Functional Screens Identify Vulnerabilities in Acute Leukemia

Ramakrishnan, Ramprasad

2020

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

Link to publication

Citation for published version (APA):
Ramakrishnan, R. (2020). Functional Screens Identify Vulnerabilities in Acute Leukemia. Lund University, Faculty of Medicine.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Functional screens identify vulnerabilities in acute leukemia

Ramprasad Ramakrishnan

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended on Friday, 18th of September 2020 at 13.00 in
Segerfalksalen, BMC A10, Sölvegatan 19, Lund

Main supervisor
Marcus Järås, PhD
Associate professor, Lund University, Lund, Sweden

Faculty opponent
Julian Walfridsson, PhD
Assistant professor, Karolinska Institute, Stockholm, Sweden
Abstract
Acute leukemia refers to a group of aggressive hematological malignancies of myeloid and lymphoid lineages termed acute myeloid leukemia (AML) and acute lymphoid leukemia (ALL) respectively. Acute leukemia is characterized by the presence of underlying genetic aberrations which alter the biology of normal hematopoietic cells resulting in the accumulation of immature abnormally differentiated blast cells. In this thesis, we have used advanced molecular techniques to identify vulnerabilities in acute leukemia.

In paper I, we performed an in vivo CRISPR-Cas9 screen targeting cell surface genes in murine AML stem cells and showed that CXCR4 is a top cell surface regulator of AML cell growth and survival. Notably, loss of CXCR4 signaling in leukemia cells leads to oxidative stress and differentiation in vivo. In contrast, the CXCR4 ligand CXCL12 is dispensable for leukemia development in recipient mice.

To identify key regulators of AML, in paper II, we performed an ex vivo cytokine screen on arrayed molecularly barcoded murine AML cells with a competitive in vivo read-out of their leukemia-initiating capacity. We identified TNFSF13 as a positive regulator of leukemia-initiating cells. We confirmed that TNFSF13 supports leukemia initiation under physiological conditions using Tnfsf13-/- mice. We further showed that TNFSF13 suppresses apoptosis and promotes AML cell proliferation in an NF-κB dependent manner.

DUX4-rearranged BCP-ALL is a recently identified molecular subtype characterized by the expression of the IGH-DUX4 fusion gene. With the aim of identifying biological dependencies of this subtype, in paper III, we performed a genome-wide CRISPR-Cas9 screen in the NALM6 cell line, driven by the IGH-DUX4 fusion gene, and two control cell lines. We showed that FNIP1, IRF4 and SYNCRIP are selectively important for the growth and survival of NALM6 cells and that their expression is under the control of the IGH-DUX4 fusion gene. While the deletion of FNIP1 led to the enrichment of transcriptional signatures associated with metabolic dysregulation, loss of IRF4 resulted in upregulation of genes involved in differentiation and apoptosis of NALM6 cells. Moreover, disruption of SYNCRIP caused downregulation of the TGFβ-SMAD signaling pathways in NALM6 cells.

In paper IV, we explored the immune-mediated anti-leukemic activity of the cytokine interleukin 4 (IL4) in a murine AML model. Overexpression of IL4 in AML cells resulted in a strong anti-leukemic effect accompanied by an expansion of macrophages in the bone marrow and spleen of the recipient mice. Depletion of macrophages in vivo eliminated the antileukemic effect of IL4. In addition, IL4 directly activates murine macrophages resulting in enhanced phagocytosis of AML cells in vitro. Interestingly, IL4 also induced Stat6-dependent upregulation of CD47 in AML cells thereby inhibiting phagocytosis. Consistent with this finding, IL4 stimulation combined with CD47 blockade enhanced macrophage-mediated phagocytosis of AML cells.

Taken together, the studies included in this thesis employed high-throughput functional screens using CRISPR-Cas9 and molecular barcoding techniques to identify key regulators of AML cells. These findings improve our understanding of the disease and may translate into the development of new therapies for acute leukemia.
Functional screens identify vulnerabilities in acute leukemia

Ramprasad Ramakrishnan

2020
Division of Clinical Genetics
Department of Laboratory Medicine, Lund
Faculty of Medicine, Lund University
Known is a mere handful, unknown is the size of the world

- Avvaiyar

1 Avvaiyar - Female Tamil poet from 10th century AD
# Table of Contents

Original papers ........................................................................................................ ix
Abbreviations ........................................................................................................ xi
Populärvetenskaplig sammanfattning .................................................................. xiii
Popular science summary ....................................................................................... xv

Hematopoiesis ........................................................................................................... 1
  Hematopoietic lineages and differentiation .......................................................... 1
  Regulation of hematopoiesis .............................................................................. 5
    Hematopoietic niche ....................................................................................... 5
    Lineage choices .............................................................................................. 5

Hematological malignancies ................................................................................. 7
  Myeloid malignancies ....................................................................................... 7
  Lymphoid malignancies .................................................................................... 10

Leukemogenesis ..................................................................................................... 11
  Rise of leukemia ............................................................................................. 11
    Pre-leukemia to overt leukemia .................................................................. 11
    Cell of origin .............................................................................................. 13
  Clonal heterogeneity and evolution ............................................................... 14

Acute myeloid leukemia ....................................................................................... 15
  Epidemiology .................................................................................................. 15
  Classification .................................................................................................. 15
  Genetic aberrations in AML .......................................................................... 17
    Cytogenetic abnormalities ........................................................................... 17
    Mutational landscape .................................................................................. 18
  AML microenvironment .................................................................................. 19
  Murine models of AML .................................................................................. 20
  Clinical aspects ............................................................................................... 21
    Diagnosis .................................................................................................... 21
    Risk stratification ......................................................................................... 22
Original papers

This thesis is based on the following original papers:

**Paper I**

**Paper II**

**Paper III**
Ramakrishnan, R., Rodriguez-Zabala, M., Lilljebjörn, H., Askmyr, M., Rissler, M., Järås, M., and Fioretos, T. **A genome-wide CRISPR screen identifies key regulators of DUX4-rearranged BCP-ALL.** *Manuscript*

**Paper IV**
Peña-Martínez, P., Ramakrishnan, R., Högborg, C., Jansson, C., Gisselsson-Nord, D., & Järås, M. **IL4 has a dual role in regulating phagocytosis of murine leukemia cells.** *Manuscript*
Papers not included in the thesis:


<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>aCML</td>
<td>Atypical chronic myeloid leukemia</td>
</tr>
<tr>
<td>ADCC</td>
<td>Antibody dependent cell cytotoxicity</td>
</tr>
<tr>
<td>ALL</td>
<td>Acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>AML</td>
<td>Acute myeloid leukemia</td>
</tr>
<tr>
<td>ATO</td>
<td>Arsenic trioxide</td>
</tr>
<tr>
<td>ATRA</td>
<td>All-trans-retinoic acid</td>
</tr>
<tr>
<td>BCP-ALL</td>
<td>B-cell precursor acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>CAR</td>
<td>Chimeric antigen receptor</td>
</tr>
<tr>
<td>Cas</td>
<td>CRISPR-associated</td>
</tr>
<tr>
<td>CBF</td>
<td>Core binding factor</td>
</tr>
<tr>
<td>CEL</td>
<td>Chronic eosinophilic leukemia</td>
</tr>
<tr>
<td>CHIP</td>
<td>Clonal hematopoiesis of indeterminate potential</td>
</tr>
<tr>
<td>CLL</td>
<td>Chronic lymphocytic leukemia</td>
</tr>
<tr>
<td>CLP</td>
<td>Common lymphoid progenitors</td>
</tr>
<tr>
<td>CML</td>
<td>Chronic myelogenous leukemia</td>
</tr>
<tr>
<td>CMML</td>
<td>Chronic myelomonocytic leukemia</td>
</tr>
<tr>
<td>CMP</td>
<td>Common myeloid progenitors</td>
</tr>
<tr>
<td>CNL</td>
<td>Chronic neutrophilic leukemia</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>COL</td>
<td>Cell of origin of leukemia</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>crRNA</td>
<td>CRISPR RNA</td>
</tr>
<tr>
<td>CXCL12</td>
<td>CXC-chemokine ligand 12</td>
</tr>
<tr>
<td>dCas9</td>
<td>dead-Cas9</td>
</tr>
<tr>
<td>DECL</td>
<td>DNA-encoded chemical libraries</td>
</tr>
<tr>
<td>DNMT3A</td>
<td>DNA methyl transferase 3A</td>
</tr>
<tr>
<td>DOT1L</td>
<td>DOT1-like histone lysine methyltransferase</td>
</tr>
<tr>
<td>DUX4</td>
<td>Double-homeobox 4</td>
</tr>
<tr>
<td>ECCITE-seq</td>
<td>Expanded CRISPR-compatible CITE-seq</td>
</tr>
<tr>
<td>EGIL</td>
<td>European group for the immunological characterization of leukemias</td>
</tr>
<tr>
<td>ELN</td>
<td>European leukemia network</td>
</tr>
<tr>
<td>EMP</td>
<td>Erythro-myeloid progenitors</td>
</tr>
<tr>
<td>ERG</td>
<td>ETS-related gene</td>
</tr>
<tr>
<td>ESA</td>
<td>Erythropoiesis stimulating agents</td>
</tr>
<tr>
<td>ET</td>
<td>Essential thrombocythemia</td>
</tr>
<tr>
<td>FAB</td>
<td>French-American-British</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescence in situ hybridization</td>
</tr>
<tr>
<td>FLT3</td>
<td>FMS-like tyrosine kinase 3</td>
</tr>
<tr>
<td>FLT3-ITD</td>
<td>FLT3-internal tandem duplication</td>
</tr>
<tr>
<td>FLT3-TKD</td>
<td>FLT3-tyrosine kinase domain</td>
</tr>
<tr>
<td>FSHD</td>
<td>Facioscapulohumeral muscular dystrophy</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>Granulocyte-macrophage colony stimulating factor</td>
</tr>
<tr>
<td>GMP</td>
<td>Granulocyte-macrophage progenitors</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>HDC</td>
<td>Histamine dihydrochloride</td>
</tr>
<tr>
<td>HDR</td>
<td>Homology-directed repair</td>
</tr>
<tr>
<td>HMA</td>
<td>Hypomethylating agents</td>
</tr>
<tr>
<td>HSC</td>
<td>Hematopoietic stem cells</td>
</tr>
<tr>
<td>HSCT</td>
<td>Hematopoietic stem cell transplantation</td>
</tr>
<tr>
<td>HSPC</td>
<td>Hematopoietic stem and progenitor cells</td>
</tr>
<tr>
<td>IL15</td>
<td>Interleukin 15</td>
</tr>
<tr>
<td>IL2</td>
<td>Interleukin 2</td>
</tr>
<tr>
<td>IL4</td>
<td>Interleukin 4</td>
</tr>
<tr>
<td>JMML</td>
<td>Juvenile myelomonocytic leukemia</td>
</tr>
<tr>
<td>LIC</td>
<td>Leukemia initiating cell</td>
</tr>
<tr>
<td>LMPPP</td>
<td>Lymphoid-primed multipotent progenitors</td>
</tr>
<tr>
<td>LSC</td>
<td>Leukemia stem cells</td>
</tr>
<tr>
<td>MDS</td>
<td>Myelodysplastic syndromes</td>
</tr>
<tr>
<td>MDS/MPN</td>
<td>Myelodysplastic/Myeloproliferative neoplasms</td>
</tr>
<tr>
<td>MDS/MPN-RS-T</td>
<td>MDS/MPN with ring sideroblasts and thrombocytosis</td>
</tr>
<tr>
<td>MEP</td>
<td>Megakaryocyte-erythroid progenitors</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MLN</td>
<td>Mature lymphoid neoplasms</td>
</tr>
<tr>
<td>MM</td>
<td>Multiple myeloma</td>
</tr>
<tr>
<td>MM6</td>
<td>Monomac 6</td>
</tr>
<tr>
<td>MPN</td>
<td>Myeloproliferative neoplasms</td>
</tr>
<tr>
<td>MPP</td>
<td>Multi-potent progenitor</td>
</tr>
<tr>
<td>NGS</td>
<td>Next generation sequencing</td>
</tr>
<tr>
<td>NHEJ</td>
<td>Non-homologous end joining</td>
</tr>
<tr>
<td>NK</td>
<td>Natural killer</td>
</tr>
<tr>
<td>NOD</td>
<td>Non-obese diabetic</td>
</tr>
<tr>
<td>NSG</td>
<td>NOD/SCID/IL2Rγnull</td>
</tr>
<tr>
<td>NSG-β2mnull</td>
<td>NSG - beta2-microglobulinnull</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PAF</td>
<td>Polymerase associated factor</td>
</tr>
<tr>
<td>PAM</td>
<td>Protoscaler adjacent motif</td>
</tr>
<tr>
<td>PDX</td>
<td>Patient-derived xenograft</td>
</tr>
<tr>
<td>PMF</td>
<td>Primary myelofibrosis</td>
</tr>
<tr>
<td>pTEFb</td>
<td>Positive transcription elongation factor b</td>
</tr>
<tr>
<td>PV</td>
<td>Polycythemia vera</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>SCF</td>
<td>Stem cell factor</td>
</tr>
<tr>
<td>SCID</td>
<td>Severe combined immunodeficient</td>
</tr>
<tr>
<td>scRNA-seq</td>
<td>Single cell RNA sequencing</td>
</tr>
<tr>
<td>sgRNA</td>
<td>Single guide RNA</td>
</tr>
<tr>
<td>shRNA</td>
<td>Short hairpin RNA</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>T-ALL</td>
<td>T-cell acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>TALEN</td>
<td>Transcription activator-like effector nucleases</td>
</tr>
<tr>
<td>TAM</td>
<td>Tumor associated macrophages</td>
</tr>
<tr>
<td>tracrRNA</td>
<td>Trans-activating CRISPR RNA</td>
</tr>
<tr>
<td>WBC</td>
<td>White blood cell</td>
</tr>
<tr>
<td>WHO</td>
<td>World health organization</td>
</tr>
<tr>
<td>ZFN</td>
<td>Zinc finger nucleases</td>
</tr>
</tbody>
</table>
Populärvetenskaplig sammanfattning


Under normala omständigheter är produktionen av blodceller felfri. Men ibland inträffar onormala genetiska förändringar i stamcellerna vilket leder till en okontrollerad tillväxt av omogna blodceller som resulterar i blodcancer. Akut leukemi är en gruppbezeichning som anger de mest aggressiva cancerformer som uppkommer från vita blodceller. Beroende på om leukemin är av myeloid eller lymfoïd typ, kategoriseras den i akut myeloid leukemi (AML) respektive akut lymfoïd leukemi (ALL). Både AML och ALL drivs av en mängd olika genetiska avvikelser som förändrar de grundläggande egenskaperna hos normala stamceller och därmed förvandlar dem till leukemistamceller. Leukemiceller återfinns främst i benmärgen och regleras av ett molekylärt nätverk som består av ett antal gener, proteiner, metaboliter etc. För att förstå funktionerna hos dessa celler krävs identifiering av nyckelelementen i detta nätverk, och deras roll i leukemiceller. Det övergripande syftet med denna avhandling är att identifiera sårbarheter i akut leukemi. För att uppnå detta undersökte vi ett antal gener och proteiner med banbrytande molekylära tekniker och utvärderade deras förmåga att påverka tillväxten av leukemiceller.

I den första studien (artikel I) använde vi CRISPR, en så kallad gensax, för att ta reda på vilka gener som är kritiska för leukemistamcellerna i AML. CRISPR är en speciell typ av gensax som kan användas för att effektivt stänga av gener. Genom CRISPR stängde vi samtidigt av cirka 100 gener i leukemistamceller och studerade effekten på leukemicellernas tillväxt i en musmodell. Vi identifierade att genen CXCR4 är absolut nödvändig för leukemiceller. När denna gen stängdes av mognade leukemistamcellerna ut till celler med begränsad livslängd.
Detta berodde delvis på oxidativ stress som involverar uppbyggnad avgifter som produceras som en biprodukt i cellerna när syre omvandlas till energi. Vi fann också att interaktionen mellan proteinerna CXCL12 och CXCR4 inte var nödvändig för leukemistamcellerna. Detta var förvånande eftersom denna interaktion är nödvändig för normal produktion av blodceller. Således identifierade vi en viktig skillnad mellan funktionen av normala blodstamceller och leukemiceller. Detta fynd kan vara användbart vid utveckling av nya terapier för AML.

I nästa studie (artikel II) sökte vi efter cytokiner som påverkar leukemi-initierande förmåga hos AML-celler. Cytokiner är proteinmolekyler som fungerar som kemiska budbärare mellan celler och påverkar deras funktioner. Vi använde en innovativ teknik där leukemistamcellerna var markerade med DNA-streckkoder. Denna teknik möjliggjorde att vi samtidigt kunde utvärdera hur cirka 100 cytokiner påverkade leukemicellerna. Genom detta tillvägagångssätt fann vi att en av cytokinerna, TNFSF13, förbättrade den leukemi-initierande förmågan hos AML-celler och att de normala blodceller inte var nödvändiga för dess överlevnad. Detta kan vara användbart vid utveckling av nya terapier för AML.

I artikel III sökte vi efter gener som är kritiska för en subtyp av ALL som drivas av genen DUX4. Vi använde återigen CRISPR, denna gång för att stänga av samtliga gener i ALL cellerna och studera effekten på celltillväxt. Vi fann att tre gener: FNIP1, IRF4 och SYNCPRIP, var avgörande för överlevnaden av leukemicellerna i denna subtyp av ALL. Ytterligare studier pågår för att förstå via vilka mekanismer dessa gener påverkar leukemicellerna.

Vi hade tidigare identifierat att interleukin 4 (IL4) inhiberar AML-celler i en musmodell. Den exakta mekanismen bakom den negativa effekten av IL4 är dock inte känd. I artikel IV fann vi att IL4 aktiverar makrofager som sedan attackerar och dödar AML-celler. Makrofager är en typ av immunceller som vanligtvis äter skadade celler och bakterier. Intressant nog fann vi också att IL4 delvis skyddar leukemicellerna från de attackerande makrofagerna genom att öka nivåerna av CD47, ett protein på ytan av AML-celler. Blockering av CD47 tillsammans med IL4-stimulering ökade makrofagarernas dödande av AML-celler. Dessa fynd avslöjar den komplexa roll som IL4 spelar i AML.

Sammantaget har studierna som ingår i denna avhandling använt flera av de senaste molekylära teknikerna för att identifiera nyckelfaktorer som reglerar akut leukemi. Denna nya kunskap förbättrar vår förståelse av sjukdomen och kan leda till utvecklingen av nya behandlingsmetoder.
Blood is a vital component of almost all multicellular organisms and is often a symbol for life itself. Blood transports nutrients and oxygen to the different organs in our body and removes waste materials from them. It is also important for maintaining our body temperature and for fighting against infections. Blood cells are made in a spongy tissue located inside some of our bones called bone marrow. In the bone marrow, a small number of parent cells called stem cells are present. These stem cells make copies of themselves and also produce all mature blood cells. An average human body produces about 200 billion blood cells per day. Based on their color, blood cells are divided into red and white blood cells. They are also classified into myeloid and lymphoid cells based on their place of maturation and prevalence.

Under normal circumstances, the production of blood cells is error-free. Occasionally, abnormal genetic changes occur in the stem cells which leads to an uncontrolled growth of immature blood cells resulting in blood cancer. Acute leukemia is a collective term denoting the aggressive cancers of white blood cells. Depending on whether the cancerous white blood cells are of myeloid or lymphoid type, it is categorized into acute myeloid leukemia (AML) and acute lymphoid leukemia (ALL) respectively. Both AML and ALL are driven by a wide variety of genetic abnormalities which alter the fundamental characteristics of normal immature blood cells thereby turning them into leukemia stem cells. Leukemia cells also reside in the bone marrow and are regulated by a molecular network consisting of a number of genes, proteins, metabolites etc. Identifying the key elements of this network and their role in leukemia cells could help us understand how these cells are regulated. The overarching aim of this thesis is to identify vulnerabilities in acute leukemia. In order to achieve this, we used cutting edge molecular techniques to assess the functional role of a number of genes and proteins in leukemia cells.

In the first study (Paper I), we performed a CRISPR screen to identify genes that are essential for leukemia stem cells in AML. CRISPR is a special type of gene scissors that can be used to turn-off genes effectively. Through CRISPR, we simultaneously turned off about 100 genes in leukemia stem cells and studied the effect on leukemia cell growth in a mouse model. We identified the gene CXCR4 to be essential for leukemia stem cells. When this gene was turned off, the leukemia stem cells matured into cells with limited lifespan and eventually died. This was partly due to oxidative stress which involves the build-up of toxins produced as a
by-product in the cells when oxygen is converted into energy. We also found that the interaction between the proteins CXCL12 and CXCR4 was not necessary for leukemia stem cells. This was surprising because this interaction is essential for normal blood stem cells. Thus, we identified a key difference between the functioning of normal and leukemia stem cells. This finding could be useful for designing new therapies for AML.

In the next study (Paper II), we searched for cytokines that affect the leukemia-initiating capacity of AML cells. Cytokines are protein molecules that act as chemical messengers between cells and affect their functions. We used an innovative technique in which the leukemia stem cells were marked with DNA barcodes. This technique allowed us to simultaneously assess the individual effect of about 100 cytokines on the leukemia stem cells. By using this approach, we found that one of the cytokines, TNFSF13 enhanced the leukemia initiating-capacity of AML cells by reducing the cell death and increasing the cell growth.

In paper III, we searched for genes that are critical for a subtype of ALL driven by the gene DUX4. We again used the special gene scissor, CRISPR and turned off all the genes in the ALL cells simultaneously and studied its effect on leukemia cell growth. We found three genes, FNIP1, IRF4 and SYNCRIP to be essential for the survival of leukemia cells of this subtype of ALL. Further studies are ongoing to decipher the mechanisms by which these genes affect the leukemia cells.

We had previously identified interleukin 4 (IL4), a cytokine to be harmful for AML cells in a mouse model. However, the exact mechanism behind the negative effect of IL4 is not known. In paper IV, we examined if this negative effect was mediated by the immune cells. We found that IL4 activates macrophages which then attacks and kills AML cells. Macrophages are a type of immune cells that usually eat damaged cells and pathogens such as bacteria. Interestingly, we also found that IL4 partially protects the leukemia cells from the attacking macrophages by increasing the levels of CD47, a protein on the surface of AML cells. Blocking of CD47 along with IL4 stimulation mitigated the protective effect and increased the killing of AML cells by macrophages. These findings reveal the complex role played by IL4 in AML.

Collectively, the studies included in this thesis have employed several of the latest molecular techniques to identify key factors that regulate acute leukemia. The new knowledge created through these studies improves our understanding of the disease and may translate into the development of new therapies.
Hematopoiesis

Blood is a unique tissue in the human body that has fascinated the inquisitive minds for ages. According to the Greek philosopher Aristotle (384-322 BC), blood is a homogenous fluid produced in the heart and is distributed all over the body to nourish the different organs and to induce their growth [1, 2]. Between the 17th to the 19th century AD, just two types of blood cells were known based on their color under a microscope - red blood cells and white blood cells [3]. Modern medical science has come a long way from these rudimentary understandings of blood and has identified several types of specialized blood cells and their diverse functions such as transport of oxygen and nutrients to every cell in the body, defense against pathogens, coagulation etc. The process of generation of new blood cells is termed hematopoiesis (from Greek; haimato - blood, poiein - to make). The primary site of hematopoiesis in adults is the medullary cavity of the bone marrow. In addition, extramedullary hematopoiesis can take place in the spleen [4] and lungs [5].

The human body is estimated to have about $2.7 \times 10^{13}$ blood cells amounting to ~90% of all cells in the body [6]. The turnover of blood cells in an average human adult is about 1 trillion cells per day [7]. Moreover, during severe physiological stress such as infections or blood loss, even more blood cells are produced. In order to maintain the number and diversity of the blood cells, hematopoiesis is tightly regulated. By the end of 20th century, it was well established that all of the mature blood cells are derived from a common pool of cells termed hematopoietic stem cells (HSC) through a hierarchical differentiation process.

Hematopoietic lineages and differentiation

About a dozen different types of mature blood cells have been identified. They are broadly classified into two types – myeloid and lymphoid cells based on their site of origin and prevalence. Myeloid (from Greek; muelos - marrow) cells comprise of granulocytes, monocytes, erythrocytes and megakaryocytes all of which originate and mature in the bone marrow. As the name suggests, granulocytes have large cytoplasmic granules that contain enzymes and are classified into four types – basophils, eosinophils, neutrophils and mast cells. Lymphoid cells are predominantly found in the lymphatic (from Latin; lympha - clear water) circulation and include B-cells, T-cells and Natural Killer (NK) cells. While mature B-cells and
NK-cells are produced in the bone marrow, immature T-cells migrate to the thymus where they undergo maturation. Both myeloid and lymphoid cells arise from progenitor populations which in turn are derived from HSCs through complex regulations of differentiation.

The classical model of hematopoietic lineage commitment is depicted as a hierarchical tree involving successive binary fate decisions as HSCs differentiate through progenitor populations into mature blood cells (Figure 1) [8]. HSCs which are at the top of the hierarchy are defined based on two essential characteristics – long-term self-renewal and multipotency, i.e., the capacity to contribute to all types of mature blood cells. HSCs lose their self-renewal capacity while differentiating into multi-potent progenitor (MPP) populations which have reduced capacity to reconstitute the entire hematopoietic system in the long-term [9]. Subsequently, MPPs differentiate into committed progenitors namely common myeloid progenitors (CMP) [10], common lymphoid progenitors (CLP) [11] and lymphoid-primed multipotent progenitors (LMPP) [12].
CMPs have traditionally been considered as an oligopotent population, which is the only source of myeloid progenitors – granulocyte-macrophage progenitors (GMP) and megakaryocyte-erythroid progenitors (MEP). Recently, LMPPs with predominantly lymphoid developmental capacities have also been shown to give rise to GMPs suggesting that the segregation of myeloid and lymphoid lineages do not occur early in the hierarchical differentiation tree as previously thought [13]. Moreover, the capacity of CMP to produce a true GMP population is being questioned as they fail to differentiate into monocytes and neutrophils. Therefore, a suggestion that has been put forward is to rename CMPs as erythro-myeloid progenitors (EMP) which generate erythrocytes, basophils, eosinophils and megakaryocytes [14]. Megakaryocyte progenitor cells are derived from MEPs which further differentiate into megakaryocytes and platelets. However, it has been reported that megakaryocyte progenitor cells also arise directly from MPPs [15]. In the lymphoid compartment, CLPs that are derived from MPPs or LMPPs further differentiate to produce NK-cells, B-cells and T-cells.

The model of the hierarchical hematopoietic differentiation has been developed based on two key experimental pillars – i) immunophenotype-based cell purification, ii) in vitro and in vivo clonal assays. Every population in the hematopoietic system is defined by its immunophenotype -i.e., a unique profile of cell surface markers that can be used to purify the cells. Subsequently, the differentiation potential of these purified populations is assessed either by in vitro colony forming assays or by in vivo transplantation assays into murine models. For example, CMPs isolated based on their immunophenotype generate myeloid, erythroid or megakaryocytic colonies in a colony forming assay. In contrast, GMPs only give rise to myeloid colonies and MEPs produce erythroid or megakaryocytic colonies. Based on these observations, a model was proposed in which CMPs differentiate into GMPs and MEPs [10]. However, this interpretation is based on a key assumption, that the cells purified based on the immunophenotype are functionally homogenous. During the past decade, several studies have established that both the HSC pool and the progenitor populations are functionally heterogenous [14, 16, 17]. Moreover, the assessment of the differentiation potential of cells transplanted into immune-deficient mice does not reflect the fundamental properties of unperturbed native hematopoiesis. Using genetic labelling, it has been shown that in an unperturbed system, hematopoiesis is mainly driven by multipotent progenitors and not HSCs [18, 19]. Due to these limitations, the validity of the current model of the hierarchical hematopoietic differentiation is being challenged.

During the last decade, there has been an explosion of single cell RNA sequencing (scRNA-seq) studies [20-25] in the hematopoietic system that has offered a snapshot of the expression state of cells at a particular time point. These studies have proposed the idea that hematopoietic differentiation is in a continuum rather than in distinct differentiation stages [26]. For example, the HSC/MPP compartment in the upper tier of the hematopoietic tree has been proposed to be characterized by a
continuous differentiation landscape [27] as opposed to distinct subpopulations [28]. Moreover, transcriptional priming towards different lineages are initiated earlier than first anticipated, already at the MPP stage [27]. Similarly, profiling of myeloid progenitors using scRNA-seq has revealed tremendous heterogeneity, identifying 18 subpopulations with different degrees of lineage priming [29].

![Diagram](image)

**Figure 2. Alternative model of hematopoietic hierarchy**

HSCs and the progenitor populations are highly heterogenous in their capacity to differentiate into specific lineages. Lineage choices are made early in the hematopoietic hierarchy and the cells follow different trajectories of differentiation to become mature cells (adapted from Laurenti, et al., [30]).

scRNA-seq studies in the HSC/MPP compartment have identified three major trajectories in the differentiation landscape – lymphoid, erythroid and granulocytic/monocytic lineages [31]. Analysis of the progenitor populations revealed that multiple types of mature cells arises from more than one trajectory [22, 27]. These recent findings have led to the proposal of a new representation of the hierarchical organization of hematopoiesis which appreciates the heterogeneity of the cells, prevalence of early lineage choices and diversity in possible routes of differentiation [30] (**Figure 2**).
Regulation of hematopoiesis

For maintenance of a robust hematopoietic system, fate decisions such as self-renewal vs differentiation, quiescence vs proliferation, survival vs death and lineage choices are tightly regulated in HSCs and other progenitor populations. These fate choices are guided both by external stimuli and cell intrinsic factors which affect the molecular circuitries of the cells. External factors that regulate hematopoiesis include cytokines, chemokines, extracellular matrix and membrane-bound signaling molecules provided by the microenvironment where the HSCs reside. The internal factors include transcription factors, signaling modulators, epigenetic modifiers and cell cycle regulators [32].

HSCs are proposed to be localized mainly in a hypoxic bone marrow microenvironment [33] termed perivascular niche, where they reside adjacent to sinusoidal blood vessels [34-36]. The major factors that promote the maintenance of HSCs in the bone marrow are stem cell factor (SCF) [37, 38], CXC-chemokine ligand 12 (CXCL12) [39, 40] and thrombopoietin [41]. Perivascular mesenchymal stromal cells and endothelial cells are the main sources of SCF and CXCL12 [42, 43] in the bone marrow niche. Other cell types in the niche that regulate HSCs are osteoblasts [44], osteoclasts [45], schwann cells [46], macrophages [47], and megakaryocytes [48]. Apart from HSC, osteoblasts and perivascular mesenchymal stromal cells also affect lymphoid progenitors [42]. During times of hematopoietic stress such as infection [49], pregnancy [4] or other hematological disorders [50], HSCs are mobilized from the bone marrow and colonize tissues – predominantly spleen and liver leading to extramedullary hematopoiesis [51].

Lineage choices

Transcription factors play a key role in the differentiation of HSCs and progenitors into specific lineages. Stochastic fluctuations in the expression of transcription factors or environmental cues such as cytokines have been proposed to control the lineage choice [52]. For example, expression of GATA1 and SPI1 is mutually exclusive during erythroid and myeloid differentiation, respectively. Moreover, GATA1 and SPI1 have been shown to repress each other and to activate themselves via a positive feedback loop, favoring a stochastic model of lineage commitment [53]. However, this model has been challenged by a recent study which suggests that the lineage choice is made at an earlier stage and the transcription factors merely execute the differentiation of the cell [54].
Maintenance of lineage choice is achieved by the sustained expression of lineage-specific transcription factors in the committed cells and their progeny. Positive autoregulation of a transcription factor while inhibiting opposing factors leads to stability of lineage commitment. The “GATA switch” is a classic example of this phenomenon where early progenitors express the transcription factor GATA2 that induces GATA1 expression which in turn represses GATA2 and activates its own expression, thereby pushing cells towards the erythropoietic lineage [55].

Apart from transcription factors, cytokines play a vital role in lineage choices, proliferation, maturation, survival and activation of cells in the hematopoietic system [56]. For example, ectopic upregulation of granulocyte-macrophage colony stimulating factor (GM-CSF) receptor has been shown to instruct the conversion of CLPs from the lymphoid to myeloid lineage [57]. Similarly, expansion of lymphoid and myeloid progenitors can be induced by high levels of the FMS-like tyrosine kinase 3 (FLT3) ligand [58]. Collectively, cytokines instruct the lineage commitment of blood cells thereby regulating hematopoiesis.

Epigenetic modifiers are emerging as important regulators of hematopoiesis apart from transcription factors and cytokines. Epigenetic modifications such as histone acetylation and methylation of cytosine residues modulates chromatin accessibility and transcription of genes involved in lineage commitment [59]. For example, loss of DNA methyl transferase 3A (DNMT3A) results in impaired differentiation and clonal expansion of HSCs, causing a pre-leukemic state [60] (see Pre-leukemia to overt leukemia).
Hematological malignancies

Hematological malignancies represent a heterogenous group of neoplasms that are characterized by the abnormal production of blood cells. Under normal conditions, key characteristics of hematopoietic stem and progenitor cells (HSPC) such as self-renewal, differentiation and proliferation are well regulated in order to ensure efficient hematopoiesis. Occasionally, unintended genetic and epigenetic changes occur in these cells which dysregulates these characteristics resulting in the accumulation of abnormally differentiated blood cells that are dysfunctional. Interestingly, some of the initiating genetic lesions are not by themselves sufficient to cause an overt disease but lead to clonal hematopoiesis [61].

The genetic changes that lead to malignant transformation of normal HSPCs include gene mutations, chromosomal translocations and other structural abnormalities such as aneuploidy, copy number variations etc. Over the past two decades, next generation sequencing (NGS) technologies have been instrumental in identifying a number of these genetic aberrations. Several of these mutations have been used for diagnosis, risk stratification, selection of treatment regimen and to predict the prognosis of the patients. Similar to normal hematopoiesis, hematological malignancies are also broadly classified into myeloid and lymphoid malignancies based on the lineage of origin of the neoplasia. Additionally, they are classified into acute and chronic malignancies based on the rate of progression of the disease. Taking into account the diversity of the hematological malignancies, the world health organization (WHO) has made a classification system for these diseases [62-64] (Figure 3).

Myeloid malignancies

Myeloid malignancies comprise of four sub-categories - myeloproliferative neoplasms (MPN), myelodysplastic syndromes (MDS), myelodysplastic/myeloproliferative neoplasms (MDS/MPN) and acute myeloid leukemia (AML), all of which are caused by genetic lesions resulting in clonal proliferation of defective blood cells of myeloid lineage (Figure 3). Genetic mutations associated with myeloid malignancies belong to five main classes: transcription factors (e.g. CEBPA, IKZF1), tumor suppressors (e.g. TP53),
Figure 3. Major hematological malignancies

Flowchart depicting the major subtypes of myeloid and lymphoid malignancies included in the WHO classification of hematological malignancies. The subtypes of AML and BCP-ALL are discussed in detail in the following chapters.
signaling pathways (e.g. FLT3, RAS), epigenetic regulators (e.g. DNMT3A, EZH2) and components of the spliceosome (e.g. SF3B1, SRSF2) [65]. Understanding the complex mutational landscape of these diseases has been instrumental in improving risk stratification and prognosis of the patients.

MPN is characterized by a relatively slow clonal expansion of hematopoietic progenitors and increased proliferation of mature myeloid cells in the bone marrow [66]. MPN is associated with thrombotic and hemorrhagic events and an increased risk of transformation into secondary AML (sAML) [67]. MPN is further classified into chronic myelogenous leukemia (CML), chronic neutrophilic leukemia (CNL), polycythemia vera (PV), primary myelofibrosis (PMF), essential thrombocythemia (ET) and chronic eosinophilic leukemia (CEL) [63]. Genetic aberrations that lead to constitutive activation of signaling cascades and cytokine-independent proliferation are hallmarks of MPN. For example, CML is characterized by a reciprocal t(9;22)(q34;q11) translocation occurring in HSCs leading to the formation of the BCR/ABL1 fusion gene, which encodes for a constitutively active tyrosine kinase that promotes cell proliferation. Development of tyrosine kinase inhibitors have revolutionized the treatment of CML patients allowing for a near-normal life [68].

MDS is a heterogenous disease characterized by morphological dysplasia of myeloid cells and inefficient hematopoiesis resulting in cytopenias [63]. Approximately one-third of MDS patients progress to high risk MDS and sAML [69]. The founding genetic event in MDS is believed to occur in HSCs resulting in clonal hematopoiesis of indeterminate potential (CHIP) (see Pre-leukemia to overt leukemia). The transformation from CHIP to MDS is suggested to involve a complex interplay between epigenetic changes in the HSCs, dysregulation of the bone marrow microenvironment and acquisition of additional driving mutations [70]. Due to the heterogeneity of the disease, a risk-adapted treatment strategy is adopted. Allogenic HSC transplantation is the only curative option which is offered based on the availability of donors and fitness of patients. Other noncurative treatments include administration of erythropoiesis stimulating agents (ESA), hypomethylating agents (HMA), immunosuppressants and red blood cell transfusions, aimed at improving cytopenias and quality of life [71].

Myeloid neoplasms that exhibit clinical and morphological characteristics overlapping with both MDS and MPN along with ineffective hematopoiesis are classified as myelodysplastic/myeloproliferative neoplasms (MDS/MPN). MDS/MPN includes five distinct subtypes - chronic myelomonocytic leukemia (CMML), atypical chronic myeloid leukemia (aCML), juvenile myelomonocytic leukemia (JMML) and MDS/MPN with ring sideroblasts and thrombocytosis (MDS/MPN-RS-T) [63]. Many of the somatic mutations and cytogenetic abnormalities found in MDS and MPN are also found in MDS/MPN [72]. The current treatment options for MDS/MPN comprise of risk- and symptom- based therapies co-opted from other myeloid malignancies [73].
AML is an aggressive clonal disorder characterized by the accumulation of immature abnormally differentiated myeloid blast cells. It is a heterogenous disease with several different subtypes classified based on their biological and prognostic characteristics. As three of the studies (Papers I, II & IV) included in this thesis focus on AML, it is discussed in detail in a separate chapter (see *Acute myeloid leukemia*).

**Lymphoid malignancies**

Lymphoid malignancies are broadly classified into mature and precursor neoplasms based on the differentiation stage of the malignant lymphoid cells (Figure 3). As the name suggests, mature lymphoid neoplasms (MLN) are characterized by the clonal expansion of mature lymphocytes and are classified based on the type of lymphocyte involved into mature T-cell neoplasms, mature NK-cell neoplasms, mature B-cell neoplasms. Among these, mature B-cell neoplasms accounts for the majority of the cases and comprise of a heterogenous group of disorders driven by different genetic aberrations [74]. The WHO classification of lymphoid malignancies lists as many as 41 mature B-cell neoplasms [64]. Some of the predominant subtypes of mature B-cell neoplasms include chronic lymphocytic leukemia (CLL), multiple myeloma (MM), diffuse large B-cell lymphoma, and follicular lymphoma.

Precursor lymphoid neoplasms are characterized by the accumulation of abnormally differentiated lymphoblasts and are historically referred as acute lymphoblastic leukemia (ALL). ALL is the most common form of cancer among children [75]. Depending on whether B- or T-lymphoblasts are involved, ALL is further classified into T-cell acute lymphoblastic leukemia (T-ALL) or B-cell precursor acute lymphoblastic leukemia (BCP-ALL) respectively. Both classes of ALL comprise of multiple subtypes which are stratified based on the genetic lesions such as structural chromosomal alterations, DNA copy number alterations and gene mutations that contribute towards leukemia development [76]. Treatment of childhood ALL is considered a success story with an overall survival rate of ~80% and with certain subtypes approaching a cure rate of ~98% [77, 78]. However, survival and outcomes of adult ALL cases (18-60 years of age) continues to remain poor and is about 35% [79].

BCP-ALL accounts for about 85% of ALL cases [80]. One of the studies (Paper III) in this thesis focuses on BCP-ALL and is therefore discussed in detail in a separate chapter (see *B-cell precursor acute lymphoblastic leukemia*).
Leukemogenesis

Acute leukemia is a collective term referring to a group of aggressive hematological malignancies of either myeloid or lymphoid lineage and is characterized by a rapid clonal expansion of blast cells with impaired differentiation. The evolutionary path of leukemogenesis starting with the acquisition of the first somatic mutation eventually leading to the development of overt leukemia is poorly understood. A detailed genetic analysis of acute leukemia using NGS technology has revealed a complex clonal architecture with multiple driver and cooperating mutations, coexistence of multiple subclones and their evolution over time [81, 82]. Moreover, the varied response of different subclones to therapies adds another layer of complexity to the disease. A clear understanding of the process of leukemogenesis in AML and BCP-ALL would be beneficial in devising effective therapies.

Rise of leukemia

Somatic mutations in human genome are estimated to occur at a rate of ~0.06 – 1.47 x 10^-9 mutations per base pair per cell division [83]. The functional effect of a particular mutation depends on the type of mutation, the cell type and the genomic region in which the mutation occurs. Many of the mutations are functionally irrelevant and do not affect the fitness of the cells. However, some mutations will alter the fundamental characteristics of the cells such as self-renewal, differentiation or proliferation, which eventually may result in a neoplasm. A widely accepted model for the development of acute leukemia is that the occurrence of the first genetic lesion in specific hematopoietic cells results in a pre-leukemic state and the subsequent acquisition of additional mutations leads to the leukemic transformation.

Pre-leukemia to overt leukemia

As we age, there is an accumulation of mutations in different cell types. In the hematopoietic cells, several recurrent mutations have been shown to drive clonal hematopoiesis which is characterized by the overrepresentation of blood cells derived from a single clone containing the mutation. This is probably due to the
acquisition of mutations in HSCs that provide them with a growth advantage [84]. As clonal hematopoiesis may or may not transform into a hematological malignancy, it is labeled as CHIP. CHIP refers to the existence of a cancer-associated variant in the hematopoietic cells in the absence of malignancy but is associated with an increased risk of progressing into malignancies [85, 86]. A majority of mutations associated with CHIP occur in the genes encoding epigenetic modifiers, such as \textit{DNMT3A} and \textit{TET2}. Other genes frequently mutated in CHIP include \textit{ASXL1}, \textit{JAK2}, \textit{SF3B1} and \textit{TP53} [87]. Many of the genes mutated in CHIP have also been identified as recurrent mutations in MDS and AML [88-91]. Therefore, CHIP has been postulated as a pre-leukemic state which could progress into MDS and subsequently into AML (Figure 4a). The committed progenitor populations of the pre-leukemic cells are susceptible to additional somatic mutations such as \textit{NPM1} and \textit{FLT3} [92-94] which then leads to leukemic transformation. In population-based cohorts, CHIP was associated with a ~10-fold increased relative risk of the development of hematological malignancies over several years of follow-up [84].
Relative to AML, BCP-ALL is more common among children where the occurrence of clonal hematopoiesis is rare suggesting that there are other mechanisms of leukemogenesis in pediatric leukemias. A number of genetic lesions associated with BCP-ALL have been shown to arise in utero [95-99]. As secondary mutations are required for the leukemic transformation of most of the subtypes, some of these prenatal genetic lesions are hypothesized to produce a pre-leukemic state (Figure 4b) [100]. The secondary genetic events often target genes critical for lymphoid development such as RUNXI, IKZF1 and PAX5 [101]. The mechanisms by which prenatal pre-leukemia transforms into postnatal BCP-ALL is unclear. Since only a part of the carriers of the pre-leukemic lesions suffers secondary mutations and subsequent leukemic transformation, it has been hypothesized that environmental factors such as infections are involved in the acquisition of the secondary mutations [102, 103]. Dysregulation of immune cells, particularly T-helper cells and NK-cells by infections has been proposed to play a major role in the leukemic transformation [104, 105]. Interestingly, MLL-rearranged BCP-ALL (official name KMT2A) have been found to occur already in newborns [106] and is the only leukemia-related genetic lesion in some patients [107, 108] suggesting that the fusion itself is sufficient for leukemia onset.

**Cell of origin**

The beginning of leukemia can be traced back to a normal hematopoietic cell termed cell of origin of leukemia (COL) in which the first initiating genetic aberration occurs. In AML, HSPCs and myeloid progenitors have been shown to be the COL through transplantation based murine leukemia models [109]. However, only HSPCs are suggested to be the COL of human AML due to the presence of somatic mutations in pre-leukemic HSPCs that lead to CHIP and subsequently into AML [110]. In BCP-ALL, the COL varies between HSCs and committed lymphoid progenitors according to the underlying genetic lesion. For example, BCR-ABL1 fusions can originate either from HSCs or from a B-cell progenitor depending on the isoform of BCR-ABL1 [111]. It is important to note that identification of COL is generally challenging as the initiating genetic lesions can alter its properties such that the cell is reprogrammed into a different stage of differentiation [112].

Acute leukemia is sustained by a population of cells termed leukemia stem cells (LSC). LSCs are functionally defined as the cells with the ability to maintain and repopulate leukemia. They share the characteristic of self-renewal with normal HSCs. LSCs are also termed as leukemia initiating cells (LIC) as they can initiate leukemia when transplanted into a host. The use of the term “LIC” has been contentious as it has also been occasionally used to denote COL. Since additional mutations are often required for leukemic transformation, LSC/LIC may not always be the COL and is important to clarify the context in which these terms are used.
In AML, LSCs were initially thought to be a small population that shares immunophenotypic characteristics with normal HSCs [113, 114]. However, with refinements in the immunophenotypic definition of HSCs, LSCs were proposed to be present in non-HSC progenitor stages [115, 116]. Subsequently, LSCs have been shown to closely mirror normal LMPPs and GMPs suggesting that LSCs acquire abnormal self-renewal potential and maintains a hierarchical structure [117]. LSCs give rise to cells that lack the self-renewal capacity and the ability to terminally differentiate termed blast cells. Transcriptome analysis of LSCs revealed the presence of a “stemness” signature which is associated with adverse prognosis in AML [118]. In BCP-ALL, several studies have shown that the ability of leukemia cells to engraft immunodeficient mice is not restricted to specific immunophenotypes suggesting a stochastic stem cell model according to which most of the leukemia cells have the ability to propagate the disease [119-121].

**Clonal heterogeneity and evolution**

Accumulation of genetic alterations in the neoplastic cells is one of the defining features of cancer. Acquisition of a mutation that confers a cell with a selective growth advantage results in clonal expansion thereby creating a population of cells termed as a clone which carries the mutation. With the accumulation of additional mutations, many subclones are created which compete with each other and with normal cells over resources in the tissue microenvironment. Thus, cancer progression is essentially a process of mutational diversification and clonal evolution [122]. Different models of clonal evolution in tumors have been proposed [123]. Linear clonal evolution involves the stepwise accumulation of mutations that confer a strong selective advantage to progeny clones which replace the parent clones in full selective sweep [124]. By contrast, branching evolution involves the coexistence of multiple subclones, each with their selective growth advantages that compete for ascendency [125].

Due to the mobile nature of leukemia cells as opposed to the fixed tissue architecture of solid tumors, they can undergo more cellular mixing and a homogenous population through a linear clonal evolution could have been expected [126]. However, several genomic studies of acute leukemia have shown the presence of a heterogenous mixture of clonal populations suggesting a branching clonal evolution [127-132]. Thus, the eventual composition of the leukemic clones is determined by the competition between the different mutant populations.
Acute myeloid leukemia

Epidemiology

The incidence of AML per 100,000 people in USA is estimated to be ~4.3 [133]. In Sweden, about 350 AML cases are diagnosed every year (www.cancerfonden.se) and the overall prevalence is estimated to be 13.7 per 100,000 people [134]. AML is more common in the elderly with a median age of 72 years and a peak in incidence at 80-84 years of age [134]. About 25% of the patients diagnosed with AML have a previous hematological disease [135].

Classification

AML is a heterogenous disease encompassing a wide range of variations in cell morphology, molecular profile, response to treatment and prognosis. A robust classification system of the disease is necessary to gain insights into biology of the leukemia cells and to develop effective therapies for the specific subtypes. An early effort to have a standardized classification system for AML was undertaken by French-American-British (FAB) co-operative group in 1976. This classification was entirely based on the morphology of the leukemia cells in the bone marrow and peripheral blood taking into account the degree of maturation and the direction of differentiation towards one or more mature cell lines [136]. During the 1980s, the FAB-classification of AML was revised to include additional subtypes and modified diagnostic criteria to account for the stage of impaired differentiation of the leukemia cells in the patient sample [137, 138] (Table 1). Although morphological heterogeneity of AML is taken into consideration in the FAB-classification, it does not always reflect the diversity of genetic lesion in the disease.

Over the past two decades, the WHO has proposed and revised a new classification system for AML based on the morphology, presence of specific genetic abnormalities, therapy-responses and prior history of myeloid malignancy [63, 139, 140] (Table 1). The recurrent genetic abnormalities included in the WHO-classification system primarily involves structural changes. However, ~50% of AML patients present with normal karyotype and exhibit variable clinical response to conventional chemotherapy indicating disease heterogeneity within the group
Advancements in sequencing technologies have led to the identification of a number of recurring mutations in AML patient samples [88, 142-144]. The present WHO-classification system does not include several of these mutations. Therefore, a new system of classification based on the genomic mutation profile has been proposed that can better address the heterogeneity and risk stratification of AML [144].

<table>
<thead>
<tr>
<th>FAB Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0</td>
</tr>
<tr>
<td>M1</td>
</tr>
<tr>
<td>M2</td>
</tr>
<tr>
<td>M3</td>
</tr>
<tr>
<td>M4</td>
</tr>
<tr>
<td>M5</td>
</tr>
<tr>
<td>M6</td>
</tr>
<tr>
<td>M7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WHO classification</th>
<th>Genetic abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>AML with recurrent genetic abnormalities</td>
<td>RUNX1-RUNX1T1 fusion gene</td>
</tr>
<tr>
<td>AML with t(8;21)(q22;q22.1)</td>
<td>RUNX1-RUNX1T1 fusion gene</td>
</tr>
<tr>
<td>AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22)</td>
<td>CBFB-MYH11 fusion gene</td>
</tr>
<tr>
<td>AML with t(15;17)(q22;q21)</td>
<td>PML-RARA fusion gene</td>
</tr>
<tr>
<td>AML with t(9;11)(p21.3;q23.3)</td>
<td>MLL-AF9 fusion gene</td>
</tr>
<tr>
<td>AML with t(6;9)(p23;q34.1)</td>
<td>DEK-NUP214 fusion gene</td>
</tr>
<tr>
<td>AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2)</td>
<td>MECom activation or RPN1-EVI1 fusion gene</td>
</tr>
<tr>
<td>AML (megakaryoblastic) with t(1;22)(p13.3;q13.3)</td>
<td>RBM15-MKL1 fusion gene</td>
</tr>
<tr>
<td>AML with mutated NPM1</td>
<td>NPM1 loss of function mutations</td>
</tr>
<tr>
<td>AML with biallelic mutations of CEBPA</td>
<td>CEBPA dominant-negative mutations</td>
</tr>
<tr>
<td>AML with BCR-ABL1 (provisional entity)</td>
<td>BCR-ABL1 fusion gene</td>
</tr>
<tr>
<td>AML with mutated RUNX1 (provisional entity)</td>
<td>RUNX1 loss of function mutations</td>
</tr>
</tbody>
</table>

AML with myelodysplasia related changes
Therapy-related myeloid neoplasms
AML not otherwise specified
Genetic aberrations in AML

Cytogenetic abnormalities

About five decades ago, the first somatic chromosomal abnormality involving a balanced translocation t(8;21) was identified in AML [145]. Since then, a number of cytogenetic aberrations have been identified which are found in nearly half of the newly diagnosed adult AML patients [141]. Karyotyping serves as the primary diagnostic tool to identify the recurrent cytogenetic abnormalities which include chromosomal deletions, inversions, aneuploidy and translocations. Based on the number of cytogenetic aberrations, the AML cases are classified into normal karyotype and complex karyotype (≥ 3 abnormalities).

Some of the major chromosomal translocations include t(8;21)(q22;q22.1), t(15;17)(q22;q21), inv(16)(p13.1q22)/t(16;16)(p13.1;q22) and 11q23 abnormalities which together are found in ~29% of adult AML patients [146]. AML arising due the t(15;17)(q22;q21) translocation is characterized by a rapid expansion of abnormal promyelocytes and is termed acute promyelocytic leukemia [147]. This translocation results in the expression of the PML-RARA fusion gene which drives the leukemic transformation by blocking differentiation through transcriptional regulation of its target genes. Both t(8;21)(q22;q22) and inv(16)(p13.1q22)/t(16;16)(p13.1;q22) affect the proteins involved in the core binding factor (CBF) transcription complex and hence are termed CBF-AML [148]. The t(8;21)(q22;q22.1) rearrangement is characterized by RUNXI-RUNXIT1 (official name – AML1-ETO) fusion gene that encodes a protein which acts as a dominant negative regulator of RUNX1 target genes [149, 150]. However, the translocation by itself is insufficient to cause leukemia and secondary cooperative mutations are likely required [151]. Such additional mutations include activating mutations in KRAS, NRAS, ASXL1, KIT, PTPN11 and/or loss-of-function mutations in NFI1, which are found in about two-thirds of t(8;21)(q22;q22.1) AML patients [152-154]. Abnormalities in chromosome 16 includes inversion (16)(p13.1q22) and translocation (16;16)(p13.1;q22), both resulting in the formation of the CBFB-MYH11 fusion gene that alters the transcriptional profile of the cells leading to leukemic transformation. The inv(16) is more common than t(16;16) and is often accompanied by aneuploidies such as +22, +8, del(7q) and +21 [155].

MLL gene (official name KMT2A), a H3K4 methyltransferase is a key epigenetic regulator playing a critical role in the emergence of fetal HSCs and maintenance of adult HSPC [156, 157]. 11q23 abnormalities involve chromosomal translocations resulting in in-frame fusions of the MLL gene to more than 100 different partner genes [158]. However, six fusion partners – AF9, ENL, AF10, ELL, PTD and AF6 constitute ~83% of the MLL-rearranged AML cases [158]. The distribution of the different MLL-rearrangements varies with age with AF10 and AF9 being the most
common MLL-fusion partners in infant and pediatric AML cases, respectively. Although the proteins encoded by the different MLL-fusion genes have diverse functions, they can be classified into two groups based on their localization, either in the cytoplasm or nucleus of the cells. Cytoplasmic fusion partners include ESP15, GAS7, SH3GL1, AFDN and FOXO4 which have been shown to be important for the dimerization of MLL which contributes to leukemic transformation [159]. Nuclear proteins constitute the majority of the MLL-fusion partners and are involved in the regulation of transcriptional elongation via recruitment of polymerase associated factor (PAF) complex, the DOT1-like histone lysine methyltransferase (DOT1L) complex and positive transcription elongation factor b (pTEFb) complex that leads to deregulated transcription [160-164]. DOT1L is also a methyltransferase that catalyzes H3K79 dimethylation which is associated with actively transcribed genes. Through its interactions with MLL-fusion protein, DOT1L mediates H3K79 methylation on the regulatory regions of later HOXA cluster genes including HOXA9. This results in a distinct gene expression signature [165] with elevated expression of HOXA cluster genes and the HOX cofactor MEIS1 which has been shown to be essential for leukemia development [166, 167].

**Mutational landscape**

The first whole cancer genome was sequenced in 2008 from an AML patient [168]. Since then, several hundred AML genomes have been sequenced by NGS which has resulted in the emergence of the mutational landscape of AML [88, 144, 169]. The most common genetic change in AML involves mutations in the NPM1 gene amounting for 27% of adult AML cases [144]. NPM1 is a nuclear protein playing a critical role in ribosome biogenesis, DNA repair and regulation of apoptosis. The mutations result in the aberrant localization of the NPM1 protein in the cytosol which contributes to leukemogenesis [170].

Other recurrent mutations found in AML can be classified into several groups based on their functional role. These include mutations in activated signaling pathway components, epigenetic/chromatin modifiers, cohesin complex, RNA splicing factors, tumor suppressor and transcription factors (Table 2). The most commonly mutated gene involved in signaling pathways is FLT3, which encodes a tyrosine kinase that acts as a cytokine receptor for the FLT3 ligand. FLT3 mutations can occur either as FLT3-internal tandem duplication (FLT3-ITD) in the juxtamembrane region or as point mutations in the tyrosine kinase domain (FLT3-TKD) resulting in constitutive activation of the tyrosine kinase. This results in the activation of RAS, MAPK and STAT5 signaling pathways which leads to increased proliferation [171]. Other genes mutated in signaling pathway components include c-KIT, KRAS, NRAS and other kinases [88]. Mutations in genes encoding epigenetic and chromatin modifiers are usually the dominant clones in AML patients suggesting that they are
### Table 2. Recurrent gene mutations in AML

<table>
<thead>
<tr>
<th>Mutation category</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activated signaling pathway components</td>
<td>FLT3, NRAS, KRAS, PTPN11, NF1, KIT, CBL</td>
</tr>
<tr>
<td>Epigenetic/chromatin modifiers</td>
<td>DNMT3A, TET2, IDH1, IDH2-R140, IDH2-R172, WT1, ASXL1, MLL-PTD, PHF6, ASXL2, BCOR, EZH2</td>
</tr>
<tr>
<td>Cohesin complex</td>
<td>RAD21, SMC1A, SMC3, STAG1, STAG2</td>
</tr>
<tr>
<td>Splicesome</td>
<td>SRSF2, SF3B1, U2AF1, ZRSR2</td>
</tr>
<tr>
<td>Tumor suppressor</td>
<td>TP53</td>
</tr>
<tr>
<td>Transcription factors</td>
<td>RUNX1, CEBPA</td>
</tr>
<tr>
<td>Other</td>
<td>NPM1</td>
</tr>
</tbody>
</table>

Founding mutations [144]. These include mutations in DNMT3A, TET2, ASXL1, IDH1 and IDH2 which independently results in the dysregulation of DNA and histone methylation. The cohesin complex plays a key role in mediating sister chromatid cohesion during mitosis. Apart from that, it is also involved in DNA damage repair and transcriptional regulation. Mutations in the genes encoding the cohesin complex such as STAG1 and STAG2 results in the deregulation of cohesin mediated gene regulation [172]. Commonly mutated genes involved in RNA splicing are SF3B1, U2AF1, SRSF2 and ZRSR2 which are also frequently found in MDS patients suggesting that they progressed into secondary AML [173]. Mutations in the transcription factors CEBPA and RUNX1 together accounts for ~14% of AML cases [144] and leads to a block in myeloid differentiation. Many of the recurrent mutations mentioned above were found to co-occur suggesting that they are cooperating genetic lesions. Most prominent co-occurring mutations are in DNMT3A, NPM1 and FLT3.

### AML microenvironment

The bone marrow microenvironment plays an important role in AML development as it does in normal hematopoiesis (see Hematopoietic niche) [174]. Genetic lesions that occur in non-hematopoietic niche cells can contribute to the leukemic transformation. For example, constitutive activation of Wnt signaling pathway in murine osteoblast cells due to an activating mutation in β-catenin (official name Ctnnb1) has been shown to initiate an AML-like disease with common chromosomal aberrations [175]. Alternatively, AML cells can also remodel the bone marrow niche to support disease progression. This is achieved through the dysregulation of signaling in the niche cells resulting in aberrant cytokine secretion that supports the expansion of AML cells. For example, the secretion of CXCL12 by the mesenchymal stromal cells derived from the bone marrow of AML patients is markedly lower compared to that of healthy subjects. This results in an overall...
reduction of CXCL12 levels in the bone marrow plasma of AML patients which could contribute to the impaired maintenance of normal HSCs [176]. A corollary of this finding is that the AML cells could be less dependent on CXCL12 for the disease progression. In line with this, in article 1, we show that CXCL12 expression in the bone marrow is dispensable for AML development in a murine model. Apart from these changes, the bone marrow microenvironment has also been shown to protect LSCs from chemotherapy thereby contributing to therapy resistance and disease relapse [177].

**Murine models of AML**

Much of what we know about the mechanism of action of different genetic alterations associated with AML is through *in vivo* mouse models. They are also an integral part of drug development pipelines by providing a platform to test different pharmaceutical compounds prior to clinical trials. The existing mouse models of AML can be broadly categorized into three groups – carcinogen induced AML models, xenograft models and genetically engineered models [178]. Carcinogens such as chemicals, radiation and murine viruses have been used for random mutagenesis in murine hematopoietic cells leading to leukemia development [179-181]. Although these models have been instrumental in the identification of proto-oncogenes and development of anti-leukemic therapies, they do not fully phenocopy human AML.

AML patient-derived xenograft (PDX) mouse models are generated by the transplantation of leukemia cells from AML patients into mice which engraft in their bone marrow and mimic the human disease. Successful engraftment and repopulation of patient cells into mice depends on factors such as the availability of niche space, intensity of murine immune responses against the graft and the presence of supportive cross-reactive signals for human cells in the mice. In order to reduce the rejection of the transplantation, mice with different levels of immunodeficiencies such as nude, severe combined immunodeficient (SCID), non-obese diabetic (NOD)/SCID and NOD/SCID/IL2Rγnull (NSG) mice have been used [182]. In addition, ablation of mouse hematopoietic cells in the bone marrow through irradiation or use of c-Kit mutant mice [183] creates open niches for transplanted human cells to home and engraft. To further support the growth of human cells, NSG mice expressing three human cytokines – IL3, GM-CSF, SCF has been developed (NSG-S mice) which allows for higher engraftment of AML patient cells [184]. Additionally, NSG strains with mutated major histocompatibility complex (MHC) class I and class II beta2-microglobulin (NSG-β2mnull) [185] have been developed which reduces the immune reactivity of human cells against host tissue. Irrespective of the manipulations in the host, efficient expansion of PDX also depends on the underlying genetic aberrations in the AML cells [186].
Genetically engineered mouse models include transgenic mice with constitutive or conditional expression of \textit{PML-RARA} [187], \textit{RUNX1-RUNX1T1} [188], \textit{MLL-AF9} [189], \textit{MLL-ENL} [190] or cooperating mutations such as \textit{NPM1} + \textit{FLT3-ITD}/\textit{NRAS-G12D} [191, 192], \textit{MLL-PTD} + \textit{FLT3-ITD} [193], \textit{TET2} + \textit{FLT3-ITD} [194] in hematopoietic lineage cells. Genome editing techniques such as CRISPR/Cas9 and TALEN have been used to generate mouse models containing AML specific mutations [195] or cytogenetic abnormalities [196, 197]. Genetically engineered mouse models have also been generated through adaptive transfer of hematopoietic cells that are virally expressing AML-associated fusion genes such as \textit{MLL-ENL} [198], \textit{MOZ-TIF2} [199] or \textit{MLL-AF9} [200].

The \textit{MLL-AF9} driven AML mouse model is well characterized and has been used extensively in this thesis. In this model, AML is initiated in committed GMPs through retroviral introduction of \textit{MLL-AF9} and the leukemia cells can be serially propagated in mice. The transformed leukemia cells possess a gene expression signature associated with self-renewal indicating the progression of the disease from committed progenitor to LSCs. Moreover, a portion of the murine gene signature associated with self-renewal of leukemia cells is found in the human \textit{MLL-AF9} rearranged AML [200]. Serial propagations of the leukemia cells in recipient mice results in reduced disease latency possibly due to the enrichment of leukemia initiating cells \textit{in vivo} [201]. For example, transplantation of $1 \times 10^6$ quaternary transplant cells results in a fully developed AML along with splenomegaly in the mice in ~2-3 weeks. The short disease latency and the recapitulation of many of the AML characteristics makes it a suitable model for studying AML [200, 202].

Despite extensive efforts in the development of different AML mouse models, they do not completely recapitulate the complexity of human AML. Nevertheless, they are valuable in studying AML pathology and development of therapies for AML. With advancements in genome editing techniques, it is expected that newer mouse strains will be developed which better models the disease.

**Clinical aspects**

**Diagnosis**

Early symptoms of the early stages of AML include fever, lethargy, fatigue, frequent infections and unusual bleedings. Diagnosis of AML is performed based on morphology, immunophenotyping, cytogenetics and gene mutations in the cells according to the recommendations from the European leukemia network (ELN) [146]. Except for AML§ with t(15;17), t(8;21), inv(16)/t(16;16), a blood blast count of \(\geq 20\%\) is required for the diagnosis of AML. Immunophenotyping is performed
### Table 3. Risk stratification of AML patients according to the underlying genetic abnormalities

<table>
<thead>
<tr>
<th>Risk category</th>
<th>Genetic abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favorable</td>
<td>t(8;21)(q22;q22.1)</td>
</tr>
<tr>
<td></td>
<td>inv(16)(p13.1q22) or t(16;16)(p13.1;q22)</td>
</tr>
<tr>
<td></td>
<td>Mutated NPM1 without FLT3-ITD or with FLT3-ITD&lt;sup&gt;low&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Biallelic mutated CEBPA</td>
</tr>
<tr>
<td>Intermediate</td>
<td>Mutated NPM1 and FLT3-ITD&lt;sup&gt;high&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Wild-type NPM1 without FLT3-ITD or with FLT3-ITD&lt;sup&gt;low&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>t(9;11)(p21.3;q23.3)</td>
</tr>
<tr>
<td></td>
<td>Cytogenetic abnormalities not classified as favorable or adverse</td>
</tr>
<tr>
<td>Adverse</td>
<td>t(6;9)(p23;q34.1)</td>
</tr>
<tr>
<td></td>
<td>t(v;11q23.3)</td>
</tr>
<tr>
<td></td>
<td>t(9;22)(q34.1;q11.2)</td>
</tr>
<tr>
<td></td>
<td>inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2)</td>
</tr>
<tr>
<td></td>
<td>−5 or del(5q); −7; −17/abn(17p)</td>
</tr>
<tr>
<td></td>
<td>Complex karyotype, monosomal karyotype</td>
</tr>
<tr>
<td></td>
<td>Wild-type NPM1 and FLT3-ITD&lt;sup&gt;high&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Mutated RUNX1</td>
</tr>
<tr>
<td></td>
<td>Mutated ASXL1</td>
</tr>
<tr>
<td></td>
<td>Mutated TP53</td>
</tr>
</tbody>
</table>

To assess the expression levels of specific lineage markers for establishing the diagnosis of AML and distinguishing it from mixed-phenotype acute leukemia [203]. Subsequently, cytogenetic analysis is performed to identify any of the recurrent chromosomal abnormalities. Additionally, screening for mutations in NPM1, CEBPA, RUNX1, FLT3, TP53 and ASXL1 genes is performed for establishing the subtype and the prognosis of the AML cases.

### Risk stratification

According to the ELN recommendations [146], AML cases are risk stratified based on the underlying genetic abnormality and their response to standard therapy (Table 3). With increased numbers of AML genomes sequenced, patterns of co-occurring mutations have been identified which could be used to further refine the risk stratification. For example, co-occurrence of FLT3-ITD, NPM1 and DNMT3A mutations has a worse prognosis compared to mutations in just FLT3-ITD and NPM1. Similarly, mutations in IDH2 or DNMT3A do not by themselves have prognostic information, but their co-occurrence results in a worse prognosis [144].
Therapies

Although AML is a heterogenous disease with several molecular subtypes, the treatment regimen is surprisingly similar involving induction chemotherapy followed by consolidation chemotherapy for the majority of the patients. Colloquially termed “7+3” regimen, the induction therapy involves administration of cytarabine for 7 days along with short infusions of an anthracycline, usually daunorubicin on each of the first 3 days [204]. While the use of such high doses of chemotherapy is aimed at achieving complete remission (less than 5% blasts in the bone marrow) by targeting rapidly dividing AML blast cells, it also exerts toxic effects on normal cells and is therefore not suitable for elderly patients. In such patients, non-intensive treatments such as hypomethylating agents, for example, azacytidine, low dose cytarabine or cytostatic drugs such as hydroxyurea are administered [146].

In the absence of additional treatments, the great majority of patients relapse after achieving complete remission following induction therapy [205]. Therefore, appropriate post-remission consolidation therapy is necessary. Consolidation therapy usually involves administration of additional intermediate-dose chemotherapy and allogeneic hematopoietic stem cell transplantation (HSCT) i.e., transplantation of HSCs from the bone marrow of a healthy matching donor. Although allogenic HSCT has a strong anti-leukemic effect, it is associated high risk of non-relapse treatment related mortality due to graft versus host disease and is therefore recommended only for intermediate- and adverse- risk patients [146].

Even after receiving standard therapy, the relapse rate is ~50% in AML patients < 60 years of age and is as high as ~85% in older patients [206]. Such high relapse rates in AML is often attributed to the inability to target quiescent LSCs using standard chemotherapy. Moreover, AML cells may also acquire additional mutations that confer them with resistance to additional chemotherapy [127]. Other possible mechanisms by which AML cells acquire chemoresistance include increased anti-apoptotic signaling through upregulation of BCL2 [207] and drug efflux [208].

Apart from standard therapy, targeted therapies have been developed for specific subgroups of AML. The use of all-trans-retinoic acid (ATRA) and arsenic trioxide (ATO) in the treatment of patients diagnosed with acute promyelocytic leukemia characterized by the PML-RARA rearrangement has been highly effective achieving a remission rate of >90 % [209]. ATRA/ATO combination therapy degrades PML-RARA fusion protein and results in the differentiation of the leukemia cells [210]. Recently, several new targeted therapies have been approved for the treatment of AML patients. These include a drug conjugated anti-CD33 antibody (gemtuzumab ozogamicin), FLT3 inhibitors (midostaurin, gilteritinib), IDH1/IDH2 inhibitors (Enasidenib, Ivosidenib) and a BCL2 inhibitor (Venetoclax) [211]. Venetoclax in combination with hypomethylating agents and low dose cytarabine has been
approved for treating AML patients who are previously untreated and unfit for intensive chemotherapy. This combination has shown promising results for patients in this group who lack effective treatment options.

**Immunotherapy**

Successful usage of allogeneic HSCT to induce a graft versus leukemia effect shows that AML cells are susceptible targets of donor immune cells, specifically NK- and T-cells [212]. Autologous immune cells have also been shown to target AML cells when stimulated with cytokines [213]. Thus, it is clear that immune cells have the ability to target AML cells which is the fundamental principle behind immunotherapy. Immunotherapies can be based either on boosting the patient’s own immune system with cytokines or vaccines or by conferring immunity by adoptive T-cell therapy, NK-cell therapy or by monoclonal antibodies.

Cytokines are key regulatory factors of leukemic cells and have also been shown to activate immune cells of AML patients. Administration of interleukin 2 (IL2) supplemented with histamine dihydrochloride (HDC) has been shown to activate NK- and T-cells and be efficacious in AML patients [214]. Interleukin 15 (IL15) has also been shown to enhance autologous NK-cell cytotoxicity in pre-clinical studies [215]. In paper IV of this thesis, we elucidate the complex regulatory role of interleukin 4 (IL4) on macrophages. Thus, it is clear that cytokines regulate immune cells and could potentially be developed as immunotherapy. Therefore, it is worth investigating the role of other cytokines in immune regulation.

Immune evasion in AML is mediated by the expression of cell surface proteins on the leukemia cells that binds to inhibitory receptors on immune cells. A classic example of this phenomenon is the upregulation of CD47 by AML cells which inhibits phagocytosis by providing a “don’t eat me” signal to macrophages. Blocking of CD47 on AML cells using monoclonal antibodies has been shown to enhance phagocytosis of leukemia cells in pre-clinical studies [216].

AML cells can also be targeted by directing monoclonal antibodies against cell surface receptors that are upregulated selectively on AML cells. Gemtuzumab ozogamicin is a drug conjugated antibody targeting CD33 which is upregulated on AML cells. Upon binding to CD33, the antibody-drug conjugate is internalized by the leukemia cells resulting in cell death [217]. IL1RAP [218], CD123 [219], CLL-1 [220] and TIM-3 [221] are some of the other cell surface receptors that have been identified as therapeutic targets of AML using monoclonal antibodies. Upon binding the target cells, the monoclonal antibodies could elicit anti-leukemic activity by modulating the signaling mediated through the receptor as well as recruit immune cells through the fragment crystallizable (Fc) region leading to antibody dependent cell cytotoxicity (ADCC) [222]. Thus, the use of monoclonal antibodies has strong anti-leukemic effect due to the dual mechanism of inhibition of AML cells.
B-cell precursor acute lymphoblastic leukemia

Epidemiology

BCP-ALL is the most common type of pediatric cancer with about 75% of cases occurring between 2-5 years of age [223, 224]. In children younger than 14 years of age, the age-adjusted incidence of BCP-ALL in Sweden between 1989 and 2001 was 3.36 – 3.59/100,000 children with a slightly higher predominance in boys [225].

Classification

Based on the morphology of the leukemic blasts, ALL was first classified into L1, L2 and L3 subtypes according to FAB classification [136, 226]. A major limitation of this system of classification was that these subtypes did not distinguish BCP-ALL from T-ALL. Subsequently, the European group for the immunological characterization of leukemias (EGIL) proposed a revised classification of BCP-ALL into B-I (pro-B), B-II (common-B) and B-III (pre-B) ALL based on the immunophenotype corresponding to the differentiation stage of the leukemic blasts [227]. With the identification of a number of cytogenetic aberrations and gene mutations associated with BCP-ALL, WHO has proposed a molecular classification system based on recurrent genetic abnormalities (Table 4) [63]. Until recently, ~20% of the BCP-ALL cases were still unclassified due to the lack any known genetic aberration. Deep genomic characterization of these cases has led to the identification of several new oncogenic subtypes [228] inclusion of which would further refine the current WHO classification system.

Genetic abnormalities

The established molecular subtypes of BCP-ALL are defined by the presence of aneuploidies – high hyperdiploidy (50-67 chromosomes) and hypodiploidy (less than 44 chromosomes) or fusion genes – BCR-ABL1, IL3-IGH, ETV6-RUNX1 (also
Table 4. WHO classification of BCP-ALL

<table>
<thead>
<tr>
<th>Category</th>
<th>Genetic abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCP-ALL with recurrent genetic abnormalities</td>
<td></td>
</tr>
<tr>
<td>BCP-ALL with t(9;22)(q34.1;q11.2)</td>
<td>BCR-ABL1 fusion gene</td>
</tr>
<tr>
<td>BCP-ALL with t(v;11q23.3)</td>
<td>MLL-rearranged fusion gene</td>
</tr>
<tr>
<td>BCP-ALL with t(12;21)(p13.2;q22.1)</td>
<td>ETV6-RUNX1 fusion gene</td>
</tr>
<tr>
<td>BCP-ALL with hyperdiploidy</td>
<td>Hyperdiploidy</td>
</tr>
<tr>
<td>BCP-ALL with hypodiploidy</td>
<td>Hypodiploidy</td>
</tr>
<tr>
<td>BCP-ALL with t(5;14)(q31.1;q32.3)</td>
<td>IL3-IGH fusion gene</td>
</tr>
<tr>
<td>BCP-ALL with t(1;19)(q23;p13.3)</td>
<td>TCF3-PBX1 fusion gene</td>
</tr>
<tr>
<td>BCP-ALL with BCR-ABL1-like</td>
<td>Translocations involving kinases - ABL1, ABL2, PGFRB, NTRK3, TYK2, CSF1R, JAK2</td>
</tr>
<tr>
<td>BCP-ALL with iAMP21</td>
<td>Amplification of a portion of chromosome 21 resulting in 5 or more copies of RUNX1</td>
</tr>
</tbody>
</table>

BCP-ALL not otherwise specified

termed TEL-AML1), TCF3-PBX1 (also termed E2A-PBX1) or MLL-rearrangements. Additionally, two more subtypes - BCR-ABL1-like and iAMP21 have been provisionally included in the classification system [63]. Among the different molecular subtypes, high hyperdiploidy and ETV6-RUNX1 together accounts for more than half of pediatric BCP-ALL cases while the other subtypes individually constitute less than 10% (Figure 5) [228].

The genetic aberrations in ALL have been shown to deregulate transcription, resulting in distinct gene expression patterns [229, 230]. For example, the t(12;21)(p13;q22) translocation characterized by the formation of the ETV6-RUNX1 fusion protein, alters the transcriptional profile by inhibiting the expression of RUNX1 target genes through the recruitment of a transcriptional co-repressor complex with histone deacetylase (HDAC) activity [231, 232]. This altered transcription profile affects the differentiation capacity of the cells as RUNX1 has been shown to be critical for normal hematopoiesis [233, 234]. Similarly, the t(1;19)(q23;p13.3) translocation characterized by the TCF3-PBX1 fusion protein results in the aberrant activation of PBX1 target genes [235]. Additionally, the fusion protein represses the expression of TCF3 target genes leading to uncontrolled cell cycle progression [236]. The genetic abnormalities of these established subtypes are assumed to be the initiating events that are accompanied by secondary genetic aberrations, which contribute to disease evolution and progression. These mutations occur in genes such as IKAROS, PAX5, EBF1, ARIDB5, CEBPE, CDKN2A, many of which encode key transcriptional regulators of B-cell development [237].
Cytogenetic abnormalities are a hallmark of pediatric BCP-ALL and are used to stratify them into molecular subtypes with unique gene expression profiles. The data presented in the pie chart is adapted from Liljebjörn et al [238].

Recent studies of BCP-ALL cases using high-resolution sequencing techniques have led to the identification of additional molecular subtypes with recurrent chromosomal rearrangements and distinct gene expression profiles [238-241]. Double-homeobox 4 (DUX4)-rearranged BCP-ALL is one of these new subtypes which constitutes ~ 4-5% of all pediatric BCP-ALL cases (Figure 5). As a newly identified subtype, DUX4-rearranged BCP-ALL is yet to be fully characterized.

**DUX4-rearranged BCP-ALL**

DUX4 is a transcription factor which is normally expressed during human embryonic development from oocyte to the 4-cell stage to initiate zygotic genome activation [242]. It is epigenetically silenced thereafter during the rest of the development and in somatic tissues. The reactivation of DUX4 expression in skeletal muscle leads to the degradation of the muscle cells resulting in facioscapulohumeral muscular dystrophy (FSHD) [243]. Apart from BCP-ALL, DUX4-rearrangements resulting in aberrant expression of DUX4 have also been reported in Ewing-like sarcoma [244] and rhabdomyosarcoma [245].

DUX4 is present in 11-100 copies on each allele in the subtelomeric D4Z4 repeat region on chromosome 4q and 10q. The leukemia-initiating event in DUX4-rearranged BCP-ALL is a chromosomal translocation which leads to the insertion of DUX4 within the IGH locus in chromosome 14. This results in the expression of
a chimeric IGH-DUX4 protein in which the C-terminus of DUX4 is truncated which leads to transcriptional deregulation and subsequent differentiation arrest in the late pro-B/pre-B-cell stage of B-cell development [246]. Although IGH translocations are common driver events in several malignancies of lymphoid lineage [247-249], the precise contribution of IGH locus in DUX4-rearranged BCP-ALL is currently unclear. Wild-type DUX4 is a transcription factor that contains a double-homeobox domain with transactivation capacity. Although IGH-DUX4 shares a similar DNA binding mechanism with the wild-type DUX4, the transactivation capacity of IGH-DUX4 is markedly reduced due to the truncation of the C-terminal domain [250]. Consistent with this, IGH-DUX4 exhibits a distinct transcription profile compared to wild-type DUX4 [251].

A large number of DUX4-rearranged BCP-ALL cases also exhibit heterozygous intragenic ETS-related gene (ERG) deletions [238, 239]. ERG encodes a transcription factor that is essential for B-cell development and ERG deletions have previously been identified as a molecular subtype of BCP-ALL [252, 253]. Interestingly, IGH-DUX4 has also been shown to induce the expression of a C-terminal truncated isoform of ERG with a non-canonical transcription start site termed ERGalt. ERGalt acts as a dominant negative inhibitor of wild-type ERG and has the potential of leukemic transformation in mice [239]. Thus, ERG deregulation seems to cooperate with the DUX4-rearrangement in disease progression.

**Models of BCP-ALL**

Studies of BCP-ALL have been greatly facilitated by the establishment of cell lines derived from patient samples. More than 150 individual BCP-ALL cell lines have been reported in the literature [254]. These include cell lines such as MHH-CALL2, REH, 697, TOM1, SEM and NALM6 each representing one of the distinct molecular subtypes of BCP-ALL. Although the NALM6 cell line was established about four decades ago [255] and has been extensively studied as a pre-B ALL cell line, the underlying genetic abnormality, a DUX4-rearrangement, was identified only a few years ago [240]. The use of cell lines as models to study BCP-ALL has several advantages such as monoclonality, ease of manipulation and can be grown in vitro in large numbers. However, cell lines do not fully reflect the disease as their physiology, phenotype and other features are altered due to adaptation to in vitro growth conditions.

Apart from cell lines, in vivo mouse models provide insights into disease processes and act as preclinical platforms for testing of therapies. The classical approach to develop mouse models of different hematological malignancies is to overexpress the oncogene of interest in normal HSPCs of mice [256]. However, development of in vivo mouse models of BCP-ALL using this approach has largely been
unsuccessful. For example, mouse models generated by overexpression of ETV6-RUNX1 in murine bone marrow cells lead to the accumulation of early progenitors in the B-cell compartment but do not result in leukemic transformation [257-259]. Notably, additional cooperating mutations caused due to low-dose irradiation along with the loss of Cdkn2a induced leukemic transformation [260]. These findings demonstrate that the expression of the ETV6-RUNX1 fusion gene by itself is not sufficient for the development of BCP-ALL which is in line with the clinical observations [96]. Similarly, transgenic mice generated by the overexpression of TCF3-PBX1 in lymphoid cells developed T-cell ALL instead of B-cell ALL [261]. Subsequently, two mouse models of TCF3-PBX1 BCP-ALL have been developed either by expression of TCF3-PBX1 specifically in B-cell lineage cells [262] or in bone marrow cells of T-cell deficient mice [263]. Additionally, mouse models for several other subtypes of BCP-ALL have been developed with varying levels of success [256]. One of the common features in the majority of these mouse models is the long latency periods with the expression of single leukemia associated aberration which makes the studies difficult and time-consuming. Engineering additional cooperating mutations in these models using advanced genome editing techniques could result in models with a more rapid manifestation of the disease.

Clinical aspects

Diagnosis

Common symptoms of BCP-ALL include B-symptoms (fever, weight loss, night sweats), easy bleedings, fatigue, dyspnea and infections [264]. Other symptoms include anemia, thrombocytopenia and leukopenia. The primary diagnostic criteria for BCP-ALL is the presence of 20% of lymphoblasts in the bone marrow during morphological analysis. Additionally, cyogenetic analysis is routinely performed to identify the common chromosomal abnormalities except for t(12;21)(p13;q22) translocation (ETV6-RUNX1) which is detected using fluorescence in situ hybridization (FISH) assay [140]. Eosinophilia in the peripheral blood is a marker of t(5;14)(q31;q32) (IL3-IGH) translocation [265] and is included in the diagnostic criteria of this subtype. Apart from this, cerebrospinal fluid analysis is performed to investigate central nervous system (CNS) involvement in the disease.

Risk stratification

Different ALL study groups around the world have risk stratified BCP-ALL patients into standard-risk, intermediate-risk and high-risk groups according to their specific criteria. Broadly, the risk stratification is based on major prognostic factors which
includes age, white blood cell (WBC) count, cytogenetic abnormalities, immunophenotype and treatment responses. Increasing age is associated with poor prognosis of BCP-ALL with patients over the age of 60 having poor long-term survival [266]. Similarly, patients with cytogenetic abnormalities such as hypodiploidy, MLL-rearrangement or BCR-ABL1 are classified in the high-risk category due to their poor prognosis.

**Therapies**

The standard treatment of BCP-ALL is one of the most effective risk-based treatment regimens with high overall survival. Although there are small differences in the standard treatment protocols adopted by different ALL study groups, they broadly consist of a multidrug regimen that is administered in three phases – induction, consolidation and maintenance with an overall duration of about 2 to 3 years [178]. The drugs administered in the induction phase are vincristine, corticosteroids, asparaginase and anthracycline and lasts for 4-6 weeks. About 95% of the patients achieve complete remission following the completion of induction therapy. Allogenic HSCT is sometimes performed upon achieving complete remission. The consolidation phase usually includes a combination of chemotherapeutic drugs that are not used in the induction phase to minimize drug resistance. The duration and intensity of consolidation therapy depends on the risk group of the patients and usually lasts for 6 to 9 months. Maintenance chemotherapy is the final and the longest phase of standard treatment regimen lasting for about 2 years and involves administration of low intensity chemotherapy and antimetabolite therapy. Additionally, therapy is directed towards the brain both as a prophylaxis and for patients with a clinical CNS disease. Despite achieving complete remission, about 10-20% of pediatric BCP-ALL patients relapse after standard therapy and the treatment options for relapsed BCP-ALL are limited [267].

In addition to the standard treatment, several targeted therapies have been developed for the treatment of BCP-ALL patients. Antibody-based immunotherapies - Rituximab [268], inotuzumab ozagomycin [269, 270] and blinatumomab [271] targeting CD20, CD22 and CD19 respectively have improved outcomes for specific groups of patients who do not respond to standard therapy. Other successful immunotherapeutic strategy in BCP-ALL is the development of chimeric antigen receptor (CAR) T-cells that combines the specificity provided by monoclonal antibodies and the cytotoxicity provided by engineered autologous T-cells. CAR T-cell therapy targeting CD19 has been highly effective in relapsed/refractory BCP-ALL patients with 70-97% of patients achieving complete remission [272].
CRISPR/Cas9 mediated genome engineering

Our ability to manipulate the DNA through genome engineering is instrumental in deciphering the functions of different coding genes and non-coding regions of the genome. Early efforts to edit the genome of eukaryotic cells was through homologous recombination [273, 274]. Although homologous recombination was a revolutionary technique of its time, it was labor-intensive, time consuming and had poor editing efficiencies. Subsequently, two programmable nucleases - zinc finger nucleases (ZFN) and transcription activator-like effector nucleases (TALEN) were developed both of which include sequence-specific DNA-binding domains attached to a non-specific DNA cleavage module inducing targeted double-strand breaks [275, 276]. Although ZFNs and TALENs can be used to make precise genome editing, both these techniques have several disadvantages which include low editing efficiency, off-target effects, and high costs. The latest addition to the list of programmable nucleases is CRISPR/Cas9, which has revolutionized the field of genome engineering due to its high efficiency, low costs and being more user-friendly than previous technologies.

CRISPR/Cas – an adaptive immune system in prokaryotes

About three decades ago, short palindromic repeat sequences with unknown function were identified in several bacterial and archaea genomes [277-279]. Subsequently, these repeats were named as clustered regularly interspaced short palindromic repeats (CRISPR) and several CRISPR-associated (Cas) genes close to the CRISPR loci were identified [280]. Interestingly, the spacer sequences between the repeats were homologous to genetic elements in bacteriophages and conjugative plasmids [281]. The presence of these specific sequences in the genome of phages prevent them from infecting bacteria containing the corresponding spacer sequences [282]. A key study revealed the presence of an inheritable adaptive immune system in bacteria against phages by showing that the Cas genes mediated the degradation of the viral genome. The degraded protospacer DNA sequences were integrated as new spacer sequences in the CRISPR locus of the bacterial genome [283]. The
spacer units are expressed as short guide CRISPR RNAs (crRNA) which are used to silence invading pathogens containing foreign nucleic acids in a sequence specific manner [284].

Based on the sequences and organization of the CRISPR repeats and CRISPR/Cas loci, the CRISPR/Cas systems are classified into three major types; type I, II and III [285]. The type II CRISPR/Cas system exists exclusively in bacteria and is characterized by the presence of Cas9, a protein with endonuclease activity. Cas9 contains HNH and RuvC-like nuclease domains each of which cleaves one of the strands of DNA. In *Streptococcus pyogenes*, a trans-activating CRISPR RNA (tracrRNA) binds to mature crRNA through base pairing resulting in a duplex structure [286]. This RNA-duplex directs the Cas9 endonuclease to cleave the invading phage-derived DNA at specific sites. The target loci is determined by sequence complementarity between the target protospacer DNA and the crRNA and a motif adjacent to the protospacer referred as the protospacer adjacent motif (PAM) [287]. This results in the degradation of the target DNA thereby conferring *Streptococcus pyogenes* with immunity against certain phages.

**CRISPR/Cas9 system as a genome editing tool**

Successful engineering of the type II CRISPR/Cas system of *Streptococcus pyogenes* into an RNA programmable nuclease was a major innovation in the field of genome editing. This involved the fusion of the tracrRNA and customized crRNA to form a chimeric RNA termed single guide RNA (sgRNA), which provides specificity for the Cas9 nuclease to cleave the target DNA [287] (Figure 6a). Subsequently, the CRISPR/Cas9 system was applied on mammalian cells to create double strand breaks in the DNA in a site-specific manner [288, 289]. There are two major pathways for repairing double-strand DNA breaks – Non-homologous end joining (NHEJ) and Homology-directed repair (HDR) (Figure 6b). The usage of the two DNA repair pathways depends on the organism, cell type and phase of the cell cycle. NHEJ-mediated repair of double-strand break is more common in vertebrate cells compared to less complex eukaryotes [290]. The process of NHEJ is error-prone and can result in the introduction of insertions/deletions (indels). HDR mediated DNA repair can be used to perform more precise genome edits such as point mutations or insertion of pieces of DNA. This is achieved through recombination of the target locus with exogenous donor DNA template. Apart from Cas9, variants of the Cas12 nuclease has also been engineered for genome editing applications [291-293].
Figure 6. CRISPR/Cas9 genome editing tool

a) The CRISPR/Cas9 system has been engineered by the fusion of crRNA and tracrRNA resulting in a synthetic sgRNA which combines with the Cas9 protein to form the Cas9-sgRNA complex. In this complex, the sgRNA guides the Cas9 nuclease to the target the DNA site to induce a double strand break. b) NHEJ and HDR are the two DNA repair mechanisms that could be triggered following the cleavage by Cas9.

Applications of the CRISPR/Cas9 system

Due to the versatility and efficacy of the CRISPR/Cas9 system, it has a wide range of applications. The most common use of the CRISPR/Cas9 technology is the disruption of the open reading frame of coding genes thereby inhibiting their protein expression. In this process, random mutations are introduced as indels through NHEJ-mediated repair at the site of the double-strand break. Alternatively, specific mutations such as single nucleotide polymorphism (SNP) can be introduced through HDR-mediated repair. The CRISPR/Cas9 system can also be used to introduce large deletions or to engineer chromosomal rearrangements [196, 294] by simultaneous cleavage of DNA at two genomic locations. Apart from these, CRISPR/Cas9-mediated gene knockout can also be performed in a multiplexed manner in high-throughput loss-of-function screens [295, 296]. In this thesis, we have used the CRISPR/Cas9 technique both for simple gene knockout and to perform in vivo and in vitro functional screens (Papers I, III).
To utilize the DNA binding capacity of the CRISPR/Cas9 system without its nuclease activity, a mutant form of Cas9, termed dead-Cas9 (dCas9) was engineered, which lacks the ability to induce double-strand breaks [297]. Tethering of various protein effectors to dCas9 opened up a new set of applications for the CRISPR/Cas9 technique. These include modulation of gene expression by fusing dCas9 to transcriptional repression or activation domains [298]. CRISPR/dCas9 can also be used to induce epigenetic changes [299], study chromatin interactions [300] and imaging specific genomic loci [301] by tethering appropriate effector proteins to dCas9.
Functional screens

In humans and other complex organisms, biological function is determined at the cellular level through different biological processes including cell division, proliferation, differentiation, apoptosis, nutrient transport etc. These processes are mediated by different entities such as genes, proteins and other molecules. A major challenge in understanding the cellular function of these entities is to identify their roles in different biological processes. Functional screens are used to determine the effect of a number of entities in specific biological processes by assaying the change in phenotypes associated with these processes [302, 303]. As the read-out of a functional screen is based on the phenotype of interest, functional screens are often termed as phenotypic screens. In general, cell-based functional screens involve two main steps – i) cell perturbation and ii) assessment of change in phenotype.

Cell perturbation

In cell biology, perturbation refers to the alteration of the functioning of a cell through external means [304]. A common method of perturbation in functional screens is by treatment of cells with molecules such as chemical compounds, peptides and cytokines. For example, in paper II, we have performed a functional screen by treating murine AML cells with a library of cytokines. Alternatively, cell perturbation can also be achieved through modulation of gene expression which has been useful in establishing links between individual genes and specific biological phenomena [303]. The two common forms of genetic screens are gain-of-function and loss-of-function screens. Gain-of function screens involve the ectopic expression of genes by introducing a library of corresponding cDNAs or open reading frames (ORF) into cells [305]. Loss-of-function genetic screens have traditionally been performed by inhibiting the expression of genes through RNA interference (RNAi). Although RNAi is a useful tool for inhibiting gene expression, it has major limitations in terms of its off-target effects and low knock-down efficiencies [306]. Currently, loss-of-function genetic screens are often performed using CRISPR/Cas9 genome editing technique due to its low off-target effects and higher inhibition of gene expression compared to RNAi. In line with this, we have performed CRISPR/Cas9 mediated loss-of-function screens in papers I & III.
Assessment of change in phenotype

The choice of the phenotype to be assessed depends on the type of cell perturbation and the biological question that is being addressed using the functional screens. Cell viability is the most common phenotype assessed in functional screens as it allows for the identification of entities that modulate cell growth. This strategy has been particularly effective in identifying active compounds or genes that are critical for the survival of cancer cells [307, 308]. In papers I & III, we have assessed cell viability as a measure of gene functionality. Cell viability could also be used as a surrogate phenotypic read-out to address more complex biological questions. For example, in paper II, we have established an ex vivo functional cytokine screen where the assessment of leukemia cell viability in vivo is a surrogate phenotypic read-out of the effect of cytokines on the leukemia initiating capacity of the cells. Other common phenotypes assayed in functional screens include proliferation, apoptosis, cell cycle, cell signaling and differentiation [303].

Arrayed vs pooled functional screens

Cell-based functional screens are typically performed either in an arrayed format or in a pooled format. In arrayed well-by-well format, a single known perturbation is applied to the cells in each well and the phenotype of interest is assayed in individual wells. Although it allows for the assessment of complex and multi-parametric read-outs, an arrayed format is usually expensive, labor intensive and requires specialized equipment for liquid and plate handling to increase reproducibility. By contrast, pooled screens are rapid, cost-effective and allows for high-throughput analysis [309]. In pooled screens, cell perturbations are performed using molecular barcoded reagents which allow for multiplexing of the assay (see Barcoded-pooled samples).
Molecular barcodes

One of the important aspects of any assay is the throughput it offers, i.e., the number of attributes that can be simultaneously measured in the assay. Advancements in automation of laboratory hardware such as robotic liquid handling systems have allowed for multiplexed approaches which increase the throughput of the assays thereby reducing the time and cost. Similarly, recent innovations with the molecular barcoding technology has enabled the application of molecular tags that greatly increase the throughput of assays with a molecular read-out [310]. The molecular tags act as unique identifiers which can be used to label individual cells, cell lineages, macromolecules and pooled samples. Commonly used molecular tags are either short DNA oligonucleotides detected by NGS or fluorochromes detected by their spectral emission. Due to the limitations of spectral overlap, fluorochromes offer lower multiplexing capabilities compared to DNA oligonucleotides.

Barcoding of individual cells

A high degree of cellular heterogeneity exists both in healthy and diseased tissues [311, 312]. However, most of our current genomic studies have focused on analyzing bulk tissue samples consisting of a large number of cells. Although effective in assessing the collective effects in these population-averaged datasets, cell-to-cell variations are difficult to resolve. Single-cell sequencing offers a solution to overcome this limitation and provides a better overview of the cellular heterogeneity. DNA barcoding has been instrumental in the development of high-throughput single cell sequencing assays. Advancements in droplet microfluidics technology allow for the capture and barcoding of individual cells in liquid drops [313]. This greatly increases the number of individual cells that can be analyzed simultaneously thereby improving throughput of the sequencing assays.

DNA barcoding has also improved the throughput of lineage tracing, i.e., identification of the progeny of the cells of interest. This has traditionally relied on methods such as direct observations through microscopy [314], radioactive labels [315], dyes [316] and fluorescent markers [317]. A major drawback of these methods is that only a limited number of lineages can be traced in parallel. This limitation has been overcome through molecular barcoding where insertion of unique DNA barcodes into the genome of the cells ensures that it will be transmitted
to the next generation after cell division. Lentiviral DNA barcodes have been used to study cell lineages in the hematopoietic system in vivo [318-320].

Barcoded macromolecules

Molecular barcoding techniques can be used to label macromolecules such as antibodies or chemical compounds. Conjugating antibodies to DNA barcodes enable the simultaneous measurement of transcription and translation of individual cells in methods such as cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) [321], RNA expression and protein sequencing (REAP-seq) [322] and cell hashing [323]. Here, barcoded antibodies bind to cell surface proteins and the cells are subsequently captured for scRNA-seq. The DNA barcodes attached to the antibody allows for the coupling of the immunophenotype to the transcriptome of individual cells.

Screening of chemical compound libraries is performed to identify binding molecules to protein targets of pharmaceutical interest. This process is limited by large library sizes and logistical issues. DNA-encoded chemical libraries (DECL) has enabled the screening of libraries consisting of millions of chemical compounds at moderate costs [324]. DECLs are collections of compounds that are individually conjugated to DNA barcodes thereby enabling multiplexing.

Barcoded pooled samples

Despite the significant reduction in the cost for NGS in recent years, it still remains an expensive assay. Molecular barcoding offers a multiplexing strategy wherein samples from different origins are barcoded and pooled prior to sequencing [325]. While processing the sequencing data, the origin of each sequencing read can be identified by its DNA barcode allowing the demultiplexing of the samples. This considerably reduces the sequencing cost and processing time per sample.

Molecular barcoding also allows for the development of pooled competitive cell growth assays. Pooled loss-of-function short hairpin RNA (shRNA) or CRISPR/Cas9 screens are classic examples of a molecular barcoded competitive cell growth assay [326]. Here, the cells are transduced with retro- or lentiviral shRNAs or sgRNAs which gets integrated in the genome and functions both as a means of genetic perturbation as well as a barcode. Thus, cells with unique molecular barcodes and genetic perturbations are made to compete with each other for cell growth. The change in representation of the barcodes (shRNA or sgRNA)
assessed by NGS is a measure of the effect of the corresponding genetic perturbation on cell growth.

Molecular barcoding could also be used to increase the throughput of functional screens involving non-genetic perturbations. For example, in paper II, we have assessed the ex vivo effect of a library of cytokines on the leukemia initiating capacity of AML cells in vivo. One of the technical challenges in addressing this biological question is that a large number of experimental animals are required to get a meaningful in vivo read-out. Therefore, we used molecular barcoding to multiplex the assay. Unlike lentiviral shRNA/sgRNA, cytokines cannot be used as a molecular barcode. Therefore, the classical pooled approach used in the genetic screens could not be applied. In order to overcome this limitation, we designed a hybrid functional screening assay by combining ex vivo cytokine stimulation of barcoded AML cells in an arrayed format with a competitive in vivo read-out of leukemia initiating capacity of AML cells in a pooled format. Through this multiplexing strategy, the effect of a large number of cytokines on AML cells could be assessed using a relatively small number of experimental animals.
Leukemia has been known to human-kind as a blood disorder since the times of the ancient Greeks in the 4\textsuperscript{th} - 5\textsuperscript{th} century BC [327]. Our understanding of leukemia in general and acute leukemia in particular has improved over time which can be classified in chronological order as the morphological era, the cytogenetic era and the molecular era. Each of these eras have been driven by major technological advancements. The advent of the microscope in the 17\textsuperscript{th} - 18\textsuperscript{th} century was the beginning of the morphological era which provided an opportunity to visually characterize the leukemia cells for the first time. The cytogenetic era commenced with the development of techniques such as karyotyping during the last century which allowed for the visualization of chromosomes. This facilitated the characterization of leukemia as a genetic disease and together with morphological data allowed for classification of the disease into distinct entities. The last two decades have heralded the dawning of the molecular era in which NGS technologies have been instrumental in characterizing the diversity of genetic aberrations in leukemia. Additionally, the recent development of an array of tools for genetic manipulation has revolutionized our ability to investigate the complex molecular regulation of leukemia cells.

As the acute forms of leukemia are associated with poor survival, especially among the elderly population, there is a need to identify new and more effective therapies. To achieve this, it is critical to decipher the molecular networks that sustain acute leukemia. In this thesis, we apply cutting-edge molecular tools such as CRISPR/Cas9 and molecular barcoding in functional screens combined with massively parallel sequencing to identify and characterize key molecular regulators of acute leukemia.

**Objectives**

The specific objectives of the individual studies are as follows:

Paper I - Identification of cell surface receptors critical for the survival of AML stem cells \textit{in vivo}.

Paper II - Identification of cytokines that positively regulate AML stem cells.

Paper III - Identification of key dependencies of \textit{DUX4}-rearranged BCP-ALL.

Paper IV - Deciphering the immune-dependent anti-leukemic role of IL4 in AML.
Results and Discussion

Paper I

**CXCR4 Signaling Has a CXCL12-Independent Essential Role in Murine MLL-AF9-Driven Acute Myeloid Leukemia**

AML is propagated by LSCs that reside in the bone marrow and are resistant to standard chemotherapy [93]. LSCs modulate the bone marrow microenvironment which in turn regulates the LSCs through signaling cues [174]. This interaction between the LSCs and the microenvironment is mediated by cell surface receptors on LSCs. In order to identify cell surface receptors that are biologically important for the survival of LSCs under physiological conditions, we performed a pooled in vivo CRISPR/Cas9 screen targeting 96 cell surface genes that are upregulated in murine MLL-AF9 LSCs. We identified CXCR4 as the top positive cell surface regulator of AML cell growth and survival in vivo.

CXCR4 expression is critical for normal hematopoiesis and plays a key role in the maintenance of HSCs in the bone marrow niche [328]. Although high CXCR4 expression is associated with poor prognosis of AML patients [329], its functional role in AML has remained elusive [330, 331]. Therefore, we selected CXCR4 for follow-up studies. CRISPR-mediated disruption of Cxcr4 resulted in a strong depletion of murine MLL-AF9 AML cells in vivo but only had a mild effect on the cells in vitro. Interestingly, there was no negative effect in the homing of the AML cells to the bone marrow due to the loss of CXCR4 expression as CXCR4 has previously been shown to be critical for the homing of normal HSCs to the bone marrow [332]. However, loss of CXCR4 signaling resulted in the activation of p38 MAPK and NF-κB signaling pathways accompanied by increased oxidative stress and differentiation. This suggests that CXCR4 signaling is dispensable for homing of leukemia cells to the bone marrow but is essential for the maintenance of AML cells in vivo.

CXCL12 is the main ligand for CXCR4 and is expressed by several cell types in the bone marrow. Particularly, CXCL12 expression in endothelial cells and mesenchymal progenitor cells has been shown to be critical for the retention of HSPCs in the bone marrow [42, 333]. However, it is unclear whether CXCL12 also regulates AML cells in the bone marrow. Using transgenic mice with Cxcl12 deletion either globally or specifically in endothelial cells and mesenchymal progenitor cells, we showed that CXCL12 expression in the bone marrow is dispensable for murine MLL-AF9 AML development. These findings indicate that unlike normal HSPCs, AML cells are less dependent on these niches for disease development.
To further validate the role of CXCR4 signaling in AML development, we generated Cxcr4 variants with mutations in codons critical either for down-stream signaling or for CXCL12 binding. Using these variants, we showed that CXCR4 signaling but not CXCL12 binding is critical for murine MLL-AF9 AML development. Additionally, we showed that two other known ligands of CXCR4, MIF and UBIQUITIN, do not regulate AML cells in a CXCR4-dependent manner. This suggests that CXCR4 provides baseline signaling independent of ligand stimulation which supports AML development in vivo.

In conclusion, we performed an in vivo CRISPR/Cas9 screen targeting cell surface receptors and identified key dependencies of murine MLL-AF9 AML cells. Our study reveals a critical role of CXCR4 signaling independent of CXCL12 stimulation in AML cells by protecting them from oxidative stress and differentiation (Figure 7).

Figure 7. Graphical abstract of paper I (adapted from Ramakrishnan et al., [334]).
**Paper II**

**Arrayed molecular barcoding identifies TNFSF13 as a positive regulator of acute myeloid leukemia-initiating cells**

One of the key characteristics of AML is the modification of the bone marrow niche. This includes the dysregulation of cytokines which contribute to the selective growth and survival of LSCs [174]. Cytokines such as CCL3, IL6, TGFβ have previously been shown to play critical roles in different hematological malignancies [335-337]. One of the challenges in screening for regulators of LSCs in the niche is the low-throughput of *in vivo* assays. Using arrayed molecular barcoding, we developed a high-throughput *ex vivo* cytokine screen which allowed us to assess the effect of multiple cytokines on the leukemia cells in a competitive manner with an *in vivo* read-out of leukemia-initiating capacity. With this approach, we studied the effect of 114 murine cytokines on the leukemia-initiating capacity of murine *MLL-AF9* AML cells and identified TNFSF13 as a top positive regulator of LSCs.

TNFSF13 has been shown to have a pro-tumor effect in several solid cancers [338-340] and B-cell malignancies [341-343]. Although elevated levels of TNFSF13 have been reported in AML patients [344], TNFSF13 has not been previously been associated with myelopoiesis or LSCs. Therefore, TNFSF13 was selected for follow-up studies. Measurement of TNFSF13 levels in healthy control and leukemic mice showed that although TNFSF13 was present at physiologically relevant levels in both the groups, it was secreted predominantly by normal myeloid cells and not by AML cells, suggesting that mature myeloid cells support murine *MLL-AF9* AML cell growth by secretion of TNFSF13. Characterization of the hematopoietic cells of *Tnfsf13<sup>-/-</sup>* mice revealed a significant reduction in GMP cells, accompanied by lower numbers of monocytes. Notably, TNFRSF17, one of the receptors of TNFSF13, was highly expressed in monocytes relative to other hematopoietic progenitors in wild-type mice. Taken together, our data suggests that TNFSF13 supports normal myelopoiesis by binding to TNFRSF17.

Next, we assessed the role of extrinsic TNFSF13 in leukemia initiation and progression *in vivo* by retroviral overexpression of *MLL-AF9* in c-Kit<sup>+</sup> *Tnfsf13<sup>-/-</sup>* bone marrow cells followed by serial transplantsations into *Tnfsf13<sup>+/+</sup>* or *Tnfsf13<sup>-/-</sup>* recipient mice. We found a significant reduction in the leukemia burden and increased survival in the *Tnfsf13<sup>-/-</sup>* recipient group. TNFSF13 has previously been shown to elicit an anti-apoptotic effect in B-cell malignancies [345, 346]. Consistent with this, we also observed that TNFSF13 promoted the AML cells by suppressing apoptosis and promoting cell cycle progression. Collectively, our data suggests that TNFSF13 supports leukemia initiation and maintenance *in vivo*.

Additionally, TNFSF13 stimulated the proliferation of several human myeloid leukemia cell lines. Notably, TNFSF13 promoted monomac 6 (MM6) cells by suppressing apoptosis which was consistent with our observation in murine c-Kit<sup>+</sup>
AML cells. The pro-leukemic effect of TNFSF13 on MM6 cells was mediated through its receptor TNFRSF17 and was dependent on the activation of the NF-κB signaling pathway. This is in line with studies showing that TNFSF13 activates NF-κB signaling in B-cell malignancies [342, 343]. Taken together, these findings demonstrate that TNFSF13 also supports the growth of human AML cells in an NF-κB dependent manner.

In summary, we developed an *ex vivo* screening technique using arrayed molecular barcoding and identified a previously unknown role of TNFSF13 as a positive regulator of AML cells (Figure 8). This screening methodology can be extended to other types of *ex vivo* screens where a multiplexed *in vivo* read-out of stem-cell functionality is required.

Figure 8. Graphical abstract of paper II.
A genome-wide CRISPR screen identifies key regulators of DUX4-rearranged BCP-ALL

*DUX4*-rearranged BCP-ALL is characterized by the aberrant expression of DUX4 in B-cell progenitors, leading to leukemic transformation [238-240]. As a newly identified subtype of BCP-ALL, the molecular mechanisms by which DUX4 drives leukemia progression is currently unknown. To identify key dependencies of *DUX4*-rearranged BCP-ALL, we performed genome-wide CRISPR screens in the NALM6 cell line harboring the *IGH-DUX4* rearrangement along with two non-*DUX4*-rearranged BCP-ALL reference cell lines, 697 and REH. By comparing the depletion of sgRNAs in NALM6 relative to 697 and REH, we generated a ranked gene list of positive regulators of *DUX4*-rearranged BCP-ALL.

To identify the genes regulated by DUX4 among the top hits in the screen, we silenced *DUX4* expression in NALM6 cells using shRNA and performed RNA sequencing. Among the top ranked genes in the screen, we identified *FNIP1*, *IRF4* and *SYNCRIP* as differentially expressed. Notably, *FNIP1* and *IRF4* were upregulated due to the knock-down of *DUX4* suggesting that DUX4 either directly or indirectly inhibits these genes. In contrast, *SYNCRIP* expression was downregulated upon *DUX4* knock-down, indicating that DUX4 is a positive regulator of *SYNCRIP*.

*FNIP1* is a metabolic regulator which has also been shown to be critical for B-cell development [347, 348]. Transcriptome analysis of *FNIP1*-disrupted NALM6 cells revealed an enrichment of gene sets associated with mTOR, FOXO and ROS signaling pathways, all of which are associated with cellular metabolism [349-351]. We also observed an enrichment of genes related to the B-cell receptor signaling pathway. Collectively, these findings suggest that loss of *FNIP1* results in metabolic dysregulation and affects B-cell signaling in *DUX4*-rearranged BCP-ALL cells.

*IRF4* is a transcription factor that is a well-known regulator of lymphopoiesis [352, 353] and has previously been associated with CLL and multiple myeloma [354, 355]. Disruption of *IRF4* resulted in upregulation of genes associated with apoptosis, the ROS signaling pathway and B-cell differentiation, suggesting that *IRF4* regulates multiple functions in the leukemia cells that are critical for their survival.

*SYNCRIP* is a ubiquitously expressed RNA-binding protein that plays a critical role in RNA processing [356, 357]. We observed a downregulation of genes associated with RNA processing and protein translation upon loss of *SYNCRIP* expression in NALM6 cells. In addition, we observed a downregulation of the TGFβ-SMAD signaling pathway. *SYNCRIP* has previously been shown to interact with MSI2 and regulate AML stem cells [358], as well as positively regulate the expression of TGFβR1 and SMAD3 in hematopoietic stem cells [359]. Hence, we speculate that
DUX4 induces SYNCRIP expression that interacts with MSI2 leading to activation of the TGFβ-SMAD signaling pathway, which could be responsible for the block in B-cell differentiation.

Taken together, we have performed a genome-wide CRISPR screen and identified \textit{FNIP1}, \textit{IRF4} and \textit{SYNCRIP} as key regulators of \textit{DUX4}-rearranged BCP-ALL (Figure 9). Further studies are ongoing to validate these findings at the protein level and in primary patient samples.

\textbf{Figure 9.} Graphical abstract of paper III.
Paper IV

**IL4 has a dual role in regulating phagocytosis of murine leukemia cells**

IL4 is a pleiotropic cytokine which has been shown to have both pro- and anti-tumor effects and its precise role in cancer is unclear [360]. In a cytokine screen, we previously identified IL4 as a negative regulator of AML that induces apoptosis in the leukemia cells [361]. As IL4 is also known to regulate several immunological processes [362, 363] under physiological conditions, we hypothesized that the anti-leukemic effect of IL4 is partially mediated by immune cells.

To test this hypothesis, we transplanted *MLL-AF9* AML cells overexpressing IL4 into immunocompetent wild-type mice and two strains of immunodeficient recipient mice, NOD/SCID and NSG. We observed an increase in leukemia burden in the immunodeficient mice, indicating a critical role for immune cells in the anti-leukemic effect of IL4. Interestingly, the immunocompetent mice transplanted with constitutive IL4-expressing AML cells also died despite having low levels of leukemia cells in the bone marrow. In addition, we observed an expansion of F4/80+ macrophages in the bone marrow and spleen of these mice and depletion of macrophages in vivo resulted in the elimination of the antileukemic effect of IL4. Notably, the macrophage expansion was accompanied by a reduction in white blood cells, red blood cells and platelets in the peripheral blood. Constitutive expression of IL4 in mice has previously been associated with excessive phagocytosis, decreased blood cell counts and increased mortality [364, 365]. Combined, these findings suggest that although elevated IL4 levels reduce the leukemia burden by activation of macrophages, it also triggers excessive phagocytosis resulting in bone marrow failure.

Next, we studied the effect of IL4 on macrophages by performing RNA sequencing of macrophages with or without IL4 stimulation both *in vitro* and *in vivo*. We observed that IL4 stimulation resulted in the upregulation of several genes associated with activation of macrophages derived from monocytes that is distinct from tumor associated macrophages (TAM). While TAMs which are classically associated to an M2 phenotype are known to suppress the immune system and thereby have an anti-tumor effect, other types of macrophages play a critical role in tumor immune surveillance through phagocytosis [366]. In addition, gene signatures associated with phagocytosis, inflammation and immune activation were enriched in macrophages harvested from mice in the IL4 group. Consistent with these findings, we observed an increased phagocytosis due to IL4 stimulation of murine macrophages in an *in vitro* phagocytosis assay. However, stimulation of human macrophages with human IL4 resulted in reduced phagocytosis of human AML cell lines. This result is in line with previous findings showing that IL4 differentiated human monocytes into anti-inflammatory, alternatively activated macrophages [367]. Thus, our data indicates that while IL4 stimulation boosts murine
phagocytosis due to the expansion of monocyte-derived macrophages, it suppressed phagocytosis by human macrophages.

CD47 is the main inhibitory receptor for macrophages and its upregulation protects AML cells from phagocytosis [368]. Since IL4 stimulation resulted in selective elimination of AML cells by macrophages through phagocytosis, we hypothesized that IL4 downregulates the expression of CD47 on AML cells. In contrast, we found that IL4 induces CD47 expression on AML cells in a STAT6 dependent manner. The IL4-mediated upregulation of CD47 resulted in reduced phagocytosis of AML cells. Interestingly, IL4 also upregulated CD47 transcript expression on normal c-Kit+ bone marrow cells, suggesting that this is a regulatory mechanism that possibly has evolved to protect endogenous cells from phagocytosis in environments with high IL4 levels. Consistent with our finding, blocking of CD47 on AML cells combined with IL4 stimulation of macrophages resulted in increased phagocytosis of AML cells.

In summary, we show that IL4 has opposing roles in murine phagocytosis with direct activation of macrophages thereby boosting phagocytosis while also upregulating CD47 expression on AML cells thereby inhibiting phagocytosis (Figure 10).

---

Figure 10. Graphical abstract of paper IV.
Conclusions and future perspectives

Hematological malignancies are driven by different genetic aberrations which alter the normal biology of blood cells resulting in their uncontrolled growth. Identification and characterization of factors that regulate leukemia cells in a physiological context contributes towards a better understanding of their altered biology. In this thesis, we have employed advanced molecular techniques and identified several key regulators of acute leukemia.

The CXCR4-CXCL12 axis is an example of a tumor-microenvironment interaction that is critical in several types of cancers [369] and has been assumed to be important for AML biology as well. However, we have uncovered a previously unknown critical role for CXCR4-signaling independent of CXCL12 in the maintenance of AML stem cells in vivo (Paper I). These findings suggest that therapeutic strategies targeting CXCR4 signaling in AML cells could be more effective than just the disruption of the CXCR4-CXCL12 axis. A major limitation of this study is that the experiments were performed only using the MLL-AF9 AML mouse model. It remains to be studied whether our findings also extend to other types of AML.

In papers II & IV, we deciphered important roles of the cytokines TNFSF13 and IL4 in the regulation of AML cells. Although we demonstrated that TNFSF13 positively regulates AML cells in a NF-κB dependent manner, we observed that TNFSF13 is also required for normal myelopoiesis. This indicates that there could be challenges in the development of anti-TNFSF13 therapies. However, elevated serum levels of TNFSF13 have been reported in AML patients and shown to be associated with chemoresistance of AML cells [344, 370]. Therefore, TNFSF13 could potentially be utilized as a prognostic marker in AML. For IL4, we revealed the complex regulatory role it plays on murine AML cells and macrophages. IL4 induces murine phagocytosis of AML cells through macrophage activation resulting in a negative regulation of AML cells. At the same time, IL4 inhibits murine phagocytosis by upregulating the expression of CD47 in AML cells resulting in a positive regulation of AML cells. The net effect of IL4 on AML cell survival is determined by the equilibrium of these opposing signals. In a previous study, we have shown that IL4 also induces apoptosis in both murine and human AML cells [361]. Here, we find that while IL4 activates murine macrophages which are the main effector cells of its anti-leukemic activity, it inhibits human macrophages. Therefore, the therapeutic potential of IL4 in AML patients is limited. Nevertheless, the concept of using cytokines that activate macrophages combined with the inhibition of immune checkpoint proteins such as CD47 could potentially be developed as a therapeutic approach for leukemia.

The ongoing work in paper III aims to characterize the dependencies of DUX4-rearranged BCP-ALL. Identification of FNIP1 and IRF4 as regulators of DUX4-rearranged BCP-ALL is of particular interest as both have been shown to be critical
for normal B-cell development. However, transcription of these genes is also inhibited by DUX4 suggesting that the expression levels of FNIP1 and IRF4 need to be at a moderate level for the maintenance of DUX4-rearranged BCP-ALL. In contrast, we demonstrate that DUX4 positively regulates SYNCRIP in DUX4-rearranged BCP-ALL and therefore is either a direct or indirect downstream target of DUX4 that drives leukemia. Further studies are ongoing to explore the mechanism of action of these genes in DUX4-rearranged BCP-ALL. Given that DUX4 is a transcription factor with potential binding sites to the regulatory regions of a large number of genes, there are probably multiple pathways by which DUX4 drives BCP-ALL. It is possible that some of the genes regulated by DUX4 are also critical for other types of BCP-ALL. However, these genes could not be identified in this study as the hits from the CRISPR screen were filtered to identify genes specific for the survival of DUX4-rearranged BCP-ALL.

The unifying aspect of papers I-III in this thesis is the use of high-throughput functional screens featuring CRISPR/Cas9 and molecular barcoding technologies on leukemia cells to identify factors that are critical either for the growth and survival of leukemia cells. The read-out of these screens was based on the change in growth and survival of the cells being assayed. Although effective, this screening strategy does not take into consideration the cellular heterogeneity. Moreover, the initial screens do not provide information about the mechanistic basis for the functional role of the identified factors. With the advancements in single-cell genomics, it is now possible to perform high-throughput functional screens at single cell resolution with a multi-modal read-out. One such innovation is the recently developed screening method termed Expanded CRISPR-compatible CITE-seq (ECCITE-seq) which combines the CRISPR/Cas9 technology with cell hashing, scRNA-seq and CITE-seq techniques [371]. Here, the read-out includes both the transcriptome and cell surface proteins of the assayed single cells, making it a powerful technique to address interesting research questions. For example, therapies have been developed targeting immune checkpoints on the surface of cancer cells. However, the genes and the regulatory networks that control the expression of these genes is currently unclear. Through ECCITE-seq, it would be possible to screen for the genes that modulate the expression of the immune checkpoint genes using the read-out at the transcript and protein level. Apart from this, advanced molecular barcoding techniques could also be used to perform other phenotypic screens with multi-modal read-outs.

In conclusion, this thesis demonstrates the power of high-throughput functional screens to identify key regulators of leukemia cells. The new knowledge generated through this work enhances our understanding of the vulnerabilities of acute leukemia, findings that may translate into new therapies.
Acknowledgements

There are so many great people, salutations to all of them

- Saint Thyagaraju

My PhD years have truly been an enriching life experience. I take this opportunity to express my heartfelt gratitude to everyone who have helped me in this magnificent journey.

First and foremost, I would like to thank my supervisor, Marcus. You are an inspiring scientist, an epitome of positivity and a great human being. I cherish the freedom and the space you have provided me to grow as a researcher. I am truly blessed to have worked with you.

My co-supervisor, Thoas, for accepting me as a shared PhD student and letting me be part of an exciting project. You have inspired me to always look at the big picture and not be afraid of trying new things.

Our collaborators in Sweden, Jonas Larsson and David Yudovich for your contributions in this thesis.

Our collaborators in USA, Ravi Bhatia and Puneet, for being gracious hosts when I visited your lab.

My team, it has been great working with such wonderful people. Pablo, I truly respect your scientific mind and temperament and have no doubts that you will become an excellent scientist. Thank you for all the help throughout these years. Maria, you bring so much warmth to our group and always ready to help. Hope you have enjoyed the first half of your PhD; but the best is yet to come. Somadri, a short conversation with you can cheer anyone up even on a rough day. Thank you for spreading the joy around. Kishan, thanks for helping me out when my hands were full. Katrin, I am amazed by the amount of work you do for all of us. You are a rare find! Marion, you have always been so nice to me even though I was competing with you to get recruited by Marcus 😊. I treasure the memories of all the fun we have had, especially when I “saved” you and Helene from the moose. Calle H, Mia, thank you for being a great support during the start of my PhD. You really made me feel at home in those initial days. I would also like to specially thank everyone in our team for helping me out with mouse work. And all the mice that were sacrificed for this thesis.

Fioretos group, collaborating with you all has been a great pleasure and a learning experience for me. Henrik, thank you for always being there to discuss and clarify my not-so-smart bioinformatic questions. Marianne, you are an endearing person and talking to you always makes me happy. Thank you for all your encouragement to improve my Swedish. Thanks to Maria Askmyr, Helena Å, Calle S, Noelia, Rasmus and Christina for all the help during these years.

Hagström group, Anna, your thorough approach towards science is inspiring. Thank you for all the guidance you have provided for my projects. Mattias, the deep conversations we have about almost

---

2 Thyagaraju – 18th century Telugu poet/saint. He starts his famous poem “Entharo mahanubhavulu” with these iconic words.
anything under the sun are always stimulating. I can’t wait to infect you with another earworm Tamil song. Also, thanks for being an unpaid advisor to me 😊. Ton, you have shown me that one could be a great researcher and have fun at the same time.

My fellow PhDs, Elsa, Axel, Niklas, Jakob, Sofia, Karim, Anders and Naveen, for all the fun we had making strange movies which somehow turned out to be funny. Thanks for being nice co-passengers in this journey. Saskia, start training for lundalöppet already! (P.S- I still run just once a year 😊). Hanna, Louise, Larissa, Natalie and Valeria, Good luck with your PhD and have fun while doing it.

Special thanks to all the technical staff and senior scientists, Linda M, Jenny N, Caroline J, Andrea, Eleanor, Minjun, Qirui, Helena S, Linda HM, Jenny K, Kristina, Geffroy. The labs won’t be functioning without your support and Anette, for fixing everything. Thank you for being there for us.

Amazing PIs of Clinical Genetics. Together, you have created a rich environment to do great science. Kudos to all of you! Of course, thanks to Felix for being the shining light guiding us all.

My mentors before PhD, Kunal Mukhopadhyay, Rajib Bandopadhyay, Smrithi Srivastava from BIT Ranchi, Suresh Kumar Rayala, R. Baskar from IIT Madras, Kedarnath Sastry, Nagaraj Govindappa, Suma Srinivas, Nandini Pradeep, Aparajita Dubey from Biocon, Willem Voncken from Maastricht, Susana chuva desousa lopez from Leiden. Thanks to all of you that I have come this far.

All my friends here in Sweden and all around the world, thank you all for being there for me 😊

என் குடும்பத்திற்கு, தந்கத்துமகற்காற்றும் நன்றி. என் பின்னானலா நான்கு சபண்கள். என் பின்னத்தில் அயைாமல் அக்காைின் அன்பிற்கு ஈடு இகணனயில்கல மகைைி அகமைசதல்லாம் இகறைன் சகாடுத்தைைம் என்பர். என்னுகையைைம், என் மகைைியஸ் இரியா! என்கை நிமிர்ந்து நிர்க்கச் செய்யும் னைர் நீ. உன் துகண இருந்தால் இன்சைாரு PhD செய்யவும் தயங்க மாட்னைன் செய்யட்டுமா? 😊

மாற்றாள் அவர்களுக்கு மிகவும் வேண்டும்! என் பின்னானலா நான்கு சபண்கள். என் பின்னானலா காண்க முழுமையானது! என் பின்னானலா காண்க முழுமையானது! என் பின்னானலா காண்க முழுமையானது! என் பின்னானலா காண்க முழுமையானது!
References


Kato H, Igarashi K. To be red or white: lineage commitment and maintenance of the hematopoietic system by the "inner myeloid". Haematologica. 2019;104(10):1919-27.


Stem Cell. 2015;17(6):675

Dominantly Enforce Stem Cell Programs and Impair Human Hematopoietic Stem Cell Potential


222. Schurch CM. Therapeutic Antibodies for Myeloid Neoplasms-Current Developments and Future Directions. Front Oncol. 2018;8:152.


