Molecular Interrogation and Functional Studies of Acute Leukemia

Axel Hyrenius Wittsten

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended on Friday 13th of October, at 09.00 in Belfragesalen, BMC D15, Sölvegatan 19, Lund.

Faculty opponent
Dr. Ronald W. Stam
Hematological malignancies are defined by their underlying genetic alterations, many of which are used to diagnose patients to classify them to different risk groups that dictate the therapy given. Recent advances in high-throughput sequencing have highlighted the presence of co-occurring genetic lesions and that they may form distinct genetic clones that evolve throughout disease progression. Acute leukemia is a group of diseases affecting either the lymphoid or myeloid lineage in hematopoiesis, resulting in acute lymphoblastic leukemia (ALL) or acute myeloid leukemia (AML). Certain genetic alterations are closely tied to specific leukemia types, while others are more promiscuous. In this thesis, we have used high-resolution genome-wide methods and murine models to study leukemia as a way to increase our knowledge how leukemia arises and best can be treated.

In the first study (Article I) we characterized the genetic alterations in a case presenting with a rare myelodysplastic/myeloproliferative neoplasm, unclassifiable (MDS/MPN-U) that later progressed to AML. Through comprehensive analyses of the MDS/MPN-U and AML samples, we observed that all genetic lesions detected at AML diagnosis were present already at the MDS/MPN-U stage, likely in a similar clonal composition. Further, targeted drug analysis of the AML sample suggested clinically approved drugs from which the patient could benefit at a potential relapse.

Genetic rearrangements of the epigenetic regulator KMT2A (KMT2A-R) often co-occur with activating mutations in genes involved in intracellular signaling. In the second study (Article II) we show that mutations in FLT3 and NRAS significantly accelerate KMT2A-R driven AML onset, even when present in a subclone as exemplified by the FLT3[ITD] mutation. The presence of an activating mutation affected the leukemias transcriptional profiles by further enhancing transcriptional programs previously associated with KMT2A-Rs. Genomic characterization of mouse leukemias unveiled de novo signaling mutations in several mice harboring only a KMT2A-R, emphasizing the importance of such mutations in KMT2A-R leukemogenesis.

KMT2A-Rs occur in both ALL and AML but the molecular and/or biological mechanisms determining the lineage affiliation remain largely elusive for this disease. In the third study (Article III) we demonstrated the that FLT3[ITD]R promote myeloid expansion of KMT2A-R leukemia in primary human cells. We further showed that established KMT2A-R ALL and AML cells displayed expression profiles closely linked to their respective lineage but that these cells still display a certain immunophenotypic plasticity.

Previously, a large portion of pediatric B-cell precursor ALL (BCP-ALL) patients could not be classified to any of the established molecular subtypes. Chromosomal alterations are a hallmark of BCP-ALL and in the last study (Article IV) we employed high-throughput sequencing to define the fusion gene landscape of 195 pediatric BCP-ALL. Besides identifying several novel in-frame fusion genes, we also described two new oncogenic leukemia subtypes. These two subtypes were associated with distinct genetic lesions, including genetic rearrangements of the DUX4 gene and genetic alterations of ETV6 and IKZF1.

Taken together, the work included in this thesis highlights the major impact that specific genetic alterations have on leukemogenesis, and how their autonomous and non-autonomous cooperation influence clonal evolution, disease phenotype, and molecular profiles of the leukemia.

Key words: Acute leukemia, KMT2A, cooperating lesions, clonal relationship, genomic profiling, fusion genes
Molecular Interrogation and Functional Studies of Acute Leukemia

Axel Hyrenius Wittsten

LUND UNIVERSITY

2017
Division of Clinical Genetics
Department of Laboratory Medicine, Lund
Faculty of Medicine, Lund University
Cover Mr. Bojangles by Theresa Kallrath

Copyright Axel Hyrenius Wittsten

Lund University, Faculty of Medicine
Department of Laboratory Medicine, Lund

ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2017
Sometimes science is more art than science
Rick Sanchez
The complexity of the human body is both intriguing and taunting. Discrete changes in the, evolutionary fine-tuned, machineries that controls the cells in our body can lead to diseases that distort the biological functions that we rely on for everyday life. Understanding the causes that underlies these changes will improve our understanding of the human body and increase our abilities to correct damaged molecular processes. Therefore, it is important to characterize diseases in detail and to model them in an accurate fashion.

Our blood system has a vast turnover of cells each day and the different cells in our blood perform many vital tasks including respiration, wound healing, and helps us fight infections. Complex regulatory mechanisms in these cells control the formation of specific cell types, the momentum of their generation, and their survival. Genetic lesions damaging components of these regulatory structures can result in uncontrolled growth of a blood cell and its progenies, that with time overflows the blood system and disturb its normal functions. This is the basis of leukemia. The causative genetic lesions play a fundamental role for the features of the resultant leukemia, and it is therefore important to identify these lesions to help diagnose and stratify patients so that they can be treated in a way that provides the best chance for a cure. The difficulty of leukemia, as for most cancers, is that multiple contributing genetic lesions often coincide and cooperate in the establishment and progression of the disease. Even though significant advances have been made in our understanding of the pathobiology of leukemia, we still lack a full understanding of the molecular interplay between different genetic lesions. It is therefore crucial to model leukemia in experimental models in order understand the crosstalk between the affected regulatory pathways.

This thesis summarizes my attempts to shed light on the molecular and biological interactions, and cooperation between genetic lesions, that underlie leukemia.
# Table of Contents

Original articles .......................................................................................................................... 11
Abbreviations ............................................................................................................................... 13
Background ...................................................................................................................................... 15
  Normal hematopoiesis .................................................................................................................. 15
    Blood lineages and the hematopoietic hierarchy .................................................................... 15
    Regulating factors of hematopoiesis ....................................................................................... 17
  Hematological malignancies ....................................................................................................... 20
    Clonal hematopoiesis of indeterminate potential ................................................................. 21
    Myelodysplastic syndromes .................................................................................................... 22
    Myeloproliferative neoplasms ............................................................................................... 22
    Myelodysplastic/myeloproliferative syndromes .................................................................. 23
    Chronic lymphocytic leukemia .............................................................................................. 24
    Acute leukemia ....................................................................................................................... 25
  Acute lymphoblastic leukemia .................................................................................................. 25
    Genetic alterations in BCP-ALL ............................................................................................... 26
    Genetic alterations in T-ALL ................................................................................................... 27
  Acute myeloid leukemia ............................................................................................................ 28
    Genetic alterations in AML .................................................................................................... 28
  Temporal order of mutations .................................................................................................... 30
    Rise of leukemia ..................................................................................................................... 30
    Progression from pre-leukemic to acute leukemia ............................................................... 31
    Emergence of subclones .......................................................................................................... 32
    From diagnose to relapse ........................................................................................................ 33
  Genetic rearrangements of the \textit{KMT2A} gene .................................................................... 34
    Normal function of \textit{KMT2A} .............................................................................................. 35
    Fusion partners in \textit{KMT2A-R} leukemia ............................................................................ 36
    Target cell of \textit{KMT2A-R} transformation ......................................................................... 38
    The molecular impact of \textit{KMT2A} fusion proteins ............................................................ 39
    Cooperating lesions in \textit{KMT2A-R} leukemia .................................................................... 40
    Mouse models of \textit{KMT2A-R} leukemia ............................................................................ 41
  Present investigation ................................................................................................................. 47
  Results and discussion .............................................................................................................. 48
Original articles

This thesis is based on the following articles:

Article I

Article II

Article III

Article IV
Articles not included in this thesis


Svensson, G., Hyrenius Wittsten, A., Linse, S., Mani, K. **The structural role of N-linked glycans on human glypican-1.** *Biochemistry.* 2011 Nov 1;50(43):9377-87
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>aCML</td>
<td>Atypical chronic myeloid leukemia</td>
</tr>
<tr>
<td>AEP</td>
<td>AF4 family/ENL family/P-TEFb complex</td>
</tr>
<tr>
<td>ALL</td>
<td>Acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>AML</td>
<td>Acute myeloid leukemia</td>
</tr>
<tr>
<td>APL</td>
<td>Acute promyelocytic leukemia</td>
</tr>
<tr>
<td>ATH</td>
<td>AT-hooks</td>
</tr>
<tr>
<td>BCP-ALL</td>
<td>B-cell precursor ALL</td>
</tr>
<tr>
<td>BM</td>
<td>Bone marrow</td>
</tr>
<tr>
<td>BRD4</td>
<td>Bromodomain-containing protein 4</td>
</tr>
<tr>
<td>CAR</td>
<td>C-X-C chemokine 12 abundant reticular</td>
</tr>
<tr>
<td>CB</td>
<td>Umbilical cord blood</td>
</tr>
<tr>
<td>CBF</td>
<td>Core-binding facto</td>
</tr>
<tr>
<td>CBP</td>
<td>CREB-binding protein</td>
</tr>
<tr>
<td>C/EBPα</td>
<td>CCAAT/enhancer-binding protein α</td>
</tr>
<tr>
<td>CHIP</td>
<td>Clonal hematopoiesis of indeterminate potential</td>
</tr>
<tr>
<td>ChIP-seq</td>
<td>Chromatin immunoprecipitation sequencing</td>
</tr>
<tr>
<td>CLL</td>
<td>Chronic lymphocytic leukemia</td>
</tr>
<tr>
<td>CLP</td>
<td>Common lymphoid progenitor</td>
</tr>
<tr>
<td>CML</td>
<td>Chronic myeloid leukemia</td>
</tr>
<tr>
<td>CMML</td>
<td>Chronic myelomonocytic leukemia</td>
</tr>
<tr>
<td>CMP</td>
<td>Common myeloid progenitor</td>
</tr>
<tr>
<td>CN-AML</td>
<td>Cytogenetically normal acute myeloid leukemia</td>
</tr>
<tr>
<td>CTP</td>
<td>Cytosolic translocation partners</td>
</tr>
<tr>
<td>CXCL12</td>
<td>C-X-C chemokine 12</td>
</tr>
<tr>
<td>DOT1L</td>
<td>DOT1-like histone lysine methyltransferase</td>
</tr>
<tr>
<td>DPL</td>
<td>Dual-phenotypic leukemia</td>
</tr>
<tr>
<td>DUX4-R</td>
<td>DUX4 rearranged</td>
</tr>
<tr>
<td>EMP</td>
<td>Erythroid-myeloid progenitor</td>
</tr>
<tr>
<td>ERK</td>
<td>Extracellular signaling-regulated kinases</td>
</tr>
<tr>
<td>FLT3</td>
<td>FMS-like tyrosine kinase 3</td>
</tr>
<tr>
<td>GEP</td>
<td>Gene expression profile</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>Granulocyte macrophage colony-stimulating factor</td>
</tr>
<tr>
<td>GMLP</td>
<td>Granulocyte-monocyte-lymphoid progenitors</td>
</tr>
<tr>
<td>GMP</td>
<td>Granulocyte-macrophage progenitors</td>
</tr>
<tr>
<td>H3K4</td>
<td>Histone 3 lysine 4</td>
</tr>
<tr>
<td>H3K79</td>
<td>Histone 3 lysine 79</td>
</tr>
<tr>
<td>HeH</td>
<td>High hyperdiploidy</td>
</tr>
<tr>
<td>HSC</td>
<td>Hematopoietic stem cell</td>
</tr>
<tr>
<td>HSPC</td>
<td>Hematopoietic stem or progenitor cell</td>
</tr>
<tr>
<td>ILC</td>
<td>Innate lymphoid cell</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
</tbody>
</table>
ITD  Internal-tandem duplication
JMML  Juvenile myelomonocytic leukemia
JNK  c-jun N-terminal kinases
KMT2A  Histone-lysine N-methyltransferase 2A
KMT2A-R  KMT2A rearranged
LIC  Leukemia initiating cell
LMPP  Lymphoid-primed multipotent progenitor
LepR+  Leptin receptor expressing
LSC  Leukemia stem cell
LT-HSC  Long-term hematopoietic stem cell
MAF  Mutant allele frequency
MAPK  Mitogen activated protein kinase
MBL  Monoclonal B-cell lymphocytosis
MDS  Myelodysplastic syndrome
MDS/MPN  Myelodysplastic/myeloproliferative syndrome
MDS/MPN-U  MDS/MPN, unclassifiable
MDS/MPN-RS-T  MDS/MPN with ring sideroblasts and thrombocytosis
MEP  Megakaryocyte-erythroid progenitor
MkP  Megakaryocyte committed progenitor
MLP  Multilymphoid progenitor
MPAL  Mixed phenotype acute leukemia
MPN  Myeloproliferative neoplasm
MPP  Multipotent progenitor
NK  Natural killer
NS  NOD.CB17-Prkdc^csd
NS-B2m  NOD.Cg-Prkdc^csdB2mtm1Unc
NS-SGM3  NOD.CB17/Prkdc^csdTg(CMV-IL3,CSF2,KITLG)1Eav
NSG  NOD.CB17-Prkdc^csdIl2rtm^WJ^/SzJ
NSG-SGM3  NOD.CB17-Prkdc^csdIL2rtm^WJ^/SzJg(CMV-IL3,CSF2,KITLG)1Eav/MloySzj
NTP  Nuclear translocation partner
PAFc  Polymerase Associated Factor complex
Ph  Philadelphia chromosome
PHD  Plant homology domain
PI3K  Phosphoinositide 3-kinase
RS-T  Ring sideroblasts and thrombocytosis
sAML  Secondary AML
SCF  Stem cell factor
SET  Su(var)3-9, enhancer of zeste, trithorax
SNP  Single nucleotide polymorphism
SNV  Single nucleotide variation
T-ALL  T-cell lineage ALL
TPO  Thrombopoietin
trxG  Trithorax-group
WES  Whole-exome sequencing
WGS  Whole-genome sequencing
Background

Normal hematopoiesis

Hematopoiesis (from Ancient Greek; haima, blood and poiesis, to make) is the biological process involving the formation of new blood cells. Hematopoietic cells are a necessity of life and constitute our immune response, oxygen and carbon dioxide transport, coagulation, and cellular waste removal. These processes are carried out by specialized mature blood cells, many of which are short-lived [1]. In order to maintain a continuous homeostasis and replenishment of the blood system, an estimated trillion cells are formed each day in our body [2]. These numbers are further increased during stress, such as blood loss or infections. Early work postulated that the hematopoietic system was arranged as a hierarchy, with all individual lineages originating from a shared ancestor [3,4]. Extensive work has since proven that hematopoietic stem cells (HSC) reside in the apex of this hierarchy, devoted to generating progenitors that successively differentiate along a single or several blood lineages, ultimately providing mature and functional blood cells. The HSC is defined by two main characteristics, the ability to self-renew and the inherent capacity of multipotent differentiation [5].

Blood lineages and the hematopoietic hierarchy

The blood system is typically divided into two branches, a lymphoid- and myeloid lineage. The lymphoid lineage is composed of immune cells that are part of both the adaptive immune system (B-cells and T-cells) and the innate immune system (innate lymphoid cells (ILCs) and natural killer (NK)-cells), whereas the myeloid lineage is comprised of innate immune cells (monocytes and granulocytes), thrombocytes, and erythrocytes (Figure 1). The formation of these cells result from step-wise differentiation descending from HSCs through committed, proliferative and transitional progenitors, a process designed to reduce the risk of accumulating proliferation-induced DNA damage in the mostly quiescent HSCs [6-8]. This functional formation also highlights the need for reduced self-renewal capacity in these highly proliferative progenitors. The mouse has been an indispensable tool for our understanding of hematopoiesis, and although several central biological aspects remain conserved between mouse and human, there still exist fundamental differences caused
Figure 1. The hematopoietic hierarchy
Hematopoiesis can be arranged as hierarchal structure, commonly known as the hematopoietic tree. At the top of this hierarchy is the HSC, a cell with self-renewal capacity and the ability to form all other hematopoietic cells. Upon differentiation, HSCs form committed progenitors that subsequently give rise to mature and functional blood cells.

by traits such as species-related physiological differences and an increased genetic diversity for humans [9-11]. One distinct taxonomic difference is the composition of the blood system, with humans having a majority of myeloid cells and mice having predominantly lymphoid cells [11].

The adult human hematopoietic hierarchy is arranged with the HSC at the top, with the first differentiation step being loss of self-renewal capacity and the emergence of multipotent progenitor (MPP) cells [12]. Clonal multilineage engraftment competence in immunodeficient mice is used to distinguish robust long-term HSCs (LT-HSC) from more transiently engrafting MPPs [12,13]. However, complete assessment of lineage potential is temporally restricted in mouse models of human hematopoiesis given that different mature hematopoietic cell types emerge at different time points and not always persist with time in this system [10]. The next step of differentiation marks the first major lineage bifurcation in hematopoiesis (Figure 1). Previously, this step was seen as the branching of the lymphoid and myeloid lineage by the emergence of a
common lymphoid progenitor (CLP) and a common myeloid progenitor (CMP), respectively [14,15]. However, it is now believed that there exist lymphoid primed progenitors with limited myeloid potential known either as multilymphoid progenitor (MLP) or lymphoid-primed multipotent progenitors (LMPP) as predecessors to CLPs [16,17]. These LMPPs/MLPs are suggested to have the ability to generate monocytes, neutrophils, and dendritic cells and to reside alongside CMPs, which have a more broad myeloid potential (Figure 1). Recently, it was also suggested that a megakaryocyte committed progenitor (MkP), with the potential to generate megakaryocytes and platelets, emerge from MPPs [18].

CLPs are lymphoid restricted and generate T-cells, B-cells, NK-cells, and ILCs [19-22]. Worth noting is that T-cells have been suggested to also arise from a progenitor preceding CLPs [23,24]. As for the myeloid lineage, CMPs have classically been suggested to generate granulocyte-macrophage progenitors (GMP) and megakaryocyte-erythroid progenitors (MEP), which combined generate all mature myeloid cells. However, the existence of such an oligopotent CMP has recently been challenged and instead CMPs, then termed erythroid-myeloid progenitor (EMP), were suggested to only generate basophils, eosinophils, megakaryocytes, and erythrocytes but not monocytes and neutrophils, thus being more restricted than previously believed [25]. This was further corroborated in adult hematopoiesis where single CMPs failed to generate more than one mature lineage, suggesting that CMPs rather comprise a mix of unipotent progenitors [18]. Further, GMPs are now suggested to be a progenitor of monocytes, dendritic cells, and neutrophils and to arise directly from LMPPs/MLPs and not from CMPs [18,25-27]. Also, these lineage paths of progenitors and differentiation are said to differ and change during our life [18]. The assembly of the hematopoietic circuitry will likely be further revised, given its complexity and the continuous improvements in experimental methods assessing lineage potential.

**Regulating factors of hematopoiesis**

The dynamic process by which HSCs progressively differentiate along the distinct blood lineages is complex and mainly driven by transcriptional and epigenetic changes, which are instructed and governed by several critical regulators in a context type of fashion [28]. The demand of highly proliferating and differentiating cells requires strict control mechanisms and subtle defects in this machinery may lead to development of severe blood disorders.

*The HSC niche*

Hematopoietic progenitors are to a large extent guided by external factors, such as cytokines, growth factors, and cell-to-cell signaling [5]. These cell non-autonomous signals are largely provided by specialized microenvironments, known as niches [29].
Within the bone marrow (BM), these niches are proposed to locally retain HSCs in a microenvironment that restricts their differentiation and maintains their self-renewal capacity. Much effort has been focused into defining the signals important for HSC maintenance, which has proposed two distinct sites that harbor and maintain HSCs, namely the endosteal- and perivascular niches. The endosteal niche was the first to be identified based on an observed co-localization of HSCs and the endosteal surface in the BM [30,31]. The contributing stromal cells were suggested to be osteoblasts that supported HSC by providing critical factors such as angiopoietin 1, C-X-C chemokine 12 (CXCL12, also known as SDF-1), thrombopoietin (TPO), and osteopontin [8,32-35]. Later it was shown that primitive HSCs rather associated with sinusoidal blood vessels than osteoblasts and this was thereby posed as an alternate niche [36]. The contributing cellular components of the perivascular niche is mainly endothelial cells of blood vessels, various perivascular cells including CXCL12-abundant reticular (CAR) cells and leptin receptor expressing (LepR+) cells, and sympathetic neural fibers [29]. Of these, endothelial-, CAR-, and LepR+ cells provide CXCL12 and stem cell factor (SCF) at varying levels, two important factors for HSC retention and maintenance [37-39]. Moreover, neural cells provide noradrenaline which regulate CXCL12 levels in the BM in a circadian fashion [40,41]. Given that recent imaging studies provided no confirming evidence of HSC and osteoblasts association, and the lack of effect on HSC maintenance after selective deletion of Scf or Cxcl12 in osteoblasts, the existence of an endosteal niche has been challenged [38,39,42-44].

**Extrinsic factors**

Several cytokines are suggested to be critical for lineage-specificity, however, many of these presumably have pleiotropic effects on both the hematopoietic system and other tissues [45]. The pleiotropic effects of “lineage-specific” cytokines likely reflect varying levels of cytokine receptors on hematopoietic cells, cell-type specific signaling output, interplay between different cytokine pathways, intracellular availability of transcription factors and signaling molecules, and local access to the cytokines.

Nevertheless, certain cytokines have been shown to be able to highly influence lineage commitment. For example, treatment of single GMPs with either granulocyte colony-stimulating factor or macrophage colony-stimulating factor directed them to differentiate into either granulocytes or macrophages, respectively [46]. Likewise, systemically high levels of erythropoietin suppress non-erythroid progenitors and result in erythroid-based lineage output [47]. Similarly, high levels of FMS-like tyrosine kinase 3 (FLT3) ligand expanded lymphoid- and myeloid progenitors and suppressed MEPs [48]. Knock-out of the interleukin (IL)-7 receptor and the IL-2 receptor gamma chain leads to severe impairments of B-, T-, and NK-cells [49-51]. Noteworthy is also that chemical screens have identified chemical compounds that support human HSC potential ex vivo [52,53]. Collectively, this highlights the major biological impact of external factors on hematopoietic cells.
Most cytokines activate mitogen activated protein kinase (MAPK) pathways [54]. Those that do not, includes the transforming growth factor β receptor family, the tumor necrosis factor receptor family, and the G-protein coupled receptors which comprises many chemokine receptors. The MAPK signaling pathway is a vast signaling network that can be subdivided into different groups, with the more notable being: extracellular signaling-regulated kinases (ERKs), p38MAPKs, and c-jun N-terminal kinases (JNKs). MAPK signaling promote monopoiesis as opposed to granulopoiesis in human hematopoietic stem and progenitor cells (HSPCs) through negative regulation of the CCAAT/enhancer-binding protein α (C/EBPA, encoded by CEBPA) [55]. Another cytokine regulated signaling pathway is the phosphoinositide 3-kinase (PI3K)/AKT pathway [56]. As for MAPK, constant AKT signaling also favored myelopoiesis through C/EBPα-dependent mechanisms [57]. Aberrant signaling output from these pathways have a big impact on proliferation and survival of hematopoietic cells and genetic lesions causing such constitutively signaling are common oncogenic events in hematological malignancies (see Hematological malignancies).

Intrinsic factors

The transition in cell fate is generally coupled with changes in gene expression which are orchestrated by transcription factors [58]. These changes are also associated with, and often preceded by, epigenetic changes in regulatory genomic elements. The initial triggering events determining hematopoietic differentiation commitment remains largely elusive and even though external factors may play an instructive role in this process, it has also been suggested that they merely confer non-specific and permissive survival and growth signals. Support for the latter is that differentiation can occur in the absence of cytokines when apoptosis is suppressed through overexpression of Bcl-2 [59-61]. This suggest a stochastic model in which lineage commitment is primarily driven by cell autonomous programs. Regardless, the instrumental role of intrinsic factors in lineage commitment and cell fate is well established.

A number of transcriptional regulators have been shown to maintain HSC homeostasis. Most of our biological understanding of these factors have been attained from mouse models and, although there is a certain discrepancy to the human system, several intrinsic factors are evolutionary conserved [10]. This has been confirmed through several functional studies that have highlighted a number regulatory factors pivotal for human HSC function. For example, overexpression or knockdown of the Polycomb-group gene BMI1, the Notch target HES1, and the transcription factor HLF have proved to influence the function of human HSC in experimental xenograft models [62,63]. More recently, the Cohesin complex, which guides chromatin segregation, has been suggested to regulate human HSC homeostasis [64-66]. Mouse models have also highlighted several intrinsic factors important for the embryonic development of HSCs, including Kmt2a, Etv6, Runx1, Lmo2, and Scl [9]. Importantly, genetic alterations
and/or deregulation of these genes are seen in a majority of hematological malignancies (see Hematological malignancies).

Lineage-associated transcription factors play a key role in cell fate decision and, although complex interplays between different pathways play a fundamental role, certain factors are known to instruct or reflect distinctive lineage commitments. One well established antagonistic interaction is that of PU.1 (encoded by SPI1) and GATA1 in the development of myeloerythroid-, driven by GATA1, and myelolymphoid, driven by PU.1, lineages [67-69]. Although initially suggested as instructive factors in lineage commitment, single-cell tracking have recently rather suggested them to execute or reinforce already established lineage choices, thereby further complicating the process of lineage commitment [70]. So, while certain transcription factors are assigned as essential for B-cell (e.g. EBF1, PAX5, FOXO1, and IKZF1), erythroid (e.g. GATA1 and KLF1), myeloid (e.g. C/EBPα, PU.1, TAL1, GFI1), and T-cell (e.g. HEB, E2A, GATA3, and TCF family factors) maturation, several of these are known to be more promiscuous and important for multilineage differentiation [71-75].

Hematological malignancies

The need of a constant turnover of short-lived mature blood cells requires a highly active and tightly regulated hematopoiesis. Genetic aberrations affecting this regulatory machinery and/or signaling pathways associated with hematopoietic differentiation, proliferation, or survival commonly result in hematological dysplasia or neoplasms [76]. These disorders generally disturb the output from either of the two blood lineages by generating excessive proliferation, abnormal differentiation, or a combination of the two. The perturbed dynamics gradually disturb the residual normal hematopoiesis that often result in anemia, leukopenia, and thrombocytopenia, which clinically presents as fatigue, cachexia, bleedings, and infections. Although recent advancement in molecular genetic and epigenetic technologies has greatly increased our understanding of the underlying lesions in hematological malignancies, a complete and integrated understanding is still lacking. Hematopoietic malignancies is a heterogeneous collection of diseases with differences in their clinical presentation of symptoms and outcome [77,78]. Many of the underlying hematological malignancy-associated mutations are promiscuous, in the sense that they are found in several entities, while others are more linked to a certain disease phenotype. Over the past decades, accumulated cytogenetic and genetic evidence have made it increasingly clear that several structural and numerical chromosomal changes display such pathognomonic qualities in hematological malignancies. This have allowed for their use as diagnostic markers for a subclassification of, otherwise seemingly overlapping, disease entities. This has led to improved risk stratification, treatment, and clinical outcome for many
groups of patients [79]. The advancements of high-throughput sequencing have further markedly improved our understanding of an emerging complexity in the mutational landscape for several hematological malignancies, and the clinical implications for these networks of underlying genetic lesions continues to evolve [76].

**Clonal hematopoiesis of indeterminate potential**

Over recent years, the existence of age-related clonal hematopoiesis of indeterminate potential (CHIP) has become recognized as a common occurrence in older individuals, seen in up to 10% of people older than 65 years of age [80-82]. Large sequencing efforts from tens of thousands of individuals, unselected for hematological parameters or neoplasms, identified the presence of somatic mutations in genes associated with hematological malignancies, with mutations in epigenetic modifiers DNMT3A and TET2 and the Polycomb-group gene ASXL1 accounting for about two thirds of CHIP cases. However, a recent study based on whole-genome sequencing (WGS) from >11 000 healthy Icelandic individuals confirmed the high prevalence of CHIP in older people, but could through statistical modeling of passenger mutations show that most cases of CHIP lacked evident driver mutations [83]. The reason for clonal hematopoiesis in the absence of known driver mutations might be explained by either unidentified and unknown driver genes, epigenetic alterations, or through a natural neutral drift toward clonal hematopoiesis due to an exhausted HSC compartment [83]. Regardless, CHIP-associated mutations are able to clonally expand the multilineage output from a single HSC and CHIP is associated with an increased risk of developing a subsequent hematological malignancy [80,81,83]. Given the high prevalence of leukemia-associated mutations in CHIP and the increased risk of hematological malignancies in individuals with CHIP, CHIP is suggested to act as a pre-malignant state serving as an initial genetic event that is permissive to additional cooperating lesions in downstream progenitors [84-89]. Error-corrected ultra-deep sequencing of older healthy individuals (50-60 years) identified mutations in DNMT3A and TET2 in 95% of the analyzed samples, suggesting that most of us might harbor a pre-malignant clone irrespective of CHIP [90]. Although CHIP may account as an evolutionary explanation for some hematological malignancies, this is likely not the case in all subtypes and future studies will need to decipher the preceding prevalence of CHIP and its causation throughout hematological malignancy entities.
Myelodysplastic syndromes

Much like CHIP, myelodysplastic syndrome (MDS) is a clonal disorder suggested to originate in HSCs, but with an ineffective hematopoiesis displaying morphological dysplasia of the myeloid lineage that clinically results in cytopenia [91-94]. As for CHIP, MDS is mainly a disease of the elderly [93]. The prognosis of MDS varies, with some patients requiring minimal therapy while others succumb to severe complications due to cytopenia. Further, about one-third of MDS progress to secondary acute myeloid leukemia (sAML) upon the acquisition of additional cooperating genetic lesions (Figure 2) [95]. In contrast to CHIP, MDS is a genetically more complex disease with most patients harboring at least two somatic mutations in genes associated with hematological malignancies [96,97]. Given that mutations in the epigenetic regulators DNMT3A and TET2 commonly are found as early events in MDS and the increased risk of developing hematological malignancies when suffering from CHIP, MDS has been proposed to, in certain cases, emerge from CHIP, although this will need further investigation [80,81,91,96]. Cytogenetically, del(5)(q31q33) is the most common chromosomal aberration present in around 15% of patients, which results in haploinsufficiency of a variable number of genes [98]. Other common cytogenetic lesions include del(7q) and +8, seen in 11% and 8% of patients, respectively [98]. Recurrent mutations in genes encoding splicing factors are very common and include SF3B1, SRSF2, and U2AF1 [97]. Other frequently mutated genes include the Polycomb-group genes ASXL1 and EZH2, the transcription factors RUNX1, IRF1, and the tumor suppressor TP53 [96,97].

Myeloproliferative neoplasms

As opposed to MDS, a hematological malignancy that display normal morphology and preserved differentiation but that presents with an excessive proliferation is commonly classified as a myeloproliferative neoplasm (MPN) (Figure 2). MPN comprise a heterogeneous group of disorders that not always display clinical symptoms but that is associated with an increased incidence of thrombosis, myelofibrosis, and with a risk of transformation to sAML (Figure 2) [99,100]. MPN is regarded as a HSC originating disorder with a clonal expansion at the progenitor level, together with a high expansion of mature cells in one or more myeloid lineages that manifest as hypercellularity in the BM and extramedullary hematopoiesis [99,101]. As for all hematological malignancies, genetic alterations are central in the development of MPN and in its risk-stratification [102]. Chronic myeloid leukemia (CML) is an MPN defined by the presence of the t(9;22)(q34.1;q11.2), leading to the formation of the Philadelphia chromosome (Ph) [103], resulting in the causative fusion protein BCR-ABL1 [77]. CML is today quite manageable with the use of specific tyrosine kinase inhibitors that directly target the BCR-ABL1 fusion protein, but can, especially if left untreated or through the acquisition
Myeloid malignancies are caused by genetic alterations in HSPCs. This causes excessive proliferation and/or disturbed differentiation of at least one myeloid lineage. Some myeloid malignancies are associated with one of these features, while others display a combination of the two. AML can arise without any preceding malignancies, but cases with MDS, MPN, and MDS/MPN-U can also progress to secondary AML. This figure demonstrates common features and paths in myeloid malignancies.

of resistance mutations to the specific kinase inhibitors, progress into blast crisis, a state that shares several characteristics of acute leukemia [104]. Besides CML, Ph-negative MPNs includes polycythemia vera (characterized by erythrocytosis), essential thrombocytosis (characterized by thrombocytosis), and primary myelofibrosis (characterized by the accumulation of collagen fibers in the BM) [77]. All these entities commonly harbor mutations leading to constitutively activated signaling that likely result in cytokine independent growth (see Regulating factors of hematopoiesis [77,102,105]. The driver mutations in MPN, seen in ~90% of patients, all affect the JAK-STAT signaling pathway and most commonly occur in JAK2 and CALR and more seldom the thrombopoietin receptor gene MPL [105-108].

**Myelodysplastic/myeloproliferative syndromes**

Hematological malignancies with unique clinical features overlapping with MDS and MPN, i.e. the concurrent overlap of dysplasia and an excessive expansion of mature cells in one or more myeloid lineages, are classified as a group of disorders which are termed MDS/MPN (Figure 2) [77]. Several genetic alterations associated with other
myeloid neoplasms have to be excluded in order to meet a diagnosis of MDS/MPN, such as BCR-ABL1, PCM1-JAK2, and gene fusions involving, PDGFA, PDGFRB, and FGFR1 [77,109]. MDS/MPN is comprised of juvenile myelomonocytic leukemia (JMML), chronic myelomonocytic leukemia (CMML), atypical chronic myeloid leukemia (aCML), MDS/MPN-unclassifiable (MDS/MPN-U), and MDS/MPN with ring sideroblasts and thrombocytosis (MDS/MPN-RS-T) that all arise in late adulthood, with the exception of JMML which is a RASopathy seen in infants and young children [77]. As a RASopathy, a majority of JMML patients have somatic mutations in KRAS, NRAS, or the RAS regulatory genes CBL, NF1, or PTPN11 [110]. PTPN11 mutations induce a JMML-like disease in mouse models and, interestingly, through transgenic mouse models it has been shown that a Ptpn11 mutation in non-hematopoietic cells in the BM was sufficient to induce a hematological MPN disease through inflammatory processes [111,112].

For remaining MDS/MPN groups, aCML have recurrent mutations in SETBP1, ETNK1, KRAS, or NRAS, CMML have recurrent mutations in SRSF2, TET2, ASLX1, RUNX1, or SETBP1, MDS/MPN-RS-T is mainly associated with mutations in SF3B1, and MDS/MPN-U has been described to harbor recurrent mutations in JAK2, KRAS, and NRAS [113-120]. As for both MDS and MPN, MDS/MPN frequently progress to sAML. However, this is more commonly seen for JMML and CMML as compared to aCML and MDS/MPN-RS-T, with MDS/MPN-U being too heterogeneous to distinctly evaluate [109].

**Chronic lymphocytic leukemia**

Hematological malignancies also occur in the lymphoid branch of hematopoiesis (Figure 1). Accumulation of clonal mature B-cells are generally classified as chronic lymphocytic leukemia (CLL) and is the most common type of leukemia in adults in Western countries [78,121]. Mature T-cell and NK-cell neoplasms also exist but are rare disorders [122]. Scrutinizing work has shown that virtually all cases of CLL arise from monoclonal B-cell lymphocytosis (MBL), a precursor disease that is similar to CLL but does not fulfill the defining CLL criteria [78,123,124]. CLL is a slow-growing disease that usually do not require any upfront treatment, but involves active management of symptoms such as infections, chronic immune deficiency, autoimmune complications, and continuous evaluation for progression to secondary acute leukemia [125]. However, the clinical management of CLL is changing given a number of new approved chemotherapeutic and targeted drugs [126]. Chromosomal aberrations are common in CLL and include del(13)(q14), del(11q), trisomy 12, and del(17) [127]. More than half of CLL patients harbor del(13)(q14) which has been suggested to disturb miRNA 15a and 16-1 which is proposed to influence CLL leukemogenesis [128,129]. The most frequently mutated genes in CLL includes SF3B1, ATM, TP53,
**Acute leukemia**

Various hematological malignancies, as mentioned above, have the risk to progress into secondary acute leukemia (from Ancient Greek; leukos, white and haima, blood) (Figure 2). There is also an increased risk of developing therapy-related secondary acute leukemia as a consequence of chemotherapy and/or radiation therapy for prior neoplasms [132]. However, acute leukemia that develops without any prodrome or known therapeutic exposure is designated *de novo* acute leukemia.

In acute leukemia, genetic aberrations in HSPCs cause excessive proliferation, arrest in differentiation, and the subsequent accumulation of non-functional immature blast cells. The blast cells display phenotypic characteristics of hematopoietic progenitor cells and the leukemia is classified based on their associated hematopoietic branch, i.e. acute lymphoblastic leukemia (ALL) or acute myeloid leukemia (AML). In some rare cases, acute leukemia cells display lineage markers of both lymphoid and myeloid cells and can then be classified as mixed phenotype acute leukemia (MPAL) [133]. Below, the characteristics of ALL and AML will be discussed in more detail.

**Acute lymphoblastic leukemia**

ALL is a disease that occur in both children and adults with an estimated 150 new cases diagnosed each year in Sweden (www.cancerfonden.se). However, a majority of cases are below 20 years of age with a peak of prevalence between the age of 2 and 5 years [134]. Improved treatment regimens have resulted in an overall survival that is close to 90% in childhood ALL, whereas the overall survival for infants (<12 months of age) and adults (18-60 years of age) are poorer and varies with treatment protocols, being around 55% and 35-59%, respectively [135-137]. ALL can be of B-cell precursor (BCP-ALL) or T-cell lineage (T-ALL), with T-ALL being a less common but more aggressive entity, accounting for about 15% of childhood and 25% adult ALL cases [138]. ALL is further comprised of several entities commonly defined by distinct somatic genetic alterations [139]. Common genetic alterations in ALL include aneuploidy (i.e. numerical chromosome changes), chromosomal rearrangements (either deregulating gene expression or, more commonly, creating chimeric fusion proteins), gains or deletions of genomic bases, and single nucleotide variations (SNVs) in the genome [140]. The prevalence of various genetic alterations alters with age and the
presence of a specific genetic alteration is a crucial factor for the diagnosis and prognosis in patients [77,138].

Genetic alterations in BCP-ALL

The most common genetic alterations in childhood BCP-ALL include t(12;21)(p13;q22), generating the ETV6-RUNX1 fusion, and high hyperdiploidy (HeH, 51-67 chromosomes) each accounting for around 25-30% of cases, and both are generally associated with a favorable prognosis (Article IV) [141-143]. These alterations are considerably less frequent in adults (Figure 3) [144]. The mechanism by which ETV6 and RUNX1 are transcription factors important for normal hematopoietic differentiation and it has been hypothesized that ETV6-RUNX1 serves as an initial genetic event that generates a pre-leukemic clone susceptible for secondary transforming events [145,146]. The most common secondary event involves the loss of the non-rearranged ETV6 allele (77% of cases), but also includes other cytogenetic and submicroscopic deletions [146-149]. The pathogenic impact of the numerical changes in HeH is not clear, although it is likely linked to the resultant changes in gene dosage which would help explain the non-random gains of certain chromosomes [150,151]. In adults, Ph-positive BCP-ALL constitutes around 24% of cases but it is more uncommon in children, only accounting for an estimated 3% (Article IV) [144]. Chromosomal aberrations involving 11q23 that results in KMT2A (previously MLL) rearrangements (KMT2A-R) accounts for a majority (79%) of BCP-ALL in infants, but is also seen in children (7%) and adults (15%) (Article IV) [144,152]. KMT2A-Rs are discussed in more detail below (see Rearrangements of the KMT2A gene). Other well-known chromosomal abnormalities in BCP-ALL are t(1;19)(p23;q13) and der(19)t(1;19)(p23;q13) resulting in the fusion gene TCF3-PBX1 [134,153]. TCF3-PBX1 is associated with a more mature pre-B phenotype and was initially associated with poor prognosis but new treatment regimens have improved the outcome of these patients [154]. Both TCF3 and PBX1 are important transcription factors in lymphopoiesis and TCF3-PBX1 upregulate the expression of components in the pre B-cell receptor whose signaling is suggested to play a part in the oncogenic process [155]. Other recurrent, but less common cytogenetic alterations in BCP-ALL includes near-haploid (21-31 chromosomes), low hypodiploid (32-39 chromosomes), high hypodiploid (40-43 chromosomes), and near-diploid (44-45 chromosomes) [134].

Until recently, an estimated 25% of childhood and even more of adult BCP-ALL, lacked a risk stratifying cytogenetic alteration [156]. However, recent genome-wide studies have been able to establish new subtypes of BCP-ALL that often exhibit distinct gene expression profiles (GEPs) and diverse, often cryptic, genetic alterations. Of these,
the most prevalent includes leukemias that are Ph-negative but that display GEPs similar to that of Ph-positive BCP-ALL, appropriately termed Ph-like or BCR-ABL1-like BCP-ALL [161-164]. Ph-like BCP-ALL is seen from 10% in childhood to 27% in adult BCP-ALL and has, in 91% of cases, been shown to harbor genetic alterations that activate kinase signaling [165]. Additionally, we and others recently showed that deregulation of the double-homeobox transcription factor DUX4, mainly seen through juxtapositioning to the IGH locus, is seen in 3-7% of childhood and young adult BCP-ALL (Article IV) [166,167]. These cases also commonly harbor intragenic deletions of ERG. Further, we recently discovered a new subtype comprising 3% of our cohort of pediatric BCP-ALL, termed ETV6-RUNX1-like, with GEP similar to cases harboring ETV6-RUNXI without carrying the actual fusion but that instead harbored deletions of ETV6 and IKZF1 (Article IV). Importantly, these newly identified genetic alterations can be used to aid in the diagnostic process and to risk classify patients.

**Genetic alterations in T-ALL**

Around half of cases with T-ALL harbor rearrangements that juxtapose the T-cell receptor loci (mainly TRA and TRD at 14q11 and TRB at 7q34) to central transcription factors in normal T-cell development, such as TAL1, TLX1, TLX3, LMO2, and MYB. Two other hallmarks of T-ALL include deletion of the CDKN2A/CDKN2B loci,

---

**Figure 3. Common genomic alterations in BCP-ALL and AML**

Structural- and numerical chromosomal changes are hallmarks of leukemia and many serve as important diagnostic markers. Some of these chromosomal changes are disease-specific while others are more promiscuous and the prevalence of different genetic alterations vary with age. These diagrams depict a approximate summation of chromosomal alterations in infant-, pediatric-, and adult BCP-ALL and AML [144,152,157-160].
observed in 70-80% of T-ALL, and activating mutations of NOTCH1, seen in more than 50% of cases [168-170]. A third of patients harbor in-frame fusions with MLLT10-PICALM, KMT2A-MLLT1, and ABL1-NUP214 being the most frequent. Sequence mutations are seen at a mean of 16 per case in T-ALL and the most commonly mutated genes include FBXW7, PHF6, PTEN, USP7, and DNM2 [170].

ETP-ALL is a subtype of T-ALL which constitute around 15% of cases and is characterized by a reduced immunophenotypic expression of T-cell markers (CD1a, CD3, and CD5) and aberrant expression of myeloid and stem cell markers (CD13, CD33, CD34, and CD117) [171]. Genetically, ETP T-ALL is characterized by mutations in active signaling (e.g. NRAS, KRAS, FLT3, IL7R, JAK1, and JAK3), transcription factors of hematopoiesis (GATA3, ETV6, RUNX1, IKZF1, and EP300), or in epigenetic regulators (EZH2, EED, SUZ12, and SETD2) [172].

Acute myeloid leukemia

AML is the most common type of acute leukemia with an estimated 350 new cases diagnosed each year in Sweden (www.cancerfonden.se). As for ALL, AML occur both in children and adults but in contrast to ALL, it is instead associated with older age with 43% of patients being >65 years of age [173]. Albeit incurable 50 years ago, AML still has an overall bad prognosis with cure rates of 35-40% in patients ≤60 years and 5-15% in patients >60 years of age [174]. Whole genome sequencing has emphasized the complex and evolving genetic heterogeneity in AML [158,175]. Similar to other hematological malignancies, the underlying somatic and causative genetic alterations of AML are instrumental in the classification of different AML entities [77].

Genetic alterations in AML

Emerging data assembled from several high-throughput sequencing studies have generated a comprehensive catalogue of AML-associated genes [158,176]. This have led to an increased understanding of how single- and combinations of genetic alterations affect prognosis in AML [158,174,177,178]. Transcriptional and epigenomic studies have also been able to emphasize AML subtypes with prognostic differences based on GEP and DNA methylation patterns [179-181].

Around 53-59% of adult AML patients harbor chromosomal abnormalities of which about a third includes recurrent cytogenetic changes that define specific AML entities [77,158,177]. These include t(15;17)(q24;q21) encoding PML-RARA, t(8;21)(q22;q22) encoding RUNX1-RUNX1T1 (also known as AML1-ETO), inv(16)(p13q22)/t(16;16)(p13;q22) encoding CBFB-MYH11, t(6;9)(p23;q34)
encoding DEK-NUP214, t(9;11)(p22;q23) encoding KMT2A-MLLT3, inv(3)(q21q26)/t(3;3)(q21;q26) which affect GATA2 and MECOM regulation, and t(1;22)(p13;q13) encoding RBM15-MKL1 [77]. Chromosomal changes are more common in younger patients and their frequencies vary with age (Figure 3) [157,158]. PML-RARA is a hallmark of acute promyelocytic leukemia (APL), a disease characterized by an abnormal accumulation of immature promyelocytes. The prognosis of APL has significantly improved after the introduction of all-trans-retinoic acid [182]. Leukemias with chromosomal rearrangements involving CBF and RUNX1 are collectively known as core-binding factor (CBF) AML. These rearrangements perturb the normal regulation of hematopoiesis by CBF and/or RUNX1, creating a pre-leukemic clone [183,184]. Secondary genetic events in CBF AML are likely required for overt leukemia and these likely involve activating mutations in NRAS, KIT, FLT3, KRAS, PTPN11, and/or NF1 in a majority of cases [185]. Beside activating mutations, a clear pattern of co-occurring mutations (including DHX15, ZBTB7A, and ASXL2) have been shown to distinguish RUNX1-RUNXT1 from CBFB-MYH11 AML [185]. Genetic rearrangements involving KMT2A will be discussed in more detail below (see Rearrangements of the KMT2A gene). Of these cytogenetic alterations, t(8;21), inv(16)/t(16;16), and t(15;17) are generally associated with favorable prognosis, t(9;11) generally associated with intermediate prognosis, and inv(3)/t(3.3) and t(6;9) generally associated with adverse prognosis [182,186]. However, several other chromosomal- and genetic abnormalities also influence risk stratification in AML [186].

In addition to chromosomal rearrangements, SNVs are a prominent feature of AML, especially in cytogenetically normal AML (CN-AML). The Cancer Genome Atlas project was the first comprehensive study of AML, i.e. performing WGS and whole-exome sequencing (WES) of a total of 200 patients. This project identified 23 genes that were recurrently mutated, including FLT3, NPM1, CEBPA, DNMT3A, IDH1, IDH2, TET2, and RUNX1 [158]. Small insertions and deletions within these genes are also common in AML, with internal tandem duplications (ITD) in FLT3 being the second most common mutation in AML [176]. A large recent study including over 1500 AML patients was able to unambiguously distinguish 11 classes of AML, with distinct clinical phenotypes and outcomes, solely based on the molecular landscape [176]. Another study identified a specific mutational pattern of sAML and mutations in these genes (SRSF2, SF3B1, U2AF1, ZRSR2, ASXL1, EZH2, BCOR, or STAG2) were present in therapy-related AML or elderly de novo AML these leukemias presented clinicopathological features of sAML [187].
Temporal order of mutations

Acute leukemia constitutes a heterogeneous collection of diseases that are both complex and dynamic despite having a low number of mutations, as compared to other cancers [188]. The complexity originates from the presence of multiple influential somatic driver mutations, existence of coexisting leukemia clones with distinct genetic setups, and their evolvement over time [158,175,176,178,189]. Successive accumulation of mutations is seen as a central step in the development of overt leukemia and typically, one genetic lesion has been believed insufficient for full leukemic transformation. Despite the proposed multi-step process in leukemia, a unifying mutational pattern underlying leukemia development is lacking because of the complex distribution of genetic lesions in leukemia. Also, the need of cooperation lesions is perhaps not always required for the emergence of leukemia, depending on the specific genetic alterations involved and cellular context [190]. In addition, the ability of leukemias to evolve and adapt to conventional and targeted therapy by clonal evolution allowing the emergence of a therapy-resistant subclone and/or the acquisition of additional mutations, is a major clinical challenge.

Rise of leukemia

The term “cell of origin”, here also known as leukemia initiating cell (LIC), is used in leukemia to define the cell in which the first and initiating genetic alteration occurred. Given that additional genetic lesions often are needed for overt leukemia, the LIC might not necessarily be the cell that is actually transformed. Also, the LIC is not the same as leukemia stem cells (LSC) (transformed cells having leukemia maintaining or repopulating capabilities) although they are often used as interchangeable terms, usually given that LSCs are functionally defined by engraftment potential in immunodeficient mice (leukemia initiation). Preferably, however, these terms should be used separately in order to distinguish the initial LIC, and LSCs as a population able to maintain or repopulate leukemia [191].

Although HSCs repeatedly have been suggested as the cell in which the initial genetic event occurs, this might not always be the case for subsequent genetic alterations (see Hematological malignancies). Interestingly, normal HSC behavior change with age, with older HSCs being more myeloid-biased as opposed to immunophenotypically identical HSCs in younger individuals [192]. Therefore, it is possible that, although the LIC in theory is the same regardless of age, older individuals might hold an inherent bias toward developing myeloid hematological malignancies. This could to some extent explain the discrepancy in ALL and AML distribution seen for pediatric and adult cases of acute leukemia. If true, this could also influence if an additional hit is needed and, if so, which cooperating genetic lesions that are needed for overt leukemia.
Initially, human LSCs were viewed as a small population within the heterogeneous bulk of leukemia cells with immunophenotypic characteristics of normal HSCs in AML [193-195]. However, refined methodology later revealed defined LSCs in populations sharing immunophenotype with committed progenitors [196,197]. This was later supported by a large study of 100 AML patients (de novo-, refractory-, relapsed-, and secondary AML) in which LSC populations were traced to either a LMPP-like or GMP-like leukemic population [198]. These LSC populations transcriptionally resembled their normal counterpart, expressed certain self-renewal expression modules, and a retained hematopoietic hierarchy structure in that leukemic LMPPs could give rise to leukemic GMP and not vice versa [198]. Gene expression analysis have identified a “stemness” signature that is shared between LSCs and healthy HSCs and high expression of this signature correlated to poor prognosis [199,200]. Combined, this suggest that either LSCs are committed progenitors that aberrantly acquire self-renewal capacity or that they initially originate from HSCs and progress into LSCs upon faulty differentiation into committed progenitors.

**Progression from pre-leukemic to acute leukemia**

Insight into the order of acquired mutations has been well studied in adult CN-AML. Even before the discovery of CHIP, it has been suggested that mutations in genes that participate in epigenetic regulation serve as an initial event and generate pre-leukemic HSPCs. Early mutations in AML is believed to mainly occur in genes responsible for DNA methylation (e.g. *DNMT3A*, *IDH1/2* and *TET2*) and chromatin modification (e.g. *ASXL1*). These mutations are implied to arise in HSCs, given the multipotency and long durability of these cells over time, and to confer increased fitness over normal HSCs lacking the mutation [86]. However, CHIP clones can also remain stable at relatively low fractions for several years [81]. Further evidence of the emergence of these mutations in HSC is based on the immunophenotype of these cells, their increased fitness in xenotransplantation models, the fact that they remain present in remission, and syngeneic mouse models of affected genes [84-86,201-203]. The pre-leukemic HSPCs are in turn susceptible to additional somatic mutations commonly affecting genes that are involved in proliferative processes, such as *NPM1* and *FLT3*, and arise in or transform committed progenitors [85-87]. The chance of developing AML when harboring a CHIP mutation increases with the size of the pre-leukemic clone and it has been proposed that if the CHIP clone constitutes >10% of the hematopoietic cells the risk of developing AML is 1% per year [81].

Chromosomal rearrangements in acute leukemia are generally seen as early or initial events. Using monozygotic twins and neonatal biosamples, chromosomal rearrangements have been shown to occur prenatailly in pediatric leukemia [204]. Of
these, at least \textit{ETV6-RUNX1} is said to be insufficient to initiate leukemia by itself and rather serve as an initial hit that requires additional genetic lesions (see \textit{Genetic alterations in BCP-ALL}). Other chromosomal rearrangements shown to occur prenatally in pediatric leukemia includes HeH, \textit{RUNXI/RUNXT1} and \textit{KMT2A-R} [205-212]. Interestingly, \textit{KMT2A-Rs} have been suggested to in themselves be sufficient for the development of overt leukemia in infants. The basis would be their short pre-clinical latency, high concordance rate for BCP-ALL in monozygotic twins, and that \textit{KMT2A-Rs} seem to be the only leukemia-wide genetic lesion seen in some patients [190,213]. However, further studies are needed to confirm if this holds true and also the importance of the cellular context.

In adult lymphoid leukemia, CLL-associated mutations have been found in primitive non-lymphoid HSCPs with myeloid potential in CLL patients, indicating the presence of a leukemia-preceding population [89]. Most cases of CLL is considered to be preceded by MBL (see \textit{Chronic lymphocytic leukemia}). Noteworthy, “healthy” HSCs from CLL patients are lymphoid-primed and repeatedly established MBL in xenograft mouse models, which indicate that secondary mutations are required for the progression from MBL to CLL [88].

\textbf{Emergence of subclones}

The emergence of somatic mutations is linked to cell division and DNA replication due to slight intrinsic deficiencies in DNA replication and repair mechanisms. The mutational rate varies regionally throughout the genome and is influenced by factors such as transcriptional levels, chromatin organization, and replication timing [214-217]. Different mutational signatures have been described which are caused by specific DNA maintenance defects or mutagens [188]. In general, many mutations in both ALL and AML are caused by cytosine>thymine transitions, believed to result from spontaneous deamination of 5-methyl-cytosine. This mutation signature have strong correlation with age, implying that these mutations accumulate at a steady rate during life [188]. However, they have also been suggested to be influenced by proliferation rate [218]. ALL is also associated with a mutational signature linked to aberrant activity of APOBEC cytidine deaminases [188]. These enzymes normal function is to convert cytidine to uracil and they are involved in antibody affinity maturation and antiviral response [219].

When a mutation that confers a selective advantage has occurred in leukemia, it creates a population carrying a distinct genetic profile and is then known as a subclone. A classical model of subclone progression follows a linear route in which sequential accumulation of genetic alterations establishes increasingly fit leukemia clones that replace their ancestors through selective sweeps (Figure 4) [220]. However, several high-throughput sequencing studies have highlighted a greater complexity in
the clonal architecture of leukemia, which predominantly display clonal evolution through a branching pattern (Figure 4) [170,175,176,189,213,221-223]. In this scenario, multiple subclones co-exist within the bulk of leukemia cells and possibly compete for ascendancy. Given the motile nature of the hematopoietic cells and lack of the physical confinements seen in solid tumors, a linear evolution would be anticipated, since lack of spatial restriction should enable unrestrained clonal evolution. However, the predominance of branching evolution in leukemia implies more intricate interactions between different leukemia subclones.

**From diagnose to relapse**

Chemotherapy applies a vast amount of stress on leukemia cells and resistance to treatment has been thought to emerge in response to therapy [224]. Although this might hold true in some cases, more recent studies have highlighted the presence of the relapse clone already at the time of diagnosis in both ALL and AML [189,213,222,223]. Despite scrutinizing work comparing diagnostic and relapse samples it has been hard to distinguish any apparent patterns predicting the mutational basis that prevail treatment. Therefore, at diagnosis, a relapse-causing clone can appear either as a small subclone with additional mutations not seen in other clones, or as dominant clone lacking mutations seen in other clones, or vice versa (subclone lacking mutations seen other clones or dominant clone harboring additional mutations not seen in other clones).

Using functional studies, leukemia clones within a single patient has been shown to exhibit distinct differences in phenotype, such as repopulating capacity in immunodeficient mice [225-227]. Further, it was shown through longitudinal ultra-deep sequencing that leukemia subclones respond differently to conventional- and
targeted therapy [228]. A study of eleven AML patients suggested the presence of two different types of relapse causing populations, either a small population of primitive and possibly dormant LSCs, said to share properties with normal HSCs (such as self-renewal capacity and quiescence causing relative resistance to standard chemotherapeutic agents), or a larger bulk of blast cells displaying a transcriptional stemness-signature [87,194,195]. In line with this, an LSC transcriptional signature and engraftment potential in immunodeficient mice, a defining criteria of LSCs, have been shown predict adverse prognosis in patients [200,229,230]. However, evidence for the presence of a defined LSC population is lacking for BCP-ALL. Whereas LSCs in AML have been shown to be enriched in certain immature immunophenotypic populations, prospective “LSCs” in BCP-ALL are found throughout immunophenotypic maturing populations at similar frequencies [231]. This complicates the biology of relapse and suggests that biological mechanisms that go beyond genetics are important, and further studies are needed to decipher the relation between specific genetic aberrations to transcriptional, epigenetic, and functional profiles.

Genetic rearrangements of the KMT2A gene

Epigenetic modifications are important for the establishment and maintenance of transcriptional memory during hematopoiesis [232]. By adding or removing modification of histones and DNA it is possible to control the chromatin architecture and accessibility, and in turn influence gene expression through the recruitment of transcriptional activator or repressor complexes. The importance of epigenetic regulators and their prominent role in hematopoiesis is reflected by the high prevalence of dysregulation of such genes in hematological malignancies.

The Histone-lysine N-methyltransferase 2A (KMT2A) gene encodes a histone methyltransferase, and rearrangements involving KMT2A are commonly seen in ALL, AML, therapy-related AML and MPAL [233]. KMT2A-R are seen in all ages of acute leukemia with the highest frequency in ALL patients younger than one year of age, so called infant ALL (Figure 3) [233]. Infant ALL is characterized by KMT2A-R and an exceedingly poor prognosis [152,234]. Based on monozygotic twins with concordant leukemias having clonal chromosomal breakpoints of their KMT2A-R and the identification of the KMT2A-Rs in Guthrie card samples have provided evidence that KMT2A-Rs can arise prenatally, suggesting a high susceptibility KMT2A-Rs in prenatal hematopoietic cells [205-210].
Normal function of KMT2A

*KMT2A* is a member of the trithorax-group (trxG) of genes that positively influence gene transcription, as opposed to transcriptional repressive polycomb genes, in a manner that is heritable through multiple cell divisions [235]. This enables *KMT2A* to impact cell identity and plasticity and transgenic mouse models have indicated that *Kmt2a* plays an essential role in both the emergence of HSC during embryogenesis as well as in the maintenance of HSPCs throughout adult life [236-241]. The key transcriptional targets of KMT2A includes the *HOX* gene cluster. The *HOX* genes play an essential role during embryonic development, and their transcription is positively regulated by KMT2A-mediated histone 3 lysine 4 (H3K4) methylation [242-244]. Ectopic expression of *Hox* genes can revert the phenotype caused by a lack of *Kmt2a*, as opposed to other *Kmt2a* targets such as *Pitx2* and *Bcl-2*, emphasizing their importance in *Kmt2a* biology [245]. In line with this, loss of *Hoxa9* impair HSPC function [246,247].

KMT2A is a nuclear protein with a complex domain structure which is proteolytically cleaved by Threonine aspartase 1 into two subunits, one N-terminal (KMT2AN) and one C-terminal (KMT2AC) [248,249]. The two subunits are physically associated through intramolecular interactions mediated between the two interaction motifs FYRN, located at KMT2AN, and FYRC, located at KMT2AC (Figure 5). The KMT2AN encompasses several domains involved in DNA binding, including three short AT-hooks (ATH1-3), a Cysteine-n-Cysteine zinc-finger (CxxC) domain, and four plant homology domains (PHD) which also includes a bromodomain [250-253]. As opposed to the KMT2AN, KMT2AC contains domains involved in transcriptional regulation including a domain that recruits the histone acetyltransferase CREB-binding protein (CBP) that is a positive regulator of transcription and the Su(var)3-9, enhancer of zeste, trithorax (SET) domain that is responsible for the H3K4 methyltransferase activity of KMT2A [242,243,254].

The methyltransferase activity of KMT2A requires several additional proteins (WDR5, RbBP5, and ASH2L) which combined creates a large nuclear complex [243,255,256]. However, most of these protein interactions are dependent on sequences that are lost in KMT2A-Rs and therefore lack functional contribution for KMT2A fusion proteins (Figure 5). Additional proteins have also been suggested to aid in the localization of KMT2A to the DNA, with one of the more prominent being MENIN that binds to the very N-terminal of KMT2AN [257-260]. In the absence of MENIN, both wild-type KMT2A as well as KMT2A fusion proteins fail to regulate HOXA9 transcription [258]. Another interaction partner that is preserved between normal KMT2A and KMT2A fusion proteins is Polymerase Associated Factor complex (PAFc). PAFc has been shown to be essential for proper DNA targeting of KMT2A [261,262]. Thus, while large parts of the protein are lost in the formation of the *KMT2A*-R, some of its normal function is still retained.
The KMT2A protein is a histone methyltransferase with multiple conserved domains. After translation, KMT2A is processed through proteolytic cleavage, resulting in a N-terminal and C-terminal subunit which physically associate through the FYRN- and FYRC domains. The genomic breakpoints in KMT2A-Rs are concentrated just before the first PHD domain.

Fusion partners in KMT2A-R leukemia

Although 135 different translocations partners to KMT2A have, as of date, been identified, only eight of these (AFF1, MLLT3, MLLT1, MLLT10, ELL, MLLT4, EPS15, and MLLT11) accounts for >90% of all KMT2A-Rs [233]. Many KMT2A-fusion genes can be found in both lymphoid and myeloid disease, however, some show a clear lineage preference including KMT2A-AFF1 in ALL and KMT2A-ELL in AML (Figure 6). Further, the distribution of different KMT2A-R varies with age, for example with increased frequency of KMT2A with a partial tandem duplication (KMT2A-PTD) in older AML patients and increased frequency of KMT2A-AFF1 in adult ALL (Figure 6). Further, individual KMT2A-Rs can have a predictive impact on prognosis [177,263,264]. In infant KMT2A-R BCP-ALL it is possible to subdivide different KMT2A-Rs based on their GEP [213,265].

Despite the vast number of KMT2A translocation partners, biochemical studies have highlighted two unifying groups; nuclear translocation partners (NTP) involved in transcriptional elongation and cytosolic translocation partners (CTP) likely causing dimerization of the KMT2A part of the fusion protein [266-270]. Of these, NTPs constitute more than 80% of KMT2A-R cases and include AFF1, MLLT3, MLLT1, MLLT10, MLLT6, ELL, AFF4, and AFF3 [233]. This suggest that the loss of proper transcriptional regulation is one of the major molecular mechanisms behind KMT2A-R leukemogenesis. NTPs have been shown to be part of either the DOT1-like histone lysine methyltransferase (DOT1L) complex or the AF4 family/ENL family/P-TEFb (AEP) complex (also known as super elongation complex) that both impact transcriptional initiation and maintenance [266-268,270-272].

The DOT1L complex is mainly composed of DOT1L and the NTPs MLLT1, MLLT3, MLLT10, and MLLT6 [267]. DOT1L is an epigenetic modifier and the only enzyme known to mono-, di-, or trimethylate histone 3 lysine 79 (H3K79) [273]. These
Figure 6. KMT2A fusion partners and their distribution through age and disease
Several fusion partners have been identified in KMT2A-R leukemia, but only a handful of these accounts for a majority of cases. Some fusion partners are restricted to either lymphoid- or myeloid leukemia, while others can be found in both. Further, the frequency of the various fusion partners varies with age and leukemia type [233].

modifications correlate with active transcription and a chromatin state that is protected against histone deacetylase-dependent gene silencing [274-276]. The importance of DOT1L in KMT2A-R leukemogenesis has been shown through genetic ablation of Dot1l which abrogated the leukemogenic potential of KMT2A-MLLT3 and KMT2A-MLLT10 in mouse leukemia models [277-281]. Further, a therapeutic compound targeting the function of the DOT1L complex has shown promising effects in experimental models of KMT2A-R leukemia, and is currently in clinical studies [274,282].

The AEP complex is a multifaceted transactivator that promote both activation and elongation of transcription [283-286]. AEP complex consist of different subcomplexes that involve different NTPs (AFF1, AFF4, AFF3, MLLT1, MLLT3, and ELL) [270,284,287]. Another member of AEP is the bromodomain-containing protein 4 (BRD4), which recognizes acetylated chromatin and facilitates genomic localization and transcriptional elongation of the AEP complex [288,289]. BRD4 is required for KMT2A-R leukemia maintenance and inhibitors targeting BRD4 have shown high efficacy in numerous KMT2A-R AML preclinical models [290,291]. Further, it has recently been shown that the DOT1L complex and AEP complex has interconnected and cooperating functionalities, thereby creating a connection between these two complexes [292,293].

CTPs lack intrinsic transcriptional activity and include ESP15, GAS7, SH3GL1, AFDN, and FOXO4 which together constitutes around ~10% of KMT2A-R cases [233]. Although the CTPs lack any functional similarity, they possess a common structural feature, a coiled-coil domain that facilitates the dimerization of the KMT2A part of the fusion proteins. Removal of the coiled-coil domains in CTPs abolish their oncogenic potential and the CTP-dependent dimerization of KMT2A [294]. However, the first study corroborating the pathological impact of KMT2A dimerization was a knock-in mouse of KMT2A-LacZ which developed acute leukemia [295]. The underlying mechanism was ascribed multidimerization of KMT2A due to the LacZ
gene, which encodes for β-galactosidase that spontaneously di- or oligomerize in solution. In line with this, inducible dimerization of truncated KMT2A has been shown to block hematopoietic differentiation in a reversible fashion and to upregulate expression of KMT2A-R target genes such Hoxa7, Hoxa9, and Meis1 [294,296]. The mechanism through which dimerized KMT2A alters gene expression is still unknown. KMT2A-PTD involves the duplication of selected exons which duplicates the ATH1-3 and CxxC domains and, in contrast to all KMT2A-Rs, the retention of the KMT2A\(^C\) domains (Figure 5) [297]. The transforming mechanism of KMT2A-PTD is largely unknown. However, one study have suggested that the PTD increases the affinity of KMT2A to target loci [296]. Indeed, a knock-in mouse of Kmt2a-PTD displayed high expression of Hoxa7, Hoxa9, and Hoxa10 [298]. In addition, KMT2A-PTD has been suggested to be associated with DNA hypermethylation [299].

**Target cell of KMT2A-R transformation**

Most of our knowledge of the LIC- and LSC biology in KMT2A-R leukemia is derived from mouse models. Forced expression of KMT2A-MLLT1 or KMT2A-MLLT3 has shown that HSCs are not the only target cell for leukemia transformation and that it is sufficient to introduce KMT2A-Rs in GMPs to cause leukemia [300,301]. By combining GEPs from GMP-like LSC and normal HSPCs, it was further shown that the transformed progenitor cells harbored activation of a self-renewal signature [301]. However, transplantation of sorted GMPs expressing endogenous levels of Kmt2a-MLLT3 failed to initiate leukemia [302]. More recently, sorted pure populations of various HSPC from inducible knock-in mice expressing endogenous-like levels of KMT2A-MLLT1 showed that while e.g. GMPs, granulocyte-monocyte-lymphoid progenitors (GMLPs, similar to human LMPPs/MLPs), and CLPs were able to initiate leukemia, HSCs were intrinsically protected against transformation [303]. This is in contrast to a similar study in which KMT2A-MLLT3 was shown to establish disease both from GMPs and LT-HSCs, with a more aggressive disease stemming from LT-HSC [304]. Other studies using less refined methodology have also proposed HSCs as potent LICs in KMT2A-R leukemia [300,302,305,306]. Given that modelling of LIC formation includes ex vivo handling of HSPCs, which likely phenotypically influence the populations of interest, further studies are needed to determine possible LIC origins among different KMT2A-Rs.

The need for differentiation into progenitors seem important for LIC formation. C/EBPα is myeloid transcription factor necessary for the formation of GMPs and genetic knock-out of Cebpa ablated the ability of both KMT2A-MLLT1 and KMT2A-MLLT3 to initiate leukemia [306-308]. Although it still remains elusive whether or not the LIC in KMT2A-R is a HSCs and/or a committed progenitor and if it differs
between KMT2A-Rs, differentiation seem to play a pivotal role in KMT2A-R leukemogenesis.

**The molecular impact of KMT2A fusion proteins**

Numerous large studies have found transcriptional characteristics of KMT2A-Rs in acute leukemia (Article IV) [167, 179, 309-312]. In addition, several studies have confirmed that the key transcriptional targets of KMT2A-Rs are the genomically clustered HOX homeobox genes and their interacting partners MEIS1 and PBX3 [301, 313-316]. Experimental models based on both human- and mouse cells have demonstrated the importance of Hoxa/HOXA and Meis1/MEIS1 expression in KMT2A-R driven leukemogenesis [317-319]. In addition, forced overexpression of Hoxa9 results in enhanced hematopoietic expansion but rarely in leukemia [320, 321]. On the other hand, combined overexpression of Hoxa9 and Meis1 readily induce leukemia similar to that seen in KMT2A-R models [321]. An unbiased study of primary human samples implied that the presence of a KMT2A-R in AML do not result in overexpression of HOXA genes, but rather stabilizes them at levels seen in normal HSPCs [322].

It is unclear if additional KMT2A-R target genes, besides HOX genes, MEIS1, and PBX3, are actually required for leukemia initiation. Chromatin immunoprecipitation sequencing (ChIP-seq) studies have shown that KMT2A-Rs only have about 140-490 confirmed binding targets [281, 323, 324]. This is strikingly low given that normal KMT2A can be found at many or most active promoters [325]. One recent ChIP-seq study detected substantial difference in the overlap between target genes of Kmt2a and KMT2A-MLLT3 and their associated genomic positions, proposing that they are recruited to the DNA by different mechanisms [326]. Interestingly, sustained expression of wild-type Kmt2a is required for the initiation and maintenance of KMT2A-R leukemia, suggesting an intricate cooperation or dependence between the two [327]. The role of a number of potential “non-classical” targets have been assessed through experimental studies and thereby been implemented as contributors in KMT2A-R leukemia initiation. These genes include CDK6, EYA1, JMJD1C, MECOM locus (involving the genes EVI1 and MDS1), and MEF2C [328-334]. EVI1 has further been suggested to be more highly expressed in KMT2A-MLLT3 AML cells established from HSCs as compared to those originating from GMPs and high levels of EVI1 associated with poor prognosis in human AML [305, 335, 336].

The subordinate transcriptional targets of KMT2A-R target genes remain more elusive. However, one gene suggested to be highly upregulated in response to Hoxa9 and Meis1 coordinated transcriptional activity is Myb [337]. Myb is an multifaceted transcription factor that regulates self-renewal and differentiation in hematopoietic HSPCs [338]. Myb has been shown to be required for KMT2A-R leukemogenesis, emphasizing it as
an important oncogenic factor [337,339]. MYB has several transcriptional targets which can be divided into three groups; housekeeping genes, cell-type specific genes, and genes associated with oncogenesis, the latter including MYC [340]. MYC-associated signatures are often observed in KMT2A-R leukemias [341,342]. Overall, the GEP of KMT2A-R leukemias share large similarities to hematopoietic progenitor cells but with an aberrant self-renewal signature [301,305,342].

Cooperating lesions in KMT2A-R leukemia

A number of recent sequencing efforts have increased our knowledge of co-occurring lesions in KMT2A-R acute leukemia [176,213,343-345]. This have revealed that almost half of infant KMT2A-R BCP-ALL and adult KMT2A-R AML cases have mutations in genetic pathways that are part of kinase/PI3K/RAS signaling (hereafter referred to as activating mutations) as determined by WGS, WES, and/or targeted deep sequencing [158,213,346]. The presence of an activating mutation has been suggested to correlate with inferior prognosis in infant KMT2A-R ALL, but not in adult KMT2A-R AML [213,343,344]. Activating mutations are also frequent events in non-KMT2A-R acute leukemia and in pediatric acute leukemia in general, and activating mutations seem to be more common in younger patients as compared to older children of young adults [157,170,176,347]. Mutations in epigenetic regulators other than KMT2A are common in non-infant pediatric KMT2A-R patients (45%), with most of these mutations being found in non-infant KMT2A-R ALL, but such genes are rarely mutated in infant KMT2A-R ALL patients (14%) [213]. However, larger collected studies are needed in order to gain full insight of the full complement of co-occurring genetic lesions in KMT2A-R leukemia and their distribution throughout age and genetic subtype.

In infant KMT2A-R ALL, most of the activating mutations are subclonal, as determined by mutant allele frequencies (MAFs) <30% [213]. These activating mutations are not mutually exclusive but rather, multiple activating mutations at varying MAFs can be found in a single patient, suggesting the presence of multiple low-frequency leukemia clones [213,345]. The high incidence of activating mutations combined with the fact that some KMT2A-R patients carry multiple clones with different activating mutations indicates a high selective advantage for these mutations. When assessing clonal evolution in paired diagnostic-relapse samples in infant and pediatric ALL it has been shown that a leukemia clone containing an activating mutation may increase, be maintained in size, decrease in size, be lost, or gained at relapse [213,222,223]. Little is known if, or how, different clones influence each other in leukemia. However, in solid cancer models, cells from distinct genetic clones have been proposed to non-autonomously influence one another and thereby impact cell proliferation, drug resistance, and metastasis in these tumors [348-352].
Several studies using mouse models of KMT2A-R leukemia have emphasized the cooperative impact of activating mutations in leukemogenesis (see Mouse models of KMT2A-R leukemia), but none has investigated the biological impact of a subclonal activating mutation in KMT2A-R leukemogenesis (Article II). Moreover, relatively little is known with regards to the molecular processes that are dysregulated in the presence different cooperating genetic lesions in KMT2A-R leukemia (Article II and III).

**Mouse models of KMT2A-R leukemia**

Animal models of human disease are indispensable tools for our understanding of pathobiological mechanisms and therapeutic evaluation of existing and novel treatments. The mouse is the leading mammalian model organism given that it is cost-effective, genetically uniform due to inbred animals, sharing 85% similarity to human protein-coding sequences, and has a short life cycle. An applicable mouse model needs to at least recapitulate key aspects of the disease phenotype. Several different models of KMT2A-R leukemia recapitulating important human pathological features have been established. Based on the methodology, these models can be categorized into five groups; transgenic mice expressing KMT2A-Rs, engraftment of virally transduced mouse cells into recipient mice, engraftment of human patient-derived samples into immunodeficient recipient mice, engraftment of human cell lines into immunodeficient recipient mice, and engraftment of virally transduced or genetically engineered primary human cells into immunodeficient recipient mice. Each approach varies in their scope of use and comes with different strengths and weaknesses. Most models have focused on modelling KMT2A-AFF1, KMT2A-MLLT1, and KMT2A-MLLT3, which are among the most common KMT2A-Rs in acute leukemia, but other KMT2A-Rs have also been assessed (Figure 6).

**Transgenic mouse models**

One of the earliest mouse model of KMT2A-R leukemia was the establishment of a transgenic knock-in mouse harboring Kmt2a-AF9 controlled by the Kmt2a transcriptional control elements [353]. Homozygous mice died of embryonic lethality but chimeric and heterozygous mice developed mainly AML with a only few mice presenting with B-ALL [353,354]. Later, more sophisticated conditional translocator models were established of both Kmt2a-Mllt1 and Kmt2a-Mllt3 in which the Kmt2a-R was driven by lineage specific Cre expression [355,356]. When Kmt2a-Mllt1 and Kmt2a-Mllt3 was initiated in primitive HSPCs (Lmo2-Cre), mice developed myeloid leukemia but when they were initiated in T-cells (Lck-Cre) both myeloid- and T-cell malignancies were observed but only for Kmt2a-Mllt1 [355,357]. This both suggests that Mllt1 and Mllt3 are functionally different and that the cell of origin can influence the leukemia phenotype. Further discrepancies between KMT2A-MLLT1 and
**KMT2A-MLLT3** were shown through the establishment of inducible transgenic mouse models. By purifying distinct hematopoietic populations, it was shown that **KMT2A-MLLT3** can initiate disease from both LT-HSC and more committed progenitors (i.e. GMPs), with a more aggressive disease phenotype when induced in LT-HSCs, whereas **KMT2A-MLLT1** only was able to establish disease when induced in committed progenitors (e.g. GMLPs, GMPs and CLPs) and not in primitive cells (LT-HSC and MPPs) [303,304]. Transgenic mice with **Kmt2a-AFF1** have also been established, both a knock-in model and a conditional invertor model [358,359]. Both models resulted in B-cell lymphomas rather than leukemic malignancies as seen for **Kmt2a-Mllt1** and **Kmt2a-Mllt3** [358,359]. However, through engraftment of transgenic cells induced ex vivo to express **Kmt2a-AFF1**, recipient mice developed mainly AML but also pro-B ALL and rare cases with biphenotypic leukemia [360]. A knock-in mouse model of **KMT2A-PTD** has also been established through the introduction of a PTD in the endogenous **Kmt2a** locus, however, these mice do not develop disease but overexpress **Hoxa** genes and exhibit increased numbers of HSPCs [298].

**Syngeneic transduction models**

Several studies have used retroviral delivery of a **KMT2A-R** in order to study its leukemogenic effect in engrafted recipient mice. Among the **KMT2A-Rs** studied are **KMT2A-AFF1/AFF1-KMT2A**, **KMT2A-MLLT1**, **KMT2A-MLLT3**, **KMT2A-MLLT10**, and **KMT2A-ELL** [301,361-365]. Of these five fusion genes, four (**KMT2A-MLLT1**, **KMT2A-MLLT3**, **KMT2A-ELL**, and **KMT2A-MLLT10**) resulted in AML upon direct transplantation of transduced cells [301,361,363,365]. Using pre-culture in B-cell promoting conditions, **KMT2A-MLLT1** was later shown be able to induce B-ALL with myeloid morphology and combined expression of lymphoid- and myeloid specific genes, although this study only included a limited number of mice [366]. **KMT2A-AFF1** leukemia has been difficult to model in vivo, however, when co-expressed together with its reciprocal fusion partner **AFF1-KMT2A** engrafted mice developed pro-B ALL or MPAL with leukemias being biphenotypic for either B/T-cell markers or B/Myeloid cell markers with low penetrance and long latency [362]. Noteworthy, forced expression of solely **AFF1-KMT2A** was also able to initiate disease [362]. Later, expression of **KMT2A-Aff1** was shown stably induce AML in mice, ascribed to higher viral titers when using mouse **Aff1** [367].

Retroviral models have several caveats such as non-physiological gene dosage, risk of insertional mutagenesis, and need for ex vivo manipulation. Knock-in mouse models would therefore seem like a more physiologically appropriate approach. However, such models are time-laborious, phenotypically sprawling, and many fail to recapitulate the leukemia phenotype seen in patients, which at times make them ineffective for directed research questions. It is possible that retroviral models display enhanced ability to transform different strenuous HSPCs and thereby establishing diseases with varying levels of aggressiveness with GEPs that resemble human prognostic subgroups
Also, retroviral mouse models are able to simulate chemotherapy-resistance as commonly seen for KMT2A-R patients [368].

**Xenograft models**

Although mouse models are instrumental for expanding our biological understanding of KMT2A-R leukemogenesis, they lack the correct cellular context. A first attempt to address this employed retroviral delivery of both KMT2A-MLLT1 and KMT2A-MLLT3 in primary HSPCs derived from umbilical cord blood (CB) with subsequent engraftment of NOD.Cg-Prkdc<sup>scid</sup>B2m<sup>tm1Unc</sup> (NS-B2m) immunodeficient mice conditioned by sublethal irradiation [369]. In this model KMT2A-MLLT1 induced exclusively BCP-ALL whereas KMT2A-MLLT3 induced mainly BCP-ALL but with a subset of mice presenting with AML [369]. This is in contrast to what is seen in the syngeneic setting. KMT2A-MLLT3 was also studied in NOD.CB17/Prkdc<sup>scid</sup>(NS) mice designed to express high levels of the human cytokines SCF, Granulocyte macrophage colony-stimulating factor (GM-CSF), and IL-3 (NS-Tg(CMV-IL3,CSF2,KITLG)1Eav; NS-SGM3) and then exclusively induced AML [370]. In that study, none of the NS or NS-B2m mice engrafted with CB cells expressing KMT2A-MLLT3 developed AML [370]. Further, lymphoid- or myeloid ex vivo manipulations were shown to affect the leukemic immunophenotype when engrafted in NS or NS-B2m [370]. Forced expression of KMT2A-MLLT3 in adult HSPCs derived from BM is suggested to skew them into a myeloid lineage in NS or NS-Il2rgtm<sup>1Wj</sup>/SzJ (NSG), although the engraftment potential of these cells is very low [371]. Early attempts of modelling KMT2A-AFF1 in transduced human CB cells failed to initiate leukemia in NSG mice [372]. It was later suggested that the human AF4 cDNA sequence interfered with virus production and by utilizing KMT2A-Aff1 it was possible to efficiently induce BCP-ALL in NSG mice [367]. The BCP-ALL induced by KMT2A-Aff1 resembled pro-B-cells both at the immunophenotype- and molecular level, as opposed to BCP-ALLs induced by KMT2A-MLLT3, which more closely resembled pre-B-cells [367].

Novel approaches using different genome editing techniques to engineer the KMT2A-R directly into the KMT2A locus has recently been performed. The first used transcription activator-like effector nucleases to introduce KMT2A-MLLT1 and KMT2A-MLLT3 in CB cells. Engraftment of these cells in NSG mice resulted exclusively in BCP-ALL for KMT2A-MLLT1 and mainly in BCP-ALL or MPAL for KMT2A-MLLT3 and only rarely in cases of AML [373]. The second approach introduced KMT2A-MLLT1 into CB cells using a CRISPR-Cas9 vector system which resulted in an AML-like disease in NSG mice constitutively expressing the human cytokines SCF, GM-CSF, and IL-3 (NSG-Tg(CMV-IL3,CSF2,KITLG)1Eav/MloySzj; NSG-SGM3), albeit only a limited number of mice were included [374].

Through immunodeficient mice it is also possible to engraft KMT2A-R patient samples. Engraftment of KMT2A-R BCP-ALL samples are usually efficient in NS and
NSG mice and several studies have utilized this as preclinical models to assess the therapeutic potential of novel treatment modalities [375-377]. The establishment of patient derived xenograft repositories of genetically defined samples will likely improve the evaluation of future therapies in KMT2A-R leukemia [378].

**Mouse models involving cooperating lesions**

Given the high prevalence of mutations causing constitutively active signaling in KMT2A-R acute leukemia, many studies have focused on combining KMT2A-Rs with different genetic alterations causing enforced activation of signaling pathways in various mouse models.

Only few studies have explored solely transgenic mouse models when assessing cooperation between KMT2A-Rs and additional genetic changes. One of these, highlighted the impact of Flt3ITD for complete AML transformation in the presence of Kmt2a-PTD [379]. By combining transgenic mice overexpressing KMT2A-AFF1 with mice harboring activated KrasG12D it was shown that their offspring displayed accelerated development of lymphoid leukemia and/or lymphoma [380]. Another study employed the use of the transgenic Kmt2a-MLLT3 mice and combined it with mouse harboring a repressible NRASG12V, which resulted continuous dependency of NRASG12V for AML maintenance [381]. A similar approach, crossed mice expressing inducible levels of KMT2A-MLLT1 with mice having endogenous expression of KrasG12D, which resulted in accelerated development of AML [382].

Several studies have utilized syngeneic transduction models when evaluating the combinatorial effect of specific genetic alterations and a KMT2A-R. These have combined KMT2A-MLLT1 with NRASG12D and FLT3ITD, KMT2A-MLLT3 with FLT3ITD, NRASG12D and PTPN11E76K, KMT2A-SEPT6 with FLT3ITD and NRASG12V, and KMT2A-MLLT10 with PTPN11S503A, all of these models show an accelerated disease process in the presence of an activating mutation [368,383-388]. The use of the transduction methodology has also enabled assessment of genetic cooperation between activating mutations and KMT2A-Rs in human cells. Although a rather unexplored field, co-expression of KRASG12V and KMT2A-MLLT10 enable transformation of CB cells leading to development of AML in immunodeficient mice [389]. The same effect was, however, not seen when combining constitutively active FLT3 with KMT2A-AFF1 in NSG mice [390]. Further, KMT2A-MLLT3 have been suggested to cooperate with both NRASG12D and FLT3ITD and result in reduced disease latency of human CB AML mouse models [391].

Although activating mutations have been extensively shown to cooperate with KMT2A-Rs in many studies, their full biological effect is still not known. KMT2A-R leukemia is mainly a disease caused by improper transcriptional activation, however, the molecular impact of additional mutations is not well established. Further, the
significance of subclonal signaling mutations, often seen in patients, is lacking in mouse models of \textit{KMT2A}-R leukemia. How do such mutations impact leukemogenesis? Also, given the big impact of cooperating lesions and long latency of certain \textit{KMT2A}-Rs raises the need for investigating the potential presence and impact of unknown \textit{de novo} mutations in these models. Improving our knowledge about how genetic alteration cooperate in \textit{KMT2A}-R leukemogenesis has been one main focuses of this thesis (\textbf{Article II} and \textbf{III}).
Present investigation

The focus of this thesis has been to utilize advances in high-throughput sequencing in order to characterize hematological malignancies with high resolution in both primary patient material and in leukemia mouse models. Through the identification of genetic lesions and malignant signatures, we had three main objectives we sought to achieve;

• Describe the causative genetic lesions in a rare primary hematological malignancy and the genetic alterations that were associated with disease progression in order to explore alternative therapeutic strategies (Article I).

• Assess the biological and molecular impact of genetic cooperation between specific leukemia-associated genetic alterations through murine and human experimental model systems (Article II and III).

• Define the gene fusion landscape of pediatric BCP-ALL (Article IV).
Results and discussion

Article I

Genomic profiling and directed ex vivo drug analysis of an unclassifiable myelodysplastic/myeloproliferative neoplasm progressing into acute myeloid leukemia

Hematological malignancies presenting with both myelodysplasia and disproportionate proliferation constitute a separate category in the World Health Organization classification of myeloid malignancies [77]. This category of myeloid malignancies can be further divided into subgroups, but when lacking clear clinical and molecular characteristics they are compiled into an “exclusion” entity known as MDS/MPN-U [77,109]. Although some efforts have been made into defining the mutational pathogenesis of MDS/MPN-U, a unifying signature is lacking due to the rarity and heterogeneity of this complex disease [113,115-117]. In addition, little is known about the genetic changes that are associated with disease progression. In this study, we wanted to define the complement of genetic lesions in a young woman that initially presented with an MDS/MPN-U that rapidly progressed to AML and study the genetic lesions that were associated with clinical progression. Routine clinical genetic diagnostics had failed to identify any underlying genetic lesions and we therefore performed single nucleotide polymorphism (SNP) array and WES to detect somatic copy number changes and SNVs. Both the MDS/MPN-U and AML sample were analyzed in order to establish the genetic lesions associated with progression. Given that the patient presented with two extramedullary relapses at separate occasions after initial AML diagnosis, she was at high risk of a future BM relapse. We therefore assessed the therapeutic potential of a set of clinically approved compounds, based on the patient’s genetic lesions, and performed an ex vivo drug analysis.

Results in short

- We identified and validated a total of 12 coding genetic lesions including \textit{KMT2A}-PTD, \textit{DNMT3A}^{R882H}, and \textit{NRAS}^{G13D}.
- All of the genetic lesions were present already at MDS/MPN-U diagnosis at comparable MAFs as at AML diagnosis.
- \textit{Ex vivo} drug analysis suggested that the patient could benefit from treatment with the proteasome inhibitor Bortezomib or the MEK inhibitor Trametinib.

The patient harbored AML-associated genetic aberrations

Using SNP array, we identified a single copy number change in the AML sample, which was a \textit{KMT2A}-PTD. Validation using RT-PCR revealed that the \textit{KMT2A}-PTD was present already at MDS/MPN-U diagnosis and that it involved an in-frame duplication of exon 2-8. The \textit{KMT2A}-PTD is normally associated with older aged AML patients,
but can also be found in MDS, albeit at lower prevalence [392]. A recent study of 85 KMT2A-PTD patients identified mutations in well-defined oncogenic drivers for >90% of patients and with mutations in the epigenetic regulators DNMT3A, IDH1/2, and TET1/2 in >77% of patients [393]. Activating mutations were also identified in a majority of these patients (68%), but tended to be subclonal [393]. Given that KMT2A-PTD usually persist in relapse and that DNMT3A, IDH1/2, and TET1/2 could be detected at clinical remission in a number of patients, it was proposed that epigenetic mutations are a likely initial event (as for CHIP), with KMT2A-PTD being subsequent early cooperating lesion, and activating mutations being later events [393,394]. In line with this data, we identified an additional 11 non-silent SNVs, including DNMT3A<sup>R882H</sup> and NRAS<sup>G13D</sup> in our patient by WES. All identified mutations were present both in the MDS/MPN-U and AML samples at similar MAFs indicating that the malignant clone that progressed to AML was already established at initial MDS/MPN-U diagnosis. The genetic lesions identified in this study could potentially be used to monitor disease status of the patient.

The AML sample displayed possible sensitivity toward targeted therapy

The patient has had two extramedullary relapses following her AML diagnosis and is therefore at high risk for a future BM relapse. AML cells are difficult to keep alive in culture, but using a culture protocol optimized to maintain AML cells ex vivo, we performed a small targeted drug analysis with FDA approved compounds based on the patients underlying genetic lesions [395,396]. The compounds included two chemotherapeutics initially given as induction therapy (Cytarabine and Daunorubicin), as well as Bortezomib and Voriniostat that have been suggested to confer a high efficacy in infant KMT2A-R B-ALL, Trametinib which has been suggested to exhibit high efficacy in KMT2A-R leukemia carrying mutant RAS, and the hypomethylating agent Decitabine, suggested to improve prognosis for MDS and AML patients carrying DNMT3A mutations [345,397-401]. Our analyses showed that the patient cells were sensitive towards Trametinib and Bortezomib, suggesting that she may benefit from treatment with these compounds at a future BM relapse. Another compound which was not included, not being FDA approved, but that have shown high efficacy in KMT2A-PTD experimental models is the DOT1L inhibitor EPZ004777 [402].

General conclusion

This study highlights the potential clinical benefit of performing comprehensive genomic characterization of the full spectra of genetic lesions in hematological malignancies.
De novo activating mutations drive clonal evolution and enhances clonal fitness in KMT2A-rearranged leukemia

KMT2A-Rs are a common feature of infant ALL, a patient group with particularly poor prognosis. These patients have an overall low frequency of additional genetic lesions, but still, a high frequency of them harbor activating mutations (47%) in kinases/PI3K/RAS signaling pathways [213]. A majority of these activating mutations are found at a subclonal level (MAF <30%), indicating that they are secondary to the KMT2A-R [190,213]. Interestingly, the leukemic clone harboring the activating mutation is often reduced in size at relapse, and sometimes even lost, suggesting that these mutations confer a proliferative advantage during leukemia onset, but that they are not required for leukemia maintenance in the context of KMT2A-Rs [213,344]. We assessed the role of activating mutations in KMT2A-R leukemogenesis by using retroviral bone marrow transplantation assays, in which the KMT2A-MLLT3 fusion gene was combined with either FLT3ITD, FLT3N676K, or NRASG12D. Mutations in these genes are among the most common ones that deregulate signal transduction in AML. Further, to determine the biological impact of subclonal mutations, we optimized the co-transduction in order to establish leukemia subclones containing FLT3N676K. Molecular analyses including gene expressions profiling, quantitative proteomics, and targeted deep sequencing were performed to characterize the resultant mouse leukemias as a way to increase our understanding on how KMT2A-R develop and best can be treated.

Results in short

- Co-expression of FLT3ITD, FLT3N676K, or NRASG12D significantly accelerate KMT2A-MLLT3 leukemogenesis.
- Presence of subclonal FLT3N676K also accelerates disease latency, a majority of these subclones were selected for in secondary recipients.
- De novo mutations in Braf, Cbl, Kras, and Ptpn11 were identified in KMT2A-MLLT3 cells lacking forced expression of an activating mutation.
- Activating mutations enforce KMT2A-R associated gene signatures.
- The pro-inflammatory cytokine Mif increases the survival of KMT2A-R leukemia cells and is upregulated in the presence of an activating mutation.

Activating mutations cooperate with KMT2A-MLLT3

Recipients engrafted with cells expressing either of the activating mutations in combination with KMT2A-MLLT3 succumbed to disease at a significant shorter disease latency than those engrafted with cells expressing solely the KMT2A-R. This is in line with previous studies that have demonstrated a cooperating interplay between KMT2A-Rs and activating mutations [368,379-381,385-387,389]. However, this was
the first time, to the best of our knowledge, that a tyrosine kinase domain mutation in FLT3 have been shown cooperate with a KMT2A-R. Interestingly, when performing secondary transplantations of these leukemic cells all recipients displayed a significantly shorter disease process than their primary counterpart, with the most apparent reduction in latency seen for leukemias expressing only KMT2A-MLL3. This could imply that the impact of an activating mutation is most substantial during disease initiation, rather than for its maintenance.

Subclonal FLT3N676K accelerates AML onset

By optimizing the co-transduction and altering the ratio between the cells expressing KMT2A-MLLT3+FLT3N676K and those expressing only KMT2A-MLLT3, we were able to generate KMT2A-MLLT3 leukemia subclones harboring the FLT3N676K mutation. Engraftment of these cells revealed that a subclonal activating mutation was sufficient to influence disease latency, by accelerating disease onset. Infant ALLs harboring KMT2A-AFF1 exhibit a similar trend, in that an activating mutation confer a worse prognosis irrespective of their clonal size and the presence of an activating mutation also correlates to an average younger age at diagnosis, suggesting that they impact disease latency [213,343]. The same phenomena has been described for CLL patients where subclonal mutations have an impact on prognosis [403-405]. The reason for the reduced latency in the presence of a subclonal activating mutation is unclear, especially given that leukemia clones with activating mutations regularly disappear after treatment and therefore are not the cause of relapse. One explanation could be that the presence of subclone is a risk factor reflecting a more aggressive pathological phenotype, e.g. increased genetic instability, cell-cycle rate, or invasiveness, of leukemia cells. However, experimental studies using mouse and Drosophila melanogaster have shown that distinct solid tumor clones harboring Ras mutations can provide soluble factors that support the growth of separate cancer clones [348-352]. KMT2A-MLLT3 leukemia cells co-expressing FLT3N676K exhibited significantly higher cell-cycle rate, and secondary transplantations revealed that FLT3N676K subclones expanded to clonal dominance in most cases. However, one primary FLT3N676K subclone failed to expand in secondary recipients, suggesting that a competing somatic clone with a selective advantage had ascended from remaining KMT2A-MLLT3 leukemia cells.

KMT2A-MLLT3 cells acquire de novo mutations involved in signal transduction

In order to determine if somatic mutations in relevant genes had occurred during leukemogenesis and clonal evolution in secondary recipients, we performed targeted resequencing of 41 genes, associated with human KMT2A-R leukemia, on 92 mouse KMT2A-MLLT3 leukemias with or without activating mutations. This analysis revealed that de novo activating mutations had occurred in Braf, Cbl, Kras, and in Ptpn11. Only Ptpn11 was recurrently mutated. Notably, all these mutations occurred in KMT2A-MLLT3 leukemia cells lacking a constitutively expressed activating
mutation. This suggests that there is a lack of any further selective advantage for an additional activating mutation in the presence of a strong signaling mutation. In one leukemia with a de novo subclonal Kras<sup>G12D</sup>, which gained clonal dominance in the secondary recipient, fluorescence in situ hybridization analysis revealed a continued clonal evolution at the Kras locus with chromosomal gain of the Kras locus in subclones constituting 40% of the leukemia. This indicates that gene dosage of mutant Kras and/or its ratio toward wild-type Kras conferred increased competitive fitness. The same phenomena was recently shown in a different mouse model of AML which highlighted that increased Kras<sup>G12D</sup> gene dosage further drove leukemic outgrowth, but that loss of wild-type Kras rendered the leukemia more sensitive toward MAPK inhibition, suggesting increased oncogenic dependence on this pathway [406]. One de novo mutation, Cbl<sup>A308T</sup>, occurred in a leukemia with a subclonal FLT3<sup>N676K</sup> mutation that failed to expand in secondary recipients. Instead, the identified Cbl<sup>A308T</sup> had gained clonal dominance upon secondary transplantation, implying that it had outcompeted the FLT3 mutant subclone. Combined, this suggests that the acquisition of de novo activating mutations is limited during high clonal burden of a mutation that confer constitutively active signal transduction, and that changes in mutant gene dosage can further modulate the fitness of a leukemic clone harboring a signaling mutation.

Activating mutations enforce gene programs associated with KMT2A-R

By gene expression profiling and quantitative proteomics we could show that activating mutations further enforce signatures commonly assigned to KMT2A-Rs. These included Myb- and Myc gene programs that previously have been linked to leukemia maintenance and self-renewal in KMT2A-R leukemogenesis [339,341,342,388]. The role of MYC in self-renewal has been intensively debated and it has been suggested that Myc signaling activates dormant HSCs and drives differentiation and proliferation of progenitors [341,407]. We could also show that both the Myb- and Myc signatures were more associated with committed myeloid progenitors rather than HSCs or MPPs and it is therefore possible that these signatures reflect a highly proliferative progenitor state instead of self-renewal.

KMT2A-MLLT3 leukemias lacking an activating mutation, on the other hand, displayed evident expression and translation of genes involved in intracellular signaling pathways. It is possible that high expression of parts of signaling pathways might represent contributing oncogenic event and/or could explain the high prevalence of mutations of such genes in KMT2A-R leukemia, indicating that interference of intracellular signaling might be an effective therapeutic approach also in leukemias lacking an activating mutation [345,399]. Caution is, however, warranted when targeting a specific pathway, given that alternative pathways might exert compensatory signals. Therefore, kinase inhibition might be more beneficial as a combinatorial treatment.
**Mif is upregulated in the presence of an activating mutation**

From the expression data we could show that KMT2A-R leukemias with an activating mutation upregulated the pro-inflammatory cytokine Mif. Previous work have shown that the absence of Mif delays leukemia development in a CLL mouse model [408]. Further, it was recently shown that primary human AML secrete MIF which stimulates stromal cells to secrete IL-8, which in turn promotes survival of AML cells [409]. MIF is also known to suppress P53-induced apoptosis [410,411]. In line with this we could show that MIF improved the survival of KMT2A-MLLT3 leukemia cells *ex vivo*. Therefore, it is possible that elevated levels of Mif are provided by KMT2A-R leukemia cells harboring an activating mutation, and that exogenous Mif supports leukemogenesis of leukemia cells lacking such a mutation.

**General conclusion**

This study demonstrates the immense biological and molecular autonomous and non-autonomous impact of activating mutations in KMT2A-R leukemia. Understanding the mechanistic interplay between genetic alterations will be instrumental for improving the treatment of KMT2A-R leukemia.

**Article III**

**FLT3<sup>N676K</sup> drives myeloid leukemia in a xenograft model of KMT2A-MLLT3 leukemogenesis**

KMT2A-Rs are found in ALL, AML and MPAL, however, the frequency of different fusion partners varies between leukemia type and age. One fusion partner that is found in both BCP-ALL and AML is MLLT3 [233]. Although syngeneic mouse models are able to recapitulate KMT2A-MLLT3 driven AML, this model lack lymphoid potential. Conversely, xenograft models of human KMT2A-MLLT3 leukemia display a high lymphoid bias and only stably exhibit myeloid potential in immunodeficient mice expressing nonphysiological levels of certain human cytokines [367,370]. The cell autonomous and/or non-autonomous cues underlying the decisive rise of either ALL or AML in KMT2A-R leukemia remains elusive. Several primary AML samples have been shown to contain a primitive population similar to LMPPs, suggesting the presence of leukemia progenitors with lymphoid and myeloid potential. Further, immunophenotypically distinct leukemia cells have previously shown to switch lineage affiliation in response to forced exposure to certain lineage factors [367,370]. Increased understanding of the elemental processes determining lineage commitment in leukemia is needed, not the least given recent advancements in immunotherapy targeting lineage markers in leukemia [412]. In this study, we investigated the impact of the FLT3<sup>N676K</sup> mutation in KMT2A-MLLT3 driven leukemogenesis in human hematopoietic cells and characterized the resultant leukemia populations.
Results in short

- FLT3<sup>N676K</sup> drives myeloid expansion in a human KMT2A-R leukemia model.
- CD19<sup>+</sup>CD33<sup>+</sup> KMT2A-R leukemia cells share high GEP resemblance to KMT2A-R ALL cells.
- KMT2A-R ALL cells are more sensitive toward SMAD inhibition compared to KMT2A-R AML cells and sensitizes the cells towards glucocorticoid treatment.
- KMT2A-MLLT3 leukemia cells are immunophenotypically plastic.

FLT3<sup>N676K</sup> drives myeloid expansion of KMT2A-MLLT3 leukemia cells

Engraftment of KMT2A-MLLT3 expressing human CB cells in immunodeficient NS, NS-B2m, or NSG mice commonly results in the establishment of CD19<sup>+</sup> ALL. We could show that concurrent expression of FLT3<sup>N676K</sup> and KMT2A-MLLT3 strongly promoted the expansion of myeloid leukemia cells in human cells in these mouse systems, together with myeloid potential of KMT2A-MLLT3 leukemia cells in NS-SGM3 and NSG-SGM3 mice, suggest that FLT3<sup>N676K</sup> provide either deterministic and/or permissive signals allowing for the outgrowth of myeloid leukemia cells [10,367,369,370].

By separate survival analysis of ALL and AML leukemias, with or without the presence of FLT3<sup>N676K</sup>, we could show that AML samples displayed a significantly more aggressive phenotype in mice. This pattern is not seen for infant leukemia, were KMT2A-R BCP-ALL are associated with a slightly more dismal prognosis than KMT2A-R AML [152,234,413]. This could highlight that infant leukemia, which likely arise in ontogenically younger hematopoietic cells than CB, are phenotypically different to those established in our model system. It could also reflect shortcomings of the mouse as a host in mirroring the human hematopoietic environment. Nevertheless, our data illustrate the biological impact of co-occurring genetic lesions in KMT2A-R leukemia.

Immunophenotypically distinct leukemia populations display specific GEPs

Our xenograft samples enabled the isolation of lymphoid-, myeloid-, and dual-phenotypic (DPL, expressing both lymphoid and myeloid surface markers) leukemia populations. Global gene expression analysis showed that, as expected, ALL and AML cells, respectively, expressed befitting lineage specific genes. Interestingly, DPL cells shared high GEP resemblance with ALL, suggesting that DPL may originate from the same ancestor as ALL and thus aberrantly express myeloid surface markers. Interestingly, all leukemia populations exhibited an evident expression of the myeloid transcription factor CEBPA. Using publically available expression data of normal human populations, CEBPA was seen to be expressed in myeloid cells and most HSPCs with the clear exception being mature B-cells and lymphoid committed CLPs. This
could allow for a certain retention of myeloid potential in KMT2A-R ALL cells, and consequently that DPL are lymphoid cells with aberrant expression of a myeloid surface marker potentially influenced by CEBPα.

The expression data also allowed us to unravel disparities of potential importance in signaling pathway activity between KMT2A-R ALL and AML cells. This highlighted possible differences in SMAD signaling between lymphoid and myeloid leukemia populations, which could be further corroborated in a gene expression dataset of primary infant and pediatric BCP-ALL and AML [213]. The use of a SMAD3 specific inhibitor in cell lines revealed that KMT2A-R BCP-ALL cells displayed increased sensitivity to such treatment as compared to KMT2A-R AML cells. This emphasize the differences in reliance on certain signaling pathways between lymphoid and myeloid leukemia cells. Another well established example of this is the clinical beneficial efficacy of glucocorticoid treatment in BCP-ALL that is not seen in for AML patients [414]. Further, glucocorticoids have been shown to interfere with SMAD signaling and we could show that SMAD3 inhibition enhanced glucocorticoid sensitivity in both sensitive and resistant KMT2A-R BCP-ALL cells [415].

KMT2A-MLLT3 leukemia cells are immunophenotypically plastic
Both KMT2A-MLLT3 and KMT2A-AFF1 cells have been described to possess an inherent lineage promiscuity [367,370]. Using a similar approach of ex vivo manipulation, we could show that BCP ALL, DPL and AML cells all possessed a limited but distinct ability to express a surface marker of the opposite lineage. In contrast to what have been described for human cells transformed in vitro, our in vivo established leukemia cells appeared more lineage restricted, proposing an increased complexity in the regulation of lineage plasticity beyond external factors [370]. Combined with the myeloid expansion seen in the presence of FLT3$^{N676K}$ in vivo, this would suggest that external factors play more of a permissive role in KMT2A-R lineage determination. Deciphering the decisive mechanisms of lineage affiliation in KMT2A-R leukemia will be instrumental for our biological understanding of this disease.

General conclusion
Taken together, this study demonstrates that FLT3$^{N676K}$ can drive myeloid expansion of human KMT2A-R leukemia cells in vivo. It also highlights the molecular differences and similarities between lymphoid and myeloid KMT2A-R leukemia cells and imply distinct biological discrepancies between the two that potentially could be utilized to sensitize such leukemias to glucocorticoid treatment.
Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia

Although a majority of pediatric cases of BCP-ALL harbor established genetic alterations that currently are used for clinical risk stratification, until recently, an estimated 25% could still not be classified into any defined molecular entity. Subdivision of BCP-ALL based on the underlying characterized genetic event have played an instrumental role in the improved treatment of these patients as well as for the development of targeted therapy. Advancements in high-throughput sequencing have allowed for an unbiased and scrutinizing screening of patient samples in search of new classifying biomarkers. Recently, a new subgroup of BCP-ALL was identified through a shared GEP to that of Ph-positive BCP-ALL, which was aptly termed Ph-like BCP-ALL [161,164]. Using mRNA-sequencing we sat out to define the gene fusion landscape of 195 pediatric BCP-ALL.

Results in short

- Identification of 27 novel gene fusions in pediatric BCP-ALL.
- Identification of a novel subtype of pediatric BCP-ALL with similar GEP to ETV6-RUNX1 cases and characterized by coexisting deletions of ETV6 and IKZF1.
- Identification of recurrent rearrangements involving DUX4, constituting 4% of our cohort of pediatric BCP-ALL.

Identification of novel fusion transcripts in pediatric BCP-ALL

High-throughput sequencing of mRNA transcripts from 195 pediatric BCP-ALL identified in-frame fusion transcripts in a high proportion of the cases (65%). Of the remaining cases, all but four could be classified to an established or novel, identified herein, BCP-ALL subgroup, thus enabling classification of 98% of BCP-ALL cases. Most of the cases lacking an in-frame fusion were HeH (94%). Out of the identified and validated in-frame fusions, 27 (counting reciprocal fusions as one) fusions, most of which were non-recurrent, had not been described before. Many of the novel in-frame fusion genes found in cases lacking a genetic stratifying marker involved genes important for B-cell development or that are recurrently rearranged in BCP-ALL. This study emphasizes the significance of chromosomal aberrations in BCP-ALL and highlights the powerful capability of high-throughput sequencing as an unbiased tool for gene fusion identification. The high detection rate, low cost, and quick turnaround time of mRNA sequencing advocate it as a powerful clinical diagnostic approach.
Identification of a subset of leukemias displaying an ETV6-RUNX1-like GEP

A recent advancement in the classification of BCP-ALL was the identification of the Ph-like subtype through gene expression analysis [161,164]. This subtype was shown to confer a poor prognosis and to be characterized by kinase-activating lesions [144,165]. Using comparable clustering approaches, we were able to distinguish a number of leukemias (3%) displaying similar GEPs to those of ETV6-RUNX1 cases. All these cases, termed ETV6-RUNX1-like, were confirmed to lack any molecular support of an ETV6-RUNX1 fusion gene as well as any other stratifying traits. Further comprehensive analysis revealed that this subtype harbored concurrent genetic lesions in ETV6 and IKZF1, two genes with important roles in normal B-cell development. Additional cases of ETV6-RUNX1-like BCP-ALLs have since been described, providing confirming independent evidence on the presence of this subtype [416].

Recurrent DUX4 rearrangements in patients lacking molecular stratification

Among the novel fusion genes, we identified recurrent (4% of total cohort) rearrangements involving the transcription factor DUX4 (DUX4-R). All these fusion transcripts appeared in cases lacking established stratifying molecular markers and composed samples with a unique GEP. This GEP have previously been described to be associated with intragenic ERG deletions [163,417]. In line with this, we could confirm that identified DUX4-R cases commonly harbored ERG deletions (50-63%). Coinciding studies to our, involving pediatric and young adult BCP-ALL, also described the existence of the DUX4-R subgroup. These studies also provided functional evidence of the leukemic potential of a DUX4-R and proposed a functional importance of an alternative ERG transcript in this process [166,167].

General conclusion

This study provides a comprehensive analysis of the fusion genes underlying pediatric BCP-ALL. Through the identification of new genetic subgroups and novel fusion genes, this will help improve current classification of pediatric BCP-ALL and provide new pathogenetic insight that hopefully can improve future treatment of these patients.
Concluding remarks

Genetic lesions are a hallmark of hematological malignancies and the work included in this thesis has focused on the identification and functional characterization of coinciding and/or novel genomic alterations in acute leukemia.

Different genetic alterations often co-occur in hematological malignancies. This coexistence can center among specific genes, but also involve genetic alterations in genes with analogous biological functions. In line with this, we report that targeted treatment of a patient’s underlying mutations could serve as a beneficial clinical treatment option (Article I). By combining selected mutations in experimental models, we have been able to demonstrate and characterize the immense biological and molecular impact of this genetic crosstalk. Molecular signatures originating from these experimental models overlap with patients harboring similar genetic profiles, confirming the validity of the experimental models (Article II). Further, a causative genetic lesion is typically highly associated with the phenotype of the resultant disease. However, the processes underlying lineage commitment in leukemia with ambiguous genetic alterations remain largely unclear. We showed that the presence of an additional mutation was able to influence the lineage of the established leukemia, implying that faulty activation of certain signaling pathways can contribute to lineage determination (Article III).

Finally, we identified novel subgroups of pediatric BCP-ALL that were linked to specific genetic alterations. However, future studies are needed to elucidate the functional effect of these genetic lesions and the importance of the co-occurring intragenic deletions seen in these patients, such as for ERG in DUX4-R BCP-ALL or for ETV6 and IKZF1 in ETV6-RUNX1-like BCP-ALL (Article IV).

This thesis has hopefully contributed to our understanding of the biological and molecular crosstalk between specific genetic lesions in acute leukemia. The work also emphasizes the significance of deep genetic interrogation in patients and experimental models for full biological and diagnostic comprehension.

De genetiska förändringar som orsakar leukemi är tätt kopplade till sjukdomsbilden och används därför kliniskt för att dela upp patienter i olika diagnostiska och prognostiska grupper. Den specifika genetiska förändringen som påvisas vid diagnos har därmed ofta direkt betydelse för vilken behandling patienten ges. Genom den senaste tidens utveckling av avancerade sekvenseringstekniker så är det nu möjligt att läsa av och studera leukemicellernas arvsmassa. Detta har ökat vår förståelse över förekomsten av återkommande genetiska förändringar vid leukemi och vilka kombinationer av förändringar som förekommer. Genom att förstå hur olika genetiska förändringar tillsammans samverkar i utvecklingen av leukemi ökar vi möjligheten att förbättra och utveckla behandlingsstrategier för olika typer av leukemi.

Den övergripande målsättningen för denna avhandling har varit att öka vår förståelse för hur leukemi uppkommer och bäst kan behandlas. I en första studie (Artikel I) använde vi oss av olika typer av dessa högupplösta sekvenseringstekniker för att karakterisera de underliggande genetiska förändringarna hos en patient som upphövde en ovanlig utveckling av AML. Genom att identifiera dessa förändringar kunde vi välja ut läkemedel som var skräddarsydda för just den här patienten, och sedan testa dessa på patientens egna AML celler för att visa att patienten skulle kunna dra fördel av riktad behandling vid eventuellt återfall i sjukdomen.

Det har nyligen påvisats att en stor del av leukemi patienter med genetiska förändringar i KMT2A genen ofta har ytterligare genetiska förändringar i en typ av gener som aktiverar celldelning och förhindrar celldöd, biologiska processer som ofta är felaktigt aktiverade vid cancer. Genom att studera dessa specifika förändringar i KMT2A och i
de "aktiverande" generna i möss, lyckades vi återskapa patienternas sjukdomsbild och kunde se att de genetiska förändringarna samarbetade för att skapa en extra aggressiv leukemi (Artikel II). Förändringarna behövde inte förekomma i samma cancercell, utan detta samarbete kunde ske mellan olika cancerceller som innehöll olika uppsättningar av förändringarna. Detta stämmer väl med vad som har observerats hos patienter, där den genetiska förändringen av KMT2A finns i alla sjuka leukemiceller medan de aktiverande förändringarna oftast endast förekommer i en mindre del av leukemicellerna. Vi kunde även visa att vissa möss, som initialt bara hade KMT2A förändringen, också spontant utvecklade egna genetiska förändringar i aktiverande gener, vilket framhäver hur viktiga dessa aktiverande mutationer är för utvecklingen av leukemi.


Vissa patienter med leukemi saknar de specifika genetiska förändringar som idag används för att bestämma patientens diagnos och behandling. I den sista studien undersökte vi därför leukemiceller från runt 200 barn med ALL med högupplösta sekvenseringstekniker (Artikel IV). Förutom att vi kunde identifiera redan kända genetiska förändringar, så upptäckte vi även två helt nya grupper av ALL. Båda dessa nya grupper kunde kopplas till specifika genetiska förändringar som inte var kända tidigare och utgjorde tillsammans nästan 10% av patienterna. Framtida studier krävs nu för att förstå hur de funna genetiska förändringarna bidrar till leukemiutveckling.

Sammanfattningsvis har studierna i denna avhandling bidragit till en ökad kunskap om hur genetiska förändringar bidrar till leukemiutveckling, hur förekomsten av dem i olika celler kan påverka sjukdomen, samt hur de stör viktiga biologiska processer i cellerna. Förhoppningsvis kan denna informationen bidra till att bättre diagnostiskt klassificera patienter och, i framtiden, till utvecklingen av bättre och effektivare behandling av de patienter som idag fortfarande har en dålig klinisk prognosis.
Acknowledgements

There are many people who have contributed to this thesis, so most importantly, I would like to express my sincere gratitude to everyone who, in one way or another, have been involved in this process.

First of all, I would like to thank my supervisor Anna for taking me in as an untested wild-card-PhD-candidate. I am really glad to have been a part of establishing your research lab in all of its aspects. The amount of freedom and encouragement you have given me to pursue my own ideas is something I will always value. Your enthusiastic and optimistic approach to science is truly inspirational.

Thoas, my co-supervisor, for seeing the bigger picture in things and for the philosophical career talks, even being in the stairwell on a Sunday. I admire your sensible, but forward-thinking, and elaborate approach to science.

The members of Andersson research group. Helena S, what would it have become of me if you had not been there from start to keep track of me. With other colleagues and students coming and going, you have always been the stable and reliable person to go to. Thanks for teaching this confused guy some order. Mattias, your overall enthusiasm to things is so contagious, it can brighten any day. I am looking forward to hear your “Axel” Spotify playlist. Jian for being such a sweet and ambitious coworker. I would also like to thank the past and present students in the lab, especially Cristian for managing to interpret my incoherent instructions with good spirit.

To our collaborators at St. Jude in Memphis for allowing us to utilize their enormous bioinformatic expertise. Jing for being the calm and hospitable solver, Charles and Veronica for guiding a jetlagged Swedish guy with precious samples, and Jim for taking your valuable time to discuss my research.

Jörg and Tália for introducing me to the field and guiding me in my first shaky experimental attempts.

Marcus, how you manage to keep track of everything in the research field and all details of numerous projects is beyond me. I have really appreciated all your advice on both projects and future endeavors. I still hope that you one day will manage to introduce pizza to the journal clubs.
Mia, Sofia, and Niklas for all the road trips, AWs, turtle-sightseeings, stories, ventilating-conversations, dinner parties, concerts, Fernet endeavors, “eating” pictures, and much more, but most importantly for making the “hopplösa rummet” the best place to be in at BMC. You are the best PhD-buddies one can ask for! All hail the single cells and power point gates!

Marianne, you have somehow wrongfully been termed “lab elder”. Not that you are not the wisest person in the lab, but because it does not match your youthful/(childish) spirit. Your cheerful and heartening personality has truly brightened many grey days ☺. Never change! And for the woman with the many and dreaded (albeit relevant) questions, Maria. Your scientific curiosity and hard-hitting jokes is a true inspirational combo 😊. I still hope that we will have a project together someday. I would gladly do the legwork for the launch of “DNA-free H₂O”.

All the other PhDs, past and present, for contributing to both sciencey and leisurely activities. Linda O, Setareh and Elsa for teaching us newcomers how it is done. Jakob for being a true Spurs companion with special chicken wing sauce, Pablo, Anders, and Ram for late-night movie sessions and helpful advice, and also Naveen and Karim for your good spirits.

Henrik for always taking the time answer my clueless questions, even if we are crowd hanging at your door. Calle H for your magical and happy “Gothenburgian” way to get everything fixed. Calle S, for great input on both experiments and popular culture, and for showing me how to make a sensible Margarita order (or…?). Linda M, Jenny N, and Andrea for guiding me when I am running around lost in the lab.

All remaining coworkers on the “bloody” side of the division for all the fruitful discussions, helpful tips and instructions, and joint conference experiences.

The person behind the joyful smile that I pass between the lab and the office, Anette, for always straightening out my messy requisitions and fights with the travel agency.

All the other wonderful co-workers, past and present, at Clinical Genetics for creating the best division with a warmest of atmospheres. I still wished I could have been at the mysterious D15 Xmas party.

Till biomedicinargänget. Oskar, Maria, Julia, Anna, Nathalie och Lotta för ni alltid kryddat min tid både på och utanför BMC med glädje och galenskaper, er energi har alltid varit värld att sträva efter. ”Andra halvan” av biomedicinargänget för alla festliga stunder och för att ni fick mig att inse att Malmö inte är så dumt ändå.

Alla mina vänner utanför BMC som påmint mig att det finns ett liv bortom forskandet. Särskilt Peter, Elin, Robert och Rasmus som jag delat både glädjefulla toppar och lärorika dalar med. Jennifer, Christian, Siri och nu lilla Ester för ni skämt bort mig med otaliga och fantastiska middagar med gott sällskap (och vin… och GT…).

Mor för att du alltid haft mig i fokus och givit mig den mest händelserika och hjärtligaste uppväxten man kan be om, jag klarade mig ganska bra utan pannkakor och kanelbullar till slut. Far, för att du alltid valt att vara där för mig och delat dina djupa, och ibland meningslösa, kunskaper med mig. Och för att jag nu vet att man inte ska slänga doktorsringen i pappersåtervinningen.

Madde för att du står ut med mig och alla mina egenheter, för att du är min och allas gladjespridare, för att tiden och framtiden med dig känns både rätt och härlig, och för att vi nu ska få bränna alla våra pensionspengar på Nopa.

Mormor för att du är den härligaste och mest öppenhjärtiga människan jag har i mitt liv.
References

33 Yoshihara, H. et al. (2007) Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. *Cell Stem Cell* 1, 685–697
34 Qian, H. et al. (2007) Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. *Cell Stem Cell* 1, 671–684
36 Kiel, M.J. et al. (2005) SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* 121, 1109–1121
40 Katayama, Y. et al. (2006) Signals from the sympathetic nervous system regulate hematopoietic stem cell egress from bone marrow. *Cell* 124, 407–421


Arinobu, Y. et al. (2007) Reciprocal activation of GATA-1 and PU.1 marks initial specification of hematopoietic stem cells into myeloiderythroid and myelolymphoid lineages. *Cell Stem Cell* 1, 416–427


88 Kikushige, Y. et al. (2011) Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukaemia. *Cancer Cell* 20, 246–259
112 Mohi, M.G. et al. (2005) Prognostic, therapeutic, and mechanistic implications of a mouse model of...
leukemia evoked by Shp2 (PTPN11) mutations. *Cancer Cell* 7, 179–191


Meggendorfer, M. *et al.* (2013) SETBP1 mutations occur in 9% of MDS/MPN and in 4% of MPN cases and are strongly associated with atypical CML, monosomy 7, isochromosome i(17)(q10), ASXL1 and CBL mutations. *Leukemia* 27, 1852–1860


Torrano, V. et al. (2011) ETV6-RUNX1 promotes survival of early B lineage progenitor cells via a dysregulated erythropoietin receptor. Blood 118, 4910–4918


Cavé, H. et al. (1997) ETV6 is the target of chromosome 12p deletions in t(12;21) childhood acute lymphocytic leukemia. Leukemia 11, 1459–1464


Moorman, A.V. et al. (2007) Karyotype is an independent prognostic factor in adult acute lymphoblastic leukaemia (ALL): analysis of cytogenetic data from patients treated on the Medical Research Council (MRC) UKALLXI/II Eastern Cooperative Oncology Group (ECOG) 2993 trial. Blood 109, 3189–3197


Hebert, J. et al. (1994) Candidate tumor-suppressor genes MTS1 (p16INK4A) and MTS2 (p15INK4B) display frequent homozygous deletions in primary cells from T- but not from B-cell lineage acute lymphoblastic leukemias. Blood 84, 4038–4044


Lapidot, T. et al. (1994) A cell initiating human acute myeloid leukaemia after transplantation into SCID
mice. *Nature* 367, 645–648


Pearce, D.J. et al. (2006) AML engraftment in the NOD/SCID assay reflects the outcome of AML: implications for our understanding of the heterogeneity of AML. *Blood* 107, 1166–1173

Vargaftig, J. et al. (2012) Frequency of leukemic initiating cells does not depend on the xenotransplantation model used. *Leukemia* 26, 858–860

Rehe, K. et al. (2013) Acute B lymphoblastic leukaemia-propagating cells are present at high frequency in diverse lymphoblast populations. *EMBO Mol Med* 5, 38–51


Lawrence, H.J. et al. (1997) Mice bearing a targeted interruption of the homeobox gene HOXA9 have defects in myeloid, erythroid, and lymphoid hematopoiesis. *Blood* 89, 1922–1930


Sobhian, B. *et al.* (2010) HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol. Cell* 38, 439–451


Nguyen, A.T. *et al.* (2011) DOT1L, the H3K79 methyltransferase, is required for MLL-AF9-mediated...
leukemogenesis. Blood 117, 6912–6922

Chang, M.-J. et al. (2010) Histone H3 lysine 79 methyltransferase Dot1 is required for immortalization by MLL oncogenes. Cancer Res. 70, 10234–10242

Bernt, K.M. et al. (2011) MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. Cancer Cell 20, 66–78


Okuda, H. et al. (2015) AF4 uses the SL1 components of RNP1 machinery to initiate MLL fusion- and AEP-dependent transcription. Nat Commun 6, 8869

Okuda, H. et al. (2016) TBP loading by AF4 through SL1 is the major rate-limiting step in MLL fusion-dependent transcription. Cell Cycle 15, 2712–2722


Dawson, M.A. et al. (2011) Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature 478, 529–533


Dobson, C.L. et al. (2000) Tumorigenesis in mice with a fusion of the leukaemia oncogene Mll and the bacterial lacZ gene. EMBO J. 19, 843–851


Stavropoulou, V. et al. (2016) MLL–AF9 Expression in Hematopoietic Stem Cells Drives a Highly Invasive AML Expressing EMT-Related Genes Linked to Poor Outcome. Cancer Cell 30, 43–58


Ohlsson, E. et al. (2014) Initiation of MLL-rearranged AML is dependent on C/EBPα. J. Exp. Med. 211,


Li, Z. et al. (2013) PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood* 121, 1422–1431


Kroon, E. et al. (1998) Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1 but not Pbx1b. *EMBO J.* 17, 3714–3725


Wilkinson, A.C. et al. (2013) RUNX1 is a key target in t(4;11) leukemias that contributes to gene activation through an AF4-MLL complex interaction. *Cell Rep* 3, 116–127


Xu, J. et al. (2016) MLL1 and MLL1 fusion proteins have distinct functions in regulating leukemic transcription program. *Cell Discov* 2, 16008

Thiel, A.T. et al. (2010) MLL-AF9-induced leukemogenesis requires coexpression of the wild-type Mll allele. *Cancer Cell* 17, 148–159


van der Linden, M.H. et al. (2014) MLL fusion-driven activation of CDK6 potentiates proliferation in MLL-rearranged infant ALL. *Cell Cycle* 13, 834–844

Wang, Q.-F. et al. (2011) MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. *Blood* 117, 6895–6905


Matsuo, H. et al. (2014) EVI1 overexpression is a poor prognostic factor in pediatric patients with mixed lineage leukemia-AF9 rearranged acute myeloid leukemia. *Haematologica* 99, e225–7


Montes, R. et al. (1997) Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. *EMBO J.* 16, 4226–4237


Lin, S. et al. (2016) Instructive Role of MLL-Fusion Proteins Revealed by a Model of t(4;11) Pro-B Acute Lymphoblastic Leukemia. *Cancer Cell* 30, 737–749


Montes, R. et al. (2011) Enforced expression of MLL-AF4 fusion in cord blood CD34+ cells enhances the hematopoietic repopulating cell function and clonogenic potential but is not sufficient to initiate leukemia. *Blood* 117, 4746–4758


Fu, J.-F. et al. (2017) Cooperation of MLL/AF10(OM-LZ) with PTPN11 activating mutation induced monocytic leukemia with a shorter latency in a mouse bone marrow transplantation model. *Int. J. Cancer* 140, 1159–1172


Montes, R. et al. (2014) Ligand-independent FLT3 activation does not cooperate with MLL-AF4 to
immortalize/transform cord blood CD34+ cells. Leukemia 28, 666–674
397 Koss, C. et al. (2014) Targeted Inhibition of the MLL Transcriptional Complex By Proteosome Inhibitors Elicits a High Response Rate in Relapsed/Refractory MLL Rearranged Leukemia. ASH Abstract
399 Kerstjens, M. et al. (2016) MEK inhibition is a promising therapeutic strategy for MLL-rearranged infant acute lymphoblastic leukemia patients carrying RAS mutations. Oncotarget DOI: 10.18632/oncotarget.11730