarwal P, et al. CXCR4 Signaling Has a CXCL12-Independent Essential Role in Murine MLL-AF9-Driven Acute Myeloid Leukemia. *Cell Reports*. 2020;31(8):107684.

51. Hoppe PS, Schwarzfischer M, Loeffler D, et al. Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. *Nature*. 2016;535(7611):299–302.
52. Somasundaram R, Prasad MAJ, Ungerbäck J, Sigvardsson M. Transcription factor networks in B-cell differentiation link development to acute lymphoid leukemia. *Blood*. 2015;126(2):144–152.
53. Rosenbauer F, Tenen DG. Transcription factors in myeloid development: balancing differentiation with transformation. *Nat Rev Immunol*. 2007;7(2):105–117.
54. Liu X, Wang Y, Lu H, et al. Genome-wide analysis identifies NR4A1 as a key mediator of T cell dysfunction. *Nature*. 2019;567(7749):525–529.
55. McCutcheon SR, Swartz AM, Brown MC, et al. Transcriptional and epigenetic regulators of human CD8+ T cell function identified through orthogonal CRISPR screens. *Nat Genet*. 2023;55(12):2211–2223.
56. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, et al. Chromatin state dynamics during blood formation. *Science*. 2014;345(6199):943–949.
57. Yu VWC, Yusuf RZ, Oki T, et al. Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. *Cell*. 2017;168(5):944–945.
58. Dixon JR, Jung I, Selvaraj S, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature*. 2015;518(7539):331–336.
59. Buenrostro JD, Corces MR, Lareau CA, et al. Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. *Cell*. 2018;173(6):1535-1548.e16.
60. Georgolopoulos G, Psatha N, Iwata M, et al. Discrete regulatory modules instruct hematopoietic lineage commitment and differentiation. *Nat Commun*. 2021;12(1):6790.
61. Abdel-Wahab O, Adli M, LaFave LM, et al. ASXL1 Mutations Promote Myeloid Transformation through Loss of PRC2-Mediated Gene Repression. *Cancer Cell*. 2012;22(2):180–193.
62. Lyko F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat Rev Genet*. 2018;19(2):81–92.
63. Jeong M, Park HJ, Celik H, et al. Loss of Dnmt3a Immortalizes Hematopoietic Stem Cells In Vivo. *Cell Reports*. 2018;23(1):1–10.
64. Chan SM, Majeti R. Role of DNMT3A, TET2, and IDH1/2 mutations in pre-leukemic stem cells in acute myeloid leukemia. *Int J Hematol*. 2013;98(6):648–657.
65. Thomas D, Majeti R. Biology and relevance of human acute myeloid leukemia stem cells. *Blood*. 2017;129(12):1577–1585.
66. Geest CR, Coffer PJ. MAPK signaling pathways in the regulation of hematopoiesis. *J Leukoc Biol*. 2009;86(2):237–250.
67. Polak R, Buitenhuis M. The PI3K/PKB signaling module as key regulator of hematopoiesis: implications for therapeutic strategies in leukemia. *Blood*. 2012;119(4):911–923.

68. Buitenhuis M, Verhagen LP, Van Deutekom HWM, et al. Protein kinase B (c-akt) regulates hematopoietic lineage choice decisions during myelopoiesis. *Blood*. 2008;111(1):112–121.
69. Srinivasan L, Sasaki Y, Calado DP, et al. PI3 Kinase Signals BCR-Dependent Mature B Cell Survival. *Cell*. 2009;139(3):573–586.
70. Ramadani F, Bolland DJ, Garcon F, et al. The PI3K Isoforms p110α and p110δ Are Essential for Pre–B Cell Receptor Signaling and B Cell Development. *Sci. Signal.* 2010;3(134):.
71. Arber DA, Orazi A, Hasserjian RP, et al. International Consensus Classification of Myeloid Neoplasms and Acute Leukemias: integrating morphologic, clinical, and genomic data. *Blood*. 2022;140(11):1200–1228.
72. Alaggio R, Amador C, Anagnostopoulos I, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia*. 2022;36(7):1720–1748.
73. Khoury JD, Solary E, Abla O, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia*. 2022;36(7):1703–1719.
74. Pui C-H, Robison LL, Look AT. Acute lymphoblastic leukaemia. *Lancet*. 2008;371(9617):1030–1043.
75. Short NJ, Rytting ME, Cortes JE. Acute myeloid leukaemia. *Lancet*. 2018;392(10147):593–606.
76. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2018, National Cancer Institute. (2021). Available at: www.seer.cancer.gov.
77. Shallis RM, Wang R, Davidoff A, Ma X, Zeidan AM. Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. *Blood Rev*. 2019;36:70–87.
78. Döhner H, Wei AH, Appelbaum FR, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood*. 2022;140(12):1345–1377.
79. Juliusson G, Antunovic P, Derolf Å, et al. Age and acute myeloid leukemia: real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry. *Blood*. 2009;113(18):4179–4187.
80. Juliusson G, Lazarevic V, Hörstedt A-S, et al. Acute myeloid leukemia in the real world: why population-based registries are needed. *Blood*. 2012;119(17):3890–3899.
81. for the Swedish AML Group and the Swedish Childhood Leukemia Group, Juliusson G, Abrahamsson J, et al. Prevalence and characteristics of survivors from acute myeloid leukemia in Sweden. *Leukemia*. 2017;31(3):728–731.
82. Hunger SP, Mullighan CG. Acute Lymphoblastic Leukemia in Children. *N Engl J Med*. 2015;373(16):1541–1552.
83. Steliarova-Foucher E, Colombet M, Ries LAG, et al. International incidence of childhood cancer, 2001–10: a population-based registry study. *Lancet Oncol*. 2017;18(6):719–731.

84. Lennmyr E, Karlsson K, Ahlberg L, et al. Survival in adult acute lymphoblastic leukaemia (ALL): A report from the Swedish ALL Registry. *Eur J Haematol*. 2019;103(2):88–98.
85. Lam CG, Howard SC, Bouffet E, Pritchard-Jones K. Science and health for all children with cancer. *Science*. 2019;363(6432):1182–1186.
86. Bonaventure A, Harewood R, Stiller CA, et al. Worldwide comparison of survival from childhood leukaemia for 1995–2009, by subtype, age, and sex (CONCORD-2): a population-based study of individual data for 89 828 children from 198 registries in 53 countries. *Lancet Haematol*. 2017;4(5):e202–e217.
87. Alexander TB, Wang L, Inaba H, et al. Decreased relapsed rate and treatment-related mortality contribute to improved outcomes for pediatric acute myeloid leukemia in successive clinical trials. *Cancer*. 2017;123(19):3791–3798.
88. Iacobucci I, Mullighan CG. Genetic Basis of Acute Lymphoblastic Leukemia. *J Clin Oncol*. 2017;35(9):975–983.
89. Kantarjian H, Kadia T, DiNardo C, et al. Acute myeloid leukemia: current progress and future directions. *Blood Cancer J.* 2021;11(2):41.
90. Anderson K, Lutz C, Van Delft FW, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature*. 2011;469(7330):356–361.
91. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*. 2012;481(7382):506–510.
92. Welch JS, Ley TJ, Link DC, et al. The Origin and Evolution of Mutations in Acute Myeloid Leukemia. *Cell*. 2012;150(2):264–278.
93. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature*. 2014;506(7488):328–333.
94. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*. 2016;374(23):2209–2221.
95. for The St. Jude Children's Research Hospital–Washington University Pediatric Cancer Genome Project, Andersson AK, Ma J, et al. The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nat Genet*. 2015;47(4):330–337.
96. Ma X, Edmonson M, Yergeau D, et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. *Nat Commun*. 2015;6(1):6604.
97. Morita K, Wang F, Jahn K, et al. Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics. *Nat Commun*. 2020;11(1):5327.
98. Sandén C, Lilljebjörn H, Orsmark Pietras C, et al. Clonal competition within complex evolutionary hierarchies shapes AML over time. *Nat Commun*. 2020;11(1):579.

99. Gu Z, Churchman ML, Roberts KG, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet*. 2019;51(2):296–307.
100. Brady SW, Roberts KG, Gu Z, et al. The genomic landscape of pediatric acute lymphoblastic leukemia. *Nat Genet*. 2022;54(9):1376–1389.
101. Jaiswal S, Fontanillas P, Flannick J, et al. Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N Engl J Med*. 2014;371(26):2488–2498.
102. Genovese G, Kähler AK, Handsaker RE, et al. Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N Engl J Med*. 2014;371(26):2477–2487.
103. Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. *Science*. 2019;366(6465):eaan4673.
104. Fabre MA, De Almeida JG, Fiorillo E, et al. The longitudinal dynamics and natural history of clonal haematopoiesis. *Nature*. 2022;606(7913):335–342.
105. Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*. 2014;20(12):1472–1478.
106. Mori H, Colman SM, Xiao Z, et al. Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc Natl Acad Sci U.S.A*. 2002;99(12):8242–8247.
107. Schäfer D, Olsen M, Lähnemann D, et al. Five percent of healthy newborns have an ETV6-RUNX1 fusion as revealed by DNA-based GIPFEL screening. *Blood*. 2018;131(7):821–826.
108. Hong D, Gupta R, Ancliff P, et al. Initiating and Cancer-Propagating Cells in *TEL-AML1* -Associated Childhood Leukemia. *Science*. 2008;319(5861):336–339.
109. Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat Genet*. 2014;46(2):116–125.
110. Zink F, Stacey SN, Norddahl GL, et al. Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood*. 2017;130(6):742–752.
111. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2015;126(1):9–16.
112. Mitchell SR, Gopakumar J, Jaiswal S. Insights into clonal hematopoiesis and its relation to cancer risk. *Curr Opin Genet Dev*. 2021;66:63–69.
113. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat Commun*. 2016;7(1):12484.
114. Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. *Nature*. 2013;502(7471):333–339.
115. The Cancer Genome Atlas Research Network. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *N Engl J Med*. 2013;368(22):2059–2074.

116. Tyner JW, Tognon CE, Bottomly D, et al. Functional genomic landscape of acute myeloid leukaemia. *Nature*. 2018;562(7728):526–531.
117. Bolouri H, Farrar JE, Triche T, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nat Med*. 2018;24(1):103–112.
118. Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415–421.
119. Thatikonda V, Islam SMA, Autry RJ, et al. Comprehensive analysis of mutational signatures reveals distinct patterns and molecular processes across 27 pediatric cancers. *Nat Cancer*. 2023;4(2):276–289.
120. Ley TJ, Mardis ER, Ding L, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature*. 2008;456(7218):66–72.
121. Umeda M, Ma J, Westover T, et al. A new genomic framework to categorize pediatric acute myeloid leukemia. *Nat Genet*. 2024;56(2):281–293.
122. WHO Classification of Tumours Editorial Board. Haematolymphoid tumours [Internet]. Lyon (France): International Agency for Research on Cancer; 2024 [cited 2025 February 27]. (WHO classification of tumours series, 5th ed.; vol. 11). Available from: https://tumourclassification.iarc.who.int/chapters/63.
123. Ma X, Liu Y, Liu Y, et al. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature*. 2018;555(7696):371–376.
124. Barajas JM, Umeda M, Contreras L, et al. UBTF tandem duplications in pediatric myelodysplastic syndrome and acute myeloid leukemia: implications for clinical screening and diagnosis. *Haematologica*. 2024;109(8):2459–2468.
125. Bennett JM, Catovsky D, Daniel MT, et al. Proposed Revised Criteria for the Classification of Acute Myeloid Leukemia: A Report of the French-American-British Cooperative Group. *Ann Intern Med*. 1985;103(4):620–625.
126. Wouters BJ, Löwenberg B, Erpelinck-Verschueren CAJ, et al. Double CEBPA mutations, but not single CEBPA mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood*. 2009;113(13):3088–3091.
127. Hernández-Sánchez A, González T, Sobas M, et al. Rearrangements involving 11q23.3/KMT2A in adult AML: mutational landscape and prognostic implications – a HARMONY study. *Leukemia*. 2024;38(9):1929–1937.
128. Tian J, Zhu Y, Li J, et al. The landscape of NUP98 rearrangements clinical characteristics and treatment response from 1491 acute leukemia patients. *Blood Cancer J.* 2024;14(1):81.
129. Rubnitz JE, Lacayo NJ, Inaba H, et al. Clofarabine Can Replace Anthracyclines and Etoposide in Remission Induction Therapy for Childhood Acute Myeloid Leukemia: The AML08 Multicenter, Randomized Phase III Trial. *J Clin Oncol*. 2019;37(23):2072–2081.

130. Pollard JA, Alonzo TA, Gerbing R, et al. Sorafenib in Combination With Standard Chemotherapy for Children With High Allelic Ratio *FLT3* /ITD+ Acute Myeloid Leukemia: A Report From the Children's Oncology Group Protocol AAML1031. *J Clin Oncol*. 2022;40(18):2023–2035.
131. Reinhardt D, Antoniou E, Waack K. Pediatric Acute Myeloid Leukemia—Past, Present, and Future. *J Clin Med*. 2022;11(3):504.
132. Tomizawa D, Tsujimoto S, Tanaka S, et al. A phase III clinical trial evaluating efficacy and safety of minimal residual disease-based risk stratification for children with acute myeloid leukemia, incorporating a randomized study of gemtuzumab ozogamicin in combination with post-induction chemotherapy for non-low-risk patients (JPLSG-AML-20). *Jpn J Clin Oncol*. 2022;hyac105.
133. Jeha S, Choi J, Roberts KG, et al. Clinical Significance of Novel Subtypes of Acute Lymphoblastic Leukemia in the Context of Minimal Residual Disease–Directed Therapy. *Blood Cancer Discov*. 2021;2(4):326–337.
134. Bergfelt Lennmyr E, Engvall M, Barbany G, et al. Cytogenetic aberrations in adult acute lymphoblastic leukemia-A population-based study. *EJHaem*. 2021;2(4):813–817.
135. Iacobucci I, Kimura S, Mullighan CG. Biologic and Therapeutic Implications of Genomic Alterations in Acute Lymphoblastic Leukemia. *J Clin Med*. 2021;10(17):3792.
136. Visvader JE. Cells of origin in cancer. *Nature*. 2011;469(7330):314–322.
137. Lapidot T, Sirard C, Vormoor J, et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*. 1994;367(6464):645–648.
138. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med*. 1997;3(7):730–737.
139. Taussig DC, Vargaftig J, Miraki-Moud F, et al. Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34− fraction. *Blood*. 2010;115(10):1976–1984.
140. Martelli MP, Pettirossi V, Thiede C, et al. CD34+ cells from AML with mutated NPM1 harbor cytoplasmic mutated nucleophosmin and generate leukemia in immunocompromised mice. *Blood*. 2010;116(19):3907–3922.
141. Ries RE, Eidenschink Brodersen L, Pardo L, et al. Epigenetic Silencing of CD34 in AML and Association with Outcome in KMT2A Fusions. *Blood*. 2021;138(Supplement 1):802–802.
142. Shultz LD, Ishikawa F, Greiner DL. Humanized mice in translational biomedical research. *Nat Rev Immunol*. 2007;7(2):118–130.
143. Wunderlich M, Chou F-S, Link KA, et al. AML xenograft efficiency is significantly improved in NOD/SCID-IL2RG mice constitutively expressing human SCF, GM-CSF and IL-3. *Leukemia*. 2010;24(10):1785–1788.
144. Goyama S, Wunderlich M, Mulloy JC. Xenograft models for normal and malignant stem cells. *Blood*. 2015;125(17):2630–2640.

145. Le Viseur C, Hotfilder M, Bomken S, et al. In Childhood Acute Lymphoblastic Leukemia, Blasts at Different Stages of Immunophenotypic Maturation Have Stem Cell Properties. *Cancer Cell*. 2008;14(1):47–58.
146. Testa U. Leukemic Stem Cells in Acute Lymphoblastic Leukemia. *Cancer Stem Cells: Emerging Concepts and Future Perspectives in Translational Oncology*. 2015;333–389.
147. Castor A, Nilsson L, Åstrand-Grundström I, et al. Distinct patterns of hematopoietic stem cell involvement in acute lymphoblastic leukemia. *Nat Med*. 2005;11(6):630–637.
148. Rehe K, Wilson K, Bomken S, et al. Acute B lymphoblastic leukaemia-propagating cells are present at high frequency in diverse lymphoblast populations. *EMBO Mol Med*. 2013;5(1):38–51.
149. Goardon N, Marchi E, Atzberger A, et al. Coexistence of LMPP-like and GMP-like Leukemia Stem Cells in Acute Myeloid Leukemia. *Cancer Cell*. 2011;19(1):138–152.
150. Pang WW, Price EA, Sahoo D, et al. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci U.S.A*. 2011;108(50):20012–20017.
151. Marchand T, Pinho S. Leukemic Stem Cells: From Leukemic Niche Biology to Treatment Opportunities. *Front. Immunol.* 2021;12:775128.
152. Haubner S, Perna F, Köhnke T, et al. Coexpression profile of leukemic stem cell markers for combinatorial targeted therapy in AML. *Leukemia*. 2019;33(1):64–74.
153. von Palffy S, Thorsson H, Peña-Martínez P, et al. The complement receptor C3AR constitutes a novel therapeutic target in *NPM1*-mutated AML. *Blood Adv*. 2023;7(7):1204–1218.
154. Theunissen P, Mejstrikova E, Sedek L, et al. Standardized flow cytometry for highly sensitive MRD measurements in B-cell acute lymphoblastic leukemia. *Blood*. 2017;129(3):347–357.
155. Modvig S, Hallböök H, Madsen HO, et al. Value of flow cytometry for MRD-based relapse prediction in B-cell precursor ALL in a multicenter setting. *Leukemia*. 2021;35(7):1894–1906.
156. Angelova E, Audette C, Kovtun Y, et al. CD123 expression patterns and selective targeting with a CD123-targeted antibody-drug conjugate (IMGN632) in acute lymphoblastic leukemia. *Haematologica*. 2019;104(4):749–755.
157. Schinnerl D, Mejstrikova E, Schumich A, et al. CD371 cell surface expression: a unique feature of *DUX4* -rearranged acute lymphoblastic leukemia. *Haematologica*. 2019;104(8):e352–e355.
158. van Rhenen A, van Dongen GAMS, Kelder A, et al. The novel AML stem cell–associated antigen CLL-1 aids in discrimination between normal and leukemic stem cells. *Blood*. 2007;110(7):2659–2666.
159. Wang J, Chen S, Xiao W, et al. CAR-T cells targeting CLL-1 as an approach to treat acute myeloid leukemia. *J Hematol Oncol*. 2018;11(1):7.

160. Gentles AJ. Association of a Leukemic Stem Cell Gene Expression Signature With Clinical Outcomes in Acute Myeloid Leukemia. *JAMA*. 2010;304(24):2706.
161. Huang BJ, Smith JL, Farrar JE, et al. Integrated stem cell signature and cytomolecular risk determination in pediatric acute myeloid leukemia. *Nat Commun*. 2022;13(1):5487.
162. Ng SWK, Mitchell A, Kennedy JA, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*. 2016;540(7633):433–437.
163. Docking TR, Parker JDK, Jädersten M, et al. A clinical transcriptome approach to patient stratification and therapy selection in acute myeloid leukemia. *Nat Commun*. 2021;12(1):2474.
164. Wagner S, Vadakekolathu J, Tasian SK, et al. A parsimonious 3-gene signature predicts clinical outcomes in an acute myeloid leukemia multicohort study. *Blood Advances*. 2019;3(8):1330–1346.
165. Nehme A, Dakik H, Picou F, et al. Horizontal meta-analysis identifies common deregulated genes across AML subgroups providing a robust prognostic signature. *Blood Adv*. 2020;4(20):5322–5335.
166. Elsayed AH, Rafiee R, Cao X, et al. A six-gene leukemic stem cell score identifies high risk pediatric acute myeloid leukemia. *Leukemia*. 2020;34(3):735–745.
167. Eppert K, Takenaka K, Lechman ER, et al. Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat Med*. 2011;17(9):1086–1093.
168. Bill M, Nicolet D, Kohlschmidt J, et al. Mutations associated with a 17-gene leukemia stem cell score and the score's prognostic relevance in the context of the European LeukemiaNet classification of acute myeloid leukemia. *Haematologica*. 2020;105(3):721–729.
169. Vasseur L, Fenwarth L, Lambert J, et al. LSC17 score complements genetics and measurable residual disease in acute myeloid leukemia: an ALFA study. *Blood Adv*. 2023;7(15):4024–4034.
170. Zeng AGX, Bansal S, Jin L, et al. A cellular hierarchy framework for understanding heterogeneity and predicting drug response in acute myeloid leukemia. *Nat Med*. 2022;28(6):1212–1223.
171. Bottomly D, Long N, Schultz AR, et al. Integrative analysis of drug response and clinical outcome in acute myeloid leukemia. *Cancer Cell*. 2022;40(8):850-864.e9.
172. Zhang X, Song B, Carlino MJ, et al. An immunophenotype-coupled transcriptomic atlas of human hematopoietic progenitors. *Nat Immunol*. 2024;25(4):703–715.
173. Pei S, Pollyea DA, Gustafson A, et al. Monocytic Subclones Confer Resistance to Venetoclax-Based Therapy in Patients with Acute Myeloid Leukemia. *Cancer Discov*. 2020;10(4):536–551.
174. Pei S, Shelton IT, Gillen AE, et al. A Novel Type of Monocytic Leukemia Stem Cell Revealed by the Clinical Use of Venetoclax-Based Therapy. *Cancer Discov*. 2023;13(9):2032–2049.

175. Lambo S, Trinh DL, Ries RE, et al. A longitudinal single-cell atlas of treatment response in pediatric AML. *Cancer Cell*. 2023;41(12):2117-2135.e12.
176. Chen C, Yu W, Alikarami F, et al. Single-cell multiomics reveals increased plasticity, resistant populations, and stem-cell–like blasts in *KMT2A* -rearranged leukemia. *Blood*. 2022;139(14):2198–2211.
177. Agrawal-Singh S, Bagri J, Sakakini N, Huntly BJP. A guide to epigenetics in leukaemia stem cells. *Mol Oncol*. 2023;17(12):2493–2506.
178. Tabe Y, Konopleva M. Leukemia Stem Cells Microenvironment. *Stem Cell Microenvironments and Beyond*. 2017;1041:19–32.
179. Kellaway SG, Potluri S, Keane P, et al. Leukemic stem cells activate lineage inappropriate signalling pathways to promote their growth. *Nat Commun*. 2024;15(1):1359.
180. Thorsson H, Henningsson R, Puente-Moncada N, et al. Single-cell genomics details the maturation block in BCP-ALL and identifies therapeutic vulnerabilities in *DUX4*-r cases. *Blood*. 2024;144(13):1399–1411.
181. Risau W. Mechanisms of angiogenesis. *Nature*. 1997;386(6626):671–674.
182. Ferrara N, Kerbel RS. Angiogenesis as a therapeutic target. *Nature*. 2005;438(7070):967–974.
183. Wang Z, Xu Q, Zhang N, et al. CD146, from a melanoma cell adhesion molecule to a signaling receptor. *Sig Transduct Target Ther*. 2020;5(1):148.
184. Witkowski MT, Dolgalev I, Evensen NA, et al. Extensive Remodeling of the Immune Microenvironment in B Cell Acute Lymphoblastic Leukemia. *Cancer Cell*. 2020;37(6):867-882.e12.
185. Lyu A, Nam SH, Humphrey RS, Horton TM, Ehrlich LIR. Cells and signals of the leukemic microenvironment that support progression of T-cell acute lymphoblastic leukemia (T-ALL). *Exp Mol Med*. 2024;56(11):2337–2347.
186. Mumme H, Thomas BE, Bhasin SS, et al. Single-cell analysis reveals altered tumor microenvironments of relapse- and remission-associated pediatric acute myeloid leukemia. *Nat Commun*. 2023;14(1):6209.
187. Lasry A, Nadorp B, Fornerod M, et al. An inflammatory state remodels the immune microenvironment and improves risk stratification in acute myeloid leukemia. *Nat Cancer*. (2023;4(1):27–42):.
188. Wang X, Chen Y, Li Z, et al. Single-Cell RNA-Seq of T Cells in B-ALL Patients Reveals an Exhausted Subset with Remarkable Heterogeneity. *Adv. Sci.* 2021;8(19):2101447.
189. Zhang Z, Deng C, Zhu P, et al. Single-cell RNA-seq reveals a microenvironment and an exhaustion state of T/NK cells in acute myeloid leukemia. *Cancer Science*. 2023;114(10):3873–3883.
190. Wu X, Wu Z, Deng W, et al. Spatiotemporal evolution of AML immune microenvironment remodeling and RNF149-driven drug resistance through single-cell multidimensional analysis. *J Transl Med*. 2023;21(1):760.
191. Zanetti C, Kumar R, Ender J, et al. The age of the bone marrow microenvironment influences B-cell acute lymphoblastic leukemia progression via CXCR5-CXCL13. *Blood*. 2021;138(19):1870–1884.

192. Lee G-Y, Jeong S-Y, Lee H-R, Oh I-H. Age-related differences in the bone marrow stem cell niche generate specialized microenvironments for the distinct regulation of normal hematopoietic and leukemia stem cells. *Sci Rep*. 2019;9(1):1007.
193. Gooley TA, Chien JW, Pergam SA, et al. Reduced mortality after allogeneic hematopoietic-cell transplantation. *N Engl J Med*. 2010;363(22):2091–2101.
194. Lo-Coco F, Avvisati G, Vignetti M, et al. Retinoic Acid and Arsenic Trioxide for Acute Promyelocytic Leukemia. *N Engl J Med*. 2013;369(2):111–121.
195. Burnett AK, Russell NH, Hills RK, et al. Arsenic trioxide and all-trans retinoic acid treatment for acute promyelocytic leukaemia in all risk groups (AML17): results of a randomised, controlled, phase 3 trial. *Lancet Oncol*. 2015;16(13):1295–1305.
196. Cohen MH, Williams G, Johnson JR, et al. Approval summary for imatinib mesylate capsules in the treatment of chronic myelogenous leukemia. *Clin Cancer Res*. 2002;8(5):935–942.
197. Druker BJ, Tamura S, Buchdunger E, et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr–Abl positive cells. *Nat Med*. 1996;2(5):561–566.
198. Abedin S, Altman JK. Acute promyelocytic leukemia: preventing early complications and late toxicities. *Hematology*. 2016;2016(1):10–15.
199. Ravandi F, O'Brien S, Thomas D, et al. First report of phase 2 study of dasatinib with hyper-CVAD for the frontline treatment of patients with Philadelphia chromosome–positive (Ph+) acute lymphoblastic leukemia. *Blood*. 2010;116(12):2070–2077.
200. Chalandon Y, Thomas X, Hayette S, et al. Randomized study of reduced-intensity chemotherapy combined with imatinib in adults with Ph-positive acute lymphoblastic leukemia. *Blood*. 2015;125(24):3711–3719.
201. Lachowiez CA, DiNardo CD, Loghavi S. Molecularly Targeted Therapy in Acute Myeloid Leukemia: Current Treatment Landscape and Mechanisms of Response and Resistance. *Cancers*. 2023;15(5):1617.
202. Pollyea DA, Amaya M, Strati P, Konopleva MY. Venetoclax for AML: changing the treatment paradigm. *Blood Adv*. 2019;3(24):4326–4335.
203. Aldoss I, Issa GC, Blachly JS, et al. Updated Results and Longer Follow-up from the AUGMENT-101 Phase 2 Study of Revumenib in All Patients with Relapsed or Refractory (R/R) *KMT2Ar* Acute Leukemia. *Blood*. 2024;144(Supplement 1):211–211.
204. Egan G, Tasian SK. Relapsed pediatric acute myeloid leukaemia: state-of-the-art in 2023. *Haematol*. (2023;108(9):2275–2288):.
205. Thomas ED, Lochte HL, Lu WC, Ferrebee JW. Intravenous Infusion of Bone Marrow in Patients Receiving Radiation and Chemotherapy. *N Engl J Med*. 1957;257(11):491–496.
206. Tettamanti S, Pievani A, Biondi A, Dotti G, Serafini M. Catch me if you can: how AML and its niche escape immunotherapy. *Leukemia*. 2022;36(1):13–22.
207. Schürch CM. Therapeutic Antibodies for Myeloid Neoplasms—Current Developments and Future Directions. *Front Oncol*. 2018;8:152.

208. Goulet DR, Atkins WM. Considerations for the Design of Antibody-Based Therapeutics. *J Pharm Sci*. 2020;109(1):74–103.
209. Gül N, Van Egmond M. Antibody-Dependent Phagocytosis of Tumor Cells by Macrophages: A Potent Effector Mechanism of Monoclonal Antibody Therapy of Cancer. *Cancer Res*. 2015;75(23):5008–5013.
210. Minard-Colin V, Aupérin A, Pillon M, et al. Rituximab for High-Risk, Mature B-Cell Non-Hodgkin's Lymphoma in Children. *N Engl J Med*. 2020;382(23):2207–2219.
211. Dimopoulos MA, Oriol A, Nahi H, et al. Daratumumab, Lenalidomide, and Dexamethasone for Multiple Myeloma. *N Engl J Med*. 2016;375(14):1319–1331.
212. Järås M, Johnels P, Hansen N, et al. Isolation and killing of candidate chronic myeloid leukemia stem cells by antibody targeting of IL-1 receptor accessory protein. *Proc Natl Acad Sci U.S.A*. 2010;107(37):16280–16285.
213. Ågerstam H, Karlsson C, Hansen N, et al. Antibodies targeting human IL1RAP (IL1R3) show therapeutic effects in xenograft models of acute myeloid leukemia. *Proc Natl Acad Sci U.S.A*. 2015;112(34):10786–10791.
214. Marin-Acevedo JA, Kimbrough EO, Lou Y. Next generation of immune checkpoint inhibitors and beyond. *J Hematol Oncol*. 2021;14(1):45.
215. Kroll MH, Rojas-Hernandez C, Yee C. Hematologic complications of immune checkpoint inhibitors. *Blood*. 2022;139(25):3594–3604.
216. Goebeler M-E, Bargou RC. T cell-engaging therapies — BiTEs and beyond. *Nat Rev Clin Oncol*. 2020;17(7):418–434.
217. Kantarjian H, Stein A, Gökbuget N, et al. Blinatumomab versus Chemotherapy for Advanced Acute Lymphoblastic Leukemia. *N Engl J Med*. 2017;376(9):836–847.
218. Kantarjian HM, DeAngelo DJ, Stelljes M, et al. Inotuzumab Ozogamicin versus Standard Therapy for Acute Lymphoblastic Leukemia. *N Engl J Med*. 2016;375(8):740–753.
219. Lambert J, Pautas C, Terré C, et al. Gemtuzumab ozogamicin for *de novo* acute myeloid leukemia: final efficacy and safety updates from the open-label, phase III ALFA-0701 trial. *Haematologica*. 2019;104(1):113–119.
220. Baker DJ, Arany Z, Baur JA, Epstein JA, June CH. CAR T therapy beyond cancer: the evolution of a living drug. *Nature*. 2023;619(7971):707–715.
221. Maude SL, Frey N, Shaw PA, et al. Chimeric antigen receptor T cells for sustained remissions in leukemia. *N Engl J Med*. 2014;371(16):1507–1517.
222. Kochenderfer JN, Dudley ME, Feldman SA, et al. B-cell depletion and remissions of malignancy along with cytokine-associated toxicity in a clinical trial of anti-CD19 chimeric-antigen-receptor–transduced T cells. *Blood*. 2012;119(12):2709–2720.
223. Rabilloud T, Potier D, Pankaew S, et al. Single-cell profiling identifies pre-existing CD19-negative subclones in a B-ALL patient with CD19-negative relapse after CAR-T therapy. *Nat Commun*. 2021;12(1):865.
224. Mardiros A, Dos Santos C, McDonald T, et al. T cells expressing CD123-specific chimeric antigen receptors exhibit specific cytolytic effector functions

and antitumor effects against human acute myeloid leukemia. *Blood*. 2013;122(18):3138–3148.
225. Ehninger A, Kramer M, Röllig C, et al. Distribution and levels of cell surface expression of CD33 and CD123 in acute myeloid leukemia. *Blood Cancer J*. 2014;4(6):e218–e218.
226. Kenderian SS, Ruella M, Shestova O, et al. CD33-specific chimeric antigen receptor T cells exhibit potent preclinical activity against human acute myeloid leukemia. *Leukemia*. 2015;29(8):1637–1647.
227. Jetani H, Garcia-Cadenas I, Nerreter T, et al. CAR T-cells targeting FLT3 have potent activity against FLT3−ITD+ AML and act synergistically with the FLT3-inhibitor crenolanib. *Leukemia*. 2018;32(5):1168–1179.
228. Myburgh R, Kiefer JD, Russkamp NF, et al. Anti-human CD117 CAR T-cells efficiently eliminate healthy and malignant CD117-expressing hematopoietic cells. *Leukemia*. 2020;34(10):2688–2703.
229. Jin X, Zhang M, Sun R, et al. First-in-human phase I study of CLL-1 CAR-T cells in adults with relapsed/refractory acute myeloid leukemia. *J Hematol Oncol*. 2022;15(1):88.
230. MacKay M, Afshinnekoo E, Rub J, et al. The therapeutic landscape for cells engineered with chimeric antigen receptors. *Nat Biotechnol*. 2020;38(2):233–244.
231. Haubner S, Mansilla-Soto J, Nataraj S, et al. Cooperative CAR targeting to selectively eliminate AML and minimize escape. *Cancer Cell*. 2023;41(11):1871-1891.e6.
232. Casirati G, Cosentino A, Mucci A, et al. Epitope editing enables targeted immunotherapy of acute myeloid leukaemia. *Nature*. 2023;621(7978):404–414.
233. Rafei H, Daher M, Rezvani K. Chimeric antigen receptor (CAR) natural killer (NK)-cell therapy: leveraging the power of innate immunity. *Br J Haematol*. 2021;193(2):216–230.
234. Klebanoff CA, Chandran SS, Baker BM, Quezada SA, Ribas A. T cell receptor therapeutics: immunological targeting of the intracellular cancer proteome. *Nat Rev Drug Discov*. 2023;22(12):996–1017.
235. Ali M, Giannakopoulou E, Li Y, et al. T cells targeted to TdT kill leukemic lymphoblasts while sparing normal lymphocytes. *Nat Biotechnol*. 2022;40(4):488–498.
236. Giannakopoulou E, Lehander M, Virding Culleton S, et al. A T cell receptor targeting a recurrent driver mutation in FLT3 mediates elimination of primary human acute myeloid leukemia in vivo. *Nat Cancer*. 2023;4(10):1474–1490.
237. Marin D, Li Y, Basar R, et al. Safety, efficacy and determinants of response of allogeneic CD19-specific CAR-NK cells in CD19+ B cell tumors: a phase 1/2 trial. *Nat Med*. 2024;30(3):772–784.
238. Li Y, Basar R, Wang G, et al. Author Correction: KIR-based inhibitory CARs overcome CAR-NK cell trogocytosis-mediated fratricide and tumor escape. *Nat Med*. 2024;30(3):906.

239. Melo Garcia L, Gangadharan A, Banerjee P, et al. Overcoming CD226-related immune evasion in acute myeloid leukemia with CD38 CAR-engineered NK cells. *Cell Rep*. 2025;44(1):115122.
240. Picelli S, Faridani OR, Björklund ÅK, et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*. 2014;9(1):171–181.
241. Hagemann-Jensen M, Ziegenhain C, Chen P, et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat Biotechnol*. 2020;38(6):708–714.
242. Islam S, Kjällquist U, Moliner A, et al. Highly multiplexed and strand-specific single-cell RNA 5′ end sequencing. *Nat Protoc*. 2012;7(5):813–828.
243. Fan HC, Fu GK, Fodor SPA. Combinatorial labeling of single cells for gene expression cytometry. *Science*. 2015;347(6222):1258367.
244. Klein AM, Mazutis L, Akartuna I, et al. Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. *Cell*. 2015;161(5):1187–1201.
245. Macosko EZ, Basu A, Satija R, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*. 2015;161(5):1202–1214.
246. Zheng GXY, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*. 2017;8(1):14049.
247. Hashimshony T, Senderovich N, Avital G, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol*. 2016;17(1):77.
248. Sasagawa Y, Danno H, Takada H, et al. Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. *Genome Biol*. 2018;19(1):29.
249. Jaitin DA, Kenigsberg E, Keren-Shaul H, et al. Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types. *Science*. 2014;343(6172):776–779.
250. Gierahn TM, Wadsworth MH, Hughes TK, et al. Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods*. 2017;14(4):395–398.
251. Hayashi T, Ozaki H, Sasagawa Y, et al. Single-cell full-length total RNA sequencing uncovers dynamics of recursive splicing and enhancer RNAs. *Nat Commun*. 2018;9(1):619.
252. Hahaut V, Pavlinic D, Carbone W, et al. Fast and highly sensitive full-length single-cell RNA sequencing using FLASH-seq. *Nat Biotechnol*. 2022;40(10):1447–1451.
253. 10x Genomics. Chromium GEM-X Single Cell 3’ v4 Gene Expression User Guide. (2024). Available at: www.10xgenomics.com.
254. Stoeckius M, Zheng S, Houck-Loomis B, et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol*. 2018;19(1):224.
255. 10x Genomics. Chromium Next GEM Single Cell 3’ Reagent Kits v3.1 with CellPlex User Guide. (2021). Available at: www.10xgenomics.com. (2021).

256. Pfisterer U, Bräunig J, Brattås P, et al. Single-cell sequencing in translational cancer research and challenges to meet clinical diagnostic needs. *Genes Chromosomes Cancer*. 2021;60(7):504–524.
257. 10x Genomics. Chromium Next GEM Single Cell 3' Reagent Kits v3.1 User Guide. (2020). Available at: www.10xgenomics.com.
258. Singh M, Al-Eryani G, Carswell S, et al. High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat Commun*. 2019;10(1):3120.
259. 10x Genomics. Chromium Next GEM Single Cell V(D)J Reagent Kits v1.1 User Guide. (2021). Available at: www.10xgenomics.com.
260. Heumos L, Schaar AC, Lance C, et al. Best practices for single-cell analysis across modalities. *Nat Rev Genet*. 2023;24(8):550–572.
261. Lafzi A, Moutinho C, Picelli S, Heyn H. Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies. *Nat Protoc*. 2018;13(12):2742–2757.
262. Zappia L, Theis FJ. Over 1000 tools reveal trends in the single-cell RNA-seq analysis landscape. *Genome Biol*. 2021;22(1):301.
263. Amezquita RA, Lun ATL, Becht E, et al. Orchestrating single-cell analysis with Bioconductor. *Nat Methods*. 2020;17(2):137–145.
264. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-3587.e29.
265. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*. 2018;19(1):15.
266. van der Maaten, L. and G. Hinton, *Visualizing Data using t-SNE*. Journal of Machine Learning Research, 2008. 9(11): p. 2579-2605.
267. Becht E, McInnes L, Healy J, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2019;37(1):38–44.
268. Weinreb C, Wolock S, Klein AM. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. *Bioinformatics*. 2018;34(7):1246–1248.
269. Labib M, Kelley SO. Single-cell analysis targeting the proteome. *Nat Rev Chem*. 2020;4(3):143–158.
270. Hulett HR, Bonner WA, Barrett J, Herzenberg LA. Cell Sorting: Automated Separation of Mammalian Cells as a Function of Intracellular Fluorescence. *Science*. 1969;166(3906):747–749.
271. Stoeckius M, Hafemeister C, Stephenson W, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods*. 2017;14(9):865–868.
272. 10x Genomics. Chromium Single Cell 3' Reagent Kits User Guide (v3.1 Chemistry) with Feature Barcoding Technology for Cell Surface Protein. (2022). Available at: www.10xgenomics.com.
273. BioLegend. TotalSeq™-A Antibody Oligos for CITE-seq and Single-Cell Proteogenomics (Dual Indexing Protocol) (2022). Available at: https://www.biolegend.com/en-us/protocols/totalseq-a-dual-index-protocol.

274. Bandura DR, Baranov VI, Ornatsky OI, et al. Mass Cytometry: Technique for Real Time Single Cell Multitarget Immunoassay Based on Inductively Coupled Plasma Time-of-Flight Mass Spectrometry. *Anal Chem*. 2009;81(16):6813–6822.
275. Cong Y, Motamedchaboki K, Misal SA, et al. Ultrasensitive single-cell proteomics workflow identifies >1000 protein groups per mammalian cell. *Chem Sci*. 2021;12(3):1001–1006.
276. Schoof EM, Furtwängler B, Üresin N, et al. Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat Commun*. 2021;12(1):3341.
277. Specht H, Emmott E, Petelski AA, et al. Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biol*. 2021;22(1):50.
278. Zhang H, Delafield DG, Li L. Mass spectrometry imaging: the rise of spatially resolved single-cell omics. *Nat Methods*. 2023;20(3):327–330.
279. Hou Y, Song L, Zhu P, et al. Single-Cell Exome Sequencing and Monoclonal Evolution of a JAK2-Negative Myeloproliferative Neoplasm. *Cell*. 2012;148(5):873–885.
280. Zong C, Lu S, Chapman AR, Xie XS. Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. *Science*. 2012;338(6114):1622–1626.
281. Huang L, Ma F, Chapman A, Lu S, Xie XS. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. *Annu Rev Genom Hum Genet*. 2015;16(1):79–102.
282. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011;472(7341):90–94.
283. Wang J, Fan HC, Behr B, Quake SR. Genome-wide Single-Cell Analysis of Recombination Activity and De Novo Mutation Rates in Human Sperm. *Cell*. 2012;150(2):402–412.
284. Vitak SA, Torkenczy KA, Rosenkrantz JL, et al. Sequencing thousands of single-cell genomes with combinatorial indexing. *Nat Methods*. 2017;14(3):302–308.
285. Laks E, McPherson A, Zahn H, et al. Clonal Decomposition and DNA Replication States Defined by Scaled Single-Cell Genome Sequencing. *Cell*. 2019;179(5):1207-1221.e22.
286. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013;10(12):1213–1218.
287. Guo H, Zhu P, Guo F, et al. Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat Protoc*. 2015;10(5):645–659.
288. Rotem A, Ram O, Shoresh N, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol*. 2015;33(11):1165–1172.

289. Kaya-Okur HS, Wu SJ, Codomo CA, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun*. 2019;10(1):1930.
290. Pellegrino M, Sciambi A, Treusch S, et al. High-throughput single-cell DNA sequencing of acute myeloid leukemia tumors with droplet microfluidics. *Genome Res.* 2018;28(9):1345–1352.
291. Dillon LW, Ghannam J, Nosiri C, et al. Personalized Single-Cell Proteogenomics to Distinguish Acute Myeloid Leukemia from Nonmalignant Clonal Hematopoiesis. *Blood Cancer Discov*. 2021;2(4):319–325.
292. Rodriguez-Meira A, Buck G, Clark S-A, et al. Unravelling Intratumoral Heterogeneity through High-Sensitivity Single-Cell Mutational Analysis and Parallel RNA Sequencing. *Mol Cell*. 2019;73(6):1292-1305.e8.
293. Dey SS, Kester L, Spanjaard B, Bienko M, Van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol*. 2015;33(3):285–289.
294. Macaulay IC, Haerty W, Kumar P, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods*. 2015;12(6):519–522.
295. Han KY, Kim K-T, Joung J-G, et al. SIDR: simultaneous isolation and parallel sequencing of genomic DNA and total RNA from single cells. *Genome Res.* 2018;28(1):75–87.
296. Jaitin DA, Weiner A, Yofe I, et al. Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell*. 2016;167(7):1883-1896.e15.
297. Dixit A, Parnas O, Li B, et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell*. 2016;167(7):1853-1866.e17.
298. Datlinger P, Rendeiro AF, Schmidl C, et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods*. 2017;14(3):297–301.
299. Ricklin D, Reis ES, Mastellos DC, Gros P, Lambris JD. Complement component C3 – The "Swiss Army Knife" of innate immunity and host defense. *Immunol Rev*. 2016;274(1):33–58.
300. Wang Y, Zhang H, He Y-W. The Complement Receptors C3aR and C5aR Are a New Class of Immune Checkpoint Receptor in Cancer Immunotherapy. *Front Immunol*. 2019;10:1574.
301. Ajona D, Ortiz-Espinosa S, Pio R. Complement anaphylatoxins C3a and C5a: Emerging roles in cancer progression and treatment. *Semin Cell Dev Biol*. 2019;85:153–163.
302. Barratt J, Weitz I. Complement Factor D as a Strategic Target for Regulating the Alternative Complement Pathway. *Front Immunol*. 2021;12:712572.

About the author

This thesis is my contribution to the fight against leukemia. Hopefully, it adds a small but valuable piece to the puzzle of understanding and ultimately stopping this complex disease. However, the battle does not end here and we all have the power to make a difference. One way is by registering as a bone marrow donor. To learn more, visit Tobias Registret.

I am an engineer with a passion for science and life. When I am not deep into research, you will find me coaching at the gym. Alongside my PhD, I have also completed a "Svensk Klassiker". Why not add a little endurance to the mix?

Printed by Media-Tryck, Lund 2025 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

FACULTY OF MEDICINE

Division of Clinical Genetics
Department of Laboratory Medicine, Lund

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2025:42
ISBN 978-91-8021-695-1
ISSN 1652-8220