SAPIENZA
UNIVERSITÀ DI ROMA

Department of Molecular Medicine
PhD program in Human Biology and Medical Genetics
XXXVI cycle

The functional role of EZH2 inactivation in the pathogenesis of T-cell acute lymphoblastic leukemia

PhD student
Mattia Colucci
matr.1945588

Academic supervisor
Prof.ssa Viviana Caputo

Co-supervisor
Dott. Vincenzo Giambra

Academic year2022/2023

Table of Contents

Figure Index

Table index

Appendix

Abstract

T-cell acute lymphoblastic leukemia (T-ALL) is a hematological malignancy, characterized by frequent loss-of-function mutations in EZH2 gene. EZH2 is a histone-lysine N-methyltransferase enzyme and in mouse, its inactivation leads to the initiation and progression of T-cell leukemia. Here, we characterized the *in vitro* phenotypes associated with a long-term expansion of human CD34+ hematopoietic stem and progenitor cells (HSPCs) in response to CRISPR/Cas9-mediated inactivation of EZH2 alone or in combination with the constitutive expression of known T-ALL oncogenes after lentiviral transduction. It is known that early T cell leukemias with inactivating EZH2 mutations are closely associated with the upregulation of HOXA9 and LYL1 oncogenes. In our *in vitro* model EZH2 knockout in LYL1 or HOXA9-transduced HSPCs, expanded *in vitro* up to 31 days, exhibited an increased leukemia clonogenic capacity and promoted an enrichment of early progenitor T cells at the double negative 1 (DN1) development stage. Furthermore, EZH2-null LYL1 and EZH2-null HOXA9-transduced cells demonstrated high penetrance leukemia in *vivo* experiments. RNA-seq data from ETP-ALL-like LOUCY cells revealed that EZH2 knockout overexpressing HOXA9 or LYL1 resulted in a gene expression signature of immature T-ALL, stem/immature immunophenotypic stage, increased replication status, and enhanced dependency on the inflammation response. Lastly, RNA-seq data analysis from a large pediatric cohort of T-ALLs revealed that patients with Loss of Function (LoF) mutations in EZH2 gene shared a stem cell-associated transcriptional program and were statistically enriched for the immature HOXA cluster and ETP subtype in human T-ALL leukemia.

1 Introduction

1.1 Hematopoiesis and hematopoietic hierarchy

The comprehension of hematopoiesis, the foundational process that generates all blood cells, has undergone a significant shift in scientific thinking from a hierarchical organization to a more actual prospective. In 1909, the concept of a hierarchical system with a hematopoietic stem cell (HSC) as the common precursor for all blood cell types was introduced (1). This idea was validated in the 1960s by Becker, Till, and McCulloch, who provided empirical evidence for the existence of a common multipotent precursor for multiple blood cells (2). In humans, Hematopoietic Stem Cells (HSCs) are self-renewing, multipotent stem cells that are present in the bone marrow. The total number of HSCs in adult humans is estimated to be in the range of 50,000–200,000 (3). They produce precursors that can develop into all cell types found in peripheral blood.

The HSC initiates the differentiation pathway that splits into the long-term and short-term repopulating HSCs (LT/ST-HSCs) which progressively exhibit diminished self-renewing potential and more differentiated phenotypes (3). The multipotent progenitors (MPPs) gradually lose their self-renewal capacity and differentiate into lineage-committed progenitors. The common myeloid progenitor (CMP) differentiates in megakaryocyte-erythrocyte progenitor (MEP) leading to the production of red blood cells and platelets. The granulocyte-macrophage progenitor (GMP) gives rise to monocytes, various granulocytes (such as neutrophils, basophils, and eosinophils), and dendritic cells. Finally, the common lymphoid progenitor (CLP) is responsible for generating natural killer (NK) cells, as well as T and B lymphocytes (4,5). However, new advanced single-cell transcriptomic analysis has displayed that the epigenetic landscape plays a crucial rule in regulating HSC heterogeneity and hematopoietic cell fate decisions,making blood cell differentiation as a “heterogeneous continuum” rather than a rigid, stepwise process (4) **Figure 1**.

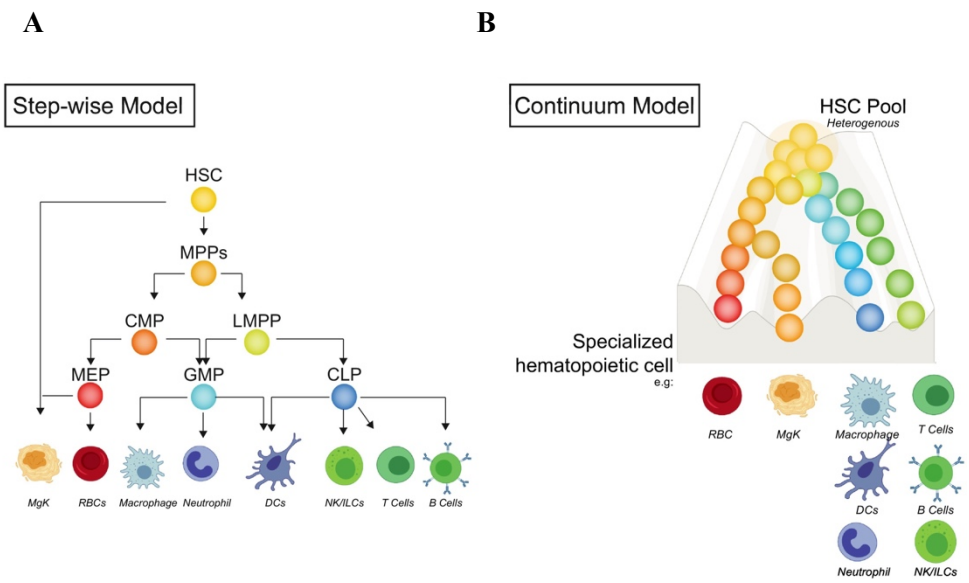


Figure 1 Scheme depicting the canonical and continuum model of hematopoiesis pattern
A) Under canonical model, HSCs are thought to contribute to each lineage equally.**B**) Revised model showing that hematopoiesis more closely represents a continuum (Rodrigues et al., 2021)

1.2 Human T-cells differentiation

In contrast with all other hematopoietic cells, T lymphocytes are not generated in the BM. T-cell development starts when hematopoietic progenitors positive for CD34 (CD – cluster of differentiation) migrate from the BM to the thymus. The subsequent expression of CD7 and CD10 characterizes CD34+ CD45RA+ precursors as oriented towards, but not fully committed to, the lymphoid lineages, thus identifying the thymic progenitors (TSPs). The presence of multipotent progenitors in the human thymus suggests that T-cell commitment occurs within this organ, marked by CD44 expression and activation of NOTCH signaling (6). Early T-lineage progenitors (ETPs) are identified within the CD34+/CD1α- population, which retain the ability to differentiate into B Cells, dendritic cells (DCs), Natural Killer cells (NK), Innate Lymphoid Cells (ILCs), and neutrophilic granulocytes (6, 7).

The release of Delta-like ligands (DLL, or DL) by thymic epithelial cells – supports the activation of T-cell differentiation programs in ETPs by activating NOTCH signaling. NOTCH receptor activates the initial lineage-specific transcription factors, with key regulators including transcription factors encoded by GATA-2, MEIS1, HOXA9 (Homeobox A9), and LYL1 (7,8). If enough NOTCH signaling is achieved the early thymic progenitors (ETPs/DN1) progress

through double-negative stages 2a (DN2a), DN2b, DN3a, DN3b, transitional DN4, followed by immature single positives (ISPs) and double positives (DPs), where DN refers to CD4- CD8- cells and DP to CD4+ CD8+ cells. An overview of this process in humans is shown in **Figure2.**

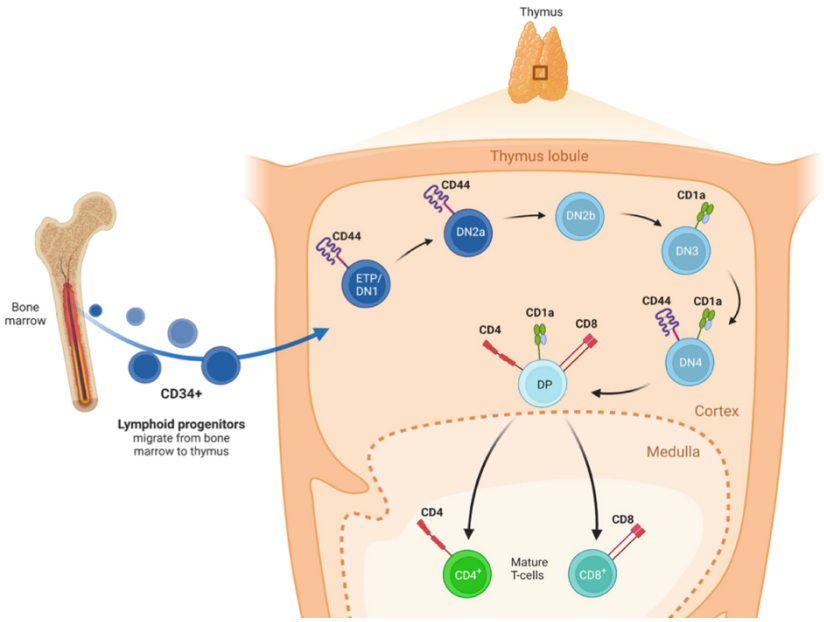


Figure 2 CD34+ progenitor cells move to and populate the thymus
Within the thymic cortex, Early T-cell progenitors (ETPs) encounter NOTCH signaling, releasing essential factors for T-cell differentiation. Phenotypically, distinct stages of T-cell differentiation can be observed, and the main surface markers found in humans are depicted. Commitment to the T-cell lineage is believed to occur at the DN2b stage (reviewed from Chopp et al; 2023)

The T cell-specific transcription factors, including *RAG1* and *RAG2,* mediate the rearrangement of TCR loci which involves the recombination of variable (V), diversity (D) and joining (J) gene segments via double-stranded DNA breaks mediated by RAG proteins. Additionally, the order of TCR loci rearrangement, guided by chromatin accessibility, typically proceeds as TCRD (delta), TCRG (gamma), TCRB (beta), and finally TCRA (alpha) (9). A successful recombination of the TCRD (delta) and TCRG (gamma) loci leads to the expression of a gamma-delta (γδ) T-cell receptor (TCR), which is characteristic of a small subset of T cells. These γδ T cells are unique in that they typically bypass the double-positive (DP) stage and exit the thymus as CD4/CD8- phenotype. In contrast,αβ T cells undergo a rigorous selection process that includes both positive and negative selection in the thymus, playing a pivotal role in adaptive immune responses against pathogens (8).

1.3 Epigenetic: an overview

The chromatin environment is fundamental in the establishment and maintenance of cell identity. Accordingly, protein complexes that modulate chromatin structure are important for many aspects of mammalian development and stem cell function. It is not surprising that their deregulation is involved in the pathogenesis of different human diseases, in particular cancer (10). In general, the cancer epigenome is characterized by aberrations and misregulation at multiple layers that involve all the epigenetic actors (11). Indeed, changes in DNA methylation and histone modification patterns as well as altered expression levels or activity of chromatin-modifying enzymes can be observed in cancer (12).
Epigenetic mechanisms and players responsible for the regulation of chromatin organization and packaging comprise chromatin remodelers, histone variants incorporation, and covalent modifications of DNA and histones (13).

The establishment and maintenance of specific chromatin states are crucial during development and during all the DNA template-based processes that require rapid rearrangements of chromatin structure. Such chromatin dynamicity is achieved through the activity of different remodeling complexes (14). Chromatin-remodeling enzymes facilitate a broad spectrum of transformations within chromatin, such as moving the histone octamer along the DNA, altering the conformation of nucleosome DNA, and modifying the histone octamer's composition. In general, the nucleosome is the main structure determining DNA conformation, which in turn is regulated by histone modification. Histones can be modified by the addition of various chemical groups such as acetyl, methyl, phosphoryl, and ubiquityl (15). These modifications can either compact or relax the chromatin structure, thereby either facilitating or limiting the binding of other proteins to chromatin. In addition to the type of modification, there is diversity, as lysine can be mono-, di-, or trimethylated. Arginines can be either mono- or dimethylated, and the dimethylation of arginines can be either symmetric or asymmetric (15,16).

Lysine residue acetylation and methylation are critical histone modifications that facilitate active versus repressed states of gene expression (17). Usually, lysine acetylation at gene start sites is associated with gene transcription and deacetylation with gene repression (18). Lysine methylation is mediated by histone methyltransferases (HTMs) and removed by histone demethylases (HDMs) (19). Such modifications can induce either gene expression or repression. For example, methylation of lysine 4 on histone 3 (H3K4) marks gene activation

(20). In contrast, methylation of lysine 9 or lysine 27 (H3K9me3, H3K27me3) marks repressed gene promoters (21).

1.3.1 Polycomb Repressive Complex 2 (PRC2)

Polycomp group (PcG) proteins are members of highly conserved multiprotein complexes that can silence gene expression, particularly during development and differentiation. During the development of an organism, several epigenetic regulators, such as PcG complexes contribute to silence genes temporarily or permanently, through histone modification (21,22).
Polycomb group proteins (PcGs) assemble into several multiprotein complex families, especially in the Polycomb Repressive Complexes PRC1 and PRC2 (23,24). Both groups of proteins promote gene repression via modification of histone tails and chromatin compaction. Indeed, at the molecular level, PRC1 catalyzes the monoubiquitylation of lysine 119 on histone H2A (H2AK119ub) (25) while PRC2 is responsible for the mono-, di- and tri-methylation of lysine 27 of histone H3 (H3K27me1/me2/me3) (26,27).

PRC2 is the only identified methyltransferase catalyzing mono-, di-, and trimethylation of H3K27 (28) . Each of these methylation levels has a different function and shows a distinct genomic distribution. H3K27me1 is enriched within gene bodies of actively transcribed genes and it marks 5-10% of all H3 histones in a cell. H3K27me2 is the most abundant, it is present in inter- and intragenic regions on 50-70% of total histone H3. H3K27me3 marks 5-10% of histone H3 and it is found especially in PRC2 binding site H3K27me2 constitutes the substrate for H3K27me3 formation and prevents H3K27 acetylation, an antagonist of methylation (29,30). Trimethylation of H3K27 correlates with gene silencing. While H3K27me1 is associated with constitutive heterochromatin, di- and trimethylated H3K27 are associated with facultative heterochromatin (31).PRC2 activity is required to maintain gene repression and support cell identity, through the maintenance of the gene expression patterns (32). For this reason, PRC2 function is often deregulated in disease and is a promising candidate for therapeutic targeting in cancer.

PRC2 core complex is composed of Enhancer of Zeste Homolog 1 or 2 (Ezh1 or Ezh2), Suppressor of Zeste 12 Protein Homolog (Suz12), and Embryonic ectoderm development (Eed). These three core subunits are present in the complex in a 1:1:1 stoichiometry and contain the minimal composition necessary for the catalytic activity of PRC2 (33,34,35).

The central complex is then associated with several additional proteins (JARID2, AEBP2, PCL1-3, C17orf96, and C10o12) and ncRNAs. More precisely, PRC2 is composed of two structurally and functionally distinct modules: while the catalytic module contains the minimally active ternary complex of EZH2, EED, and the VEFS C-terminal domain of SUZ12 (SUZ12(VEFS)), the accessory unit binding module is composed of the N-terminal portion of SUZ12 (SUZ12(N)) and RBBP4, which together form attachment sites for various accessory subunits. EZH2 is folded into a series of functional domains such that EED and SUZ12(VEFS) are bound to distinct subsets of EZH2 domains, adjacent to each other (36) **Figure 3**.

The allosteric activation of the PRC2 plays a pivotal role in its catalytic functionality, intricately tied to the chromatin's structural state. In compacted and transcriptionally repressed chromatin, EED engages with trimethylated lysine 27 on histone H3 (H3K27me3) via its aromatic cage. This interaction triggers a conformational reorganization of the PRC2 subunits, effectively orchestrating the alignment of the stimulation-responsive motif (SRM) within the EZH2 subunit and securing the structural integrity of the EZH2-SET domain. In particular, the SET domain undergoes a structural reorganization from the inactive basal state to the active state, reshaping the binding groove with H3 and the binding pocket with SAM (S-adenosyl-L-methionine) for effective substrate binding. This way, the methylation of the nearby nucleosome is favored, leading to the spread of the histone mark along the linear chromatin fiber and across regions that are in spatial proximity. These molecular rearrangements facilitate the subsequent methylation of histone targets (36).

Conversely, in an open chromatin configuration conducive to active transcription, the SUZ12 component of PRC2 discerns histone marks indicative of active gene expression, such as trimethylated lysine 4 (H3K4me3) and dimethylated or trimethylated lysine 36 (H3K36me2/3) on histone H3. The recognition of these marks by SUZ12 prompts a transfer of information to EZH2, leading to the inhibition of EZH2’s enzymatic activity. This regulatory mechanism ensures that PRC2's repressive action is precisely targeted and modulated according to the chromatin context, preventing the methylation of histones in regions of active transcription (37).

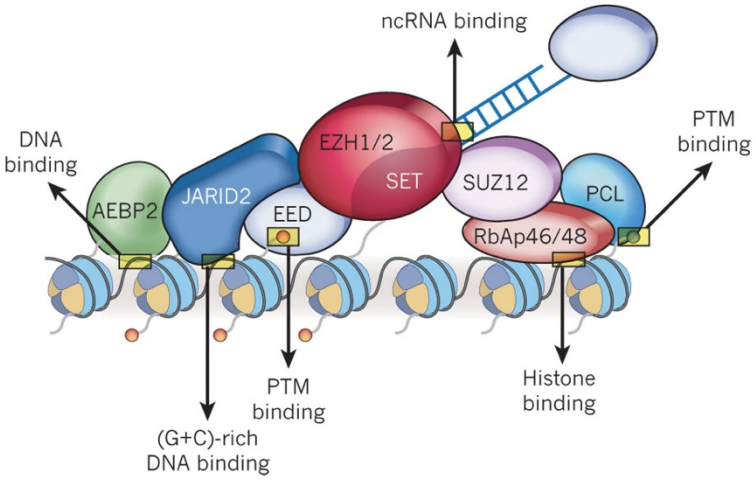


Figure 3 Schematic representation of the PRC2 complex
Putative interactions with either DNA or histones that could explain PRC2 recruitment are highlighted.(Margueron et al,2011).

ChIP-seq studies have revealed that PRC2 target genes are enriched for sequences that bind transcription factors and components of signaling pathways promoting cell proliferation while repressing genes involved in cellular differentiation. This correlation suggests that PRC2 acts as a key corepressor in embryonic stem cells, balancing proliferation and differentiation (38). The high overlap between genes that bind proliferation-promoting transcription factors and those targeted by PRC2 implies that the complex plays a central role in maintaining the undifferentiated state and proliferation of embryonic stem cells (39).

However, the role of PRC2 in cancer is complex and seemingly contradictory. Alterations leading to both gain and loss of function of PRC2 have been linked to oncogenesis (40).In particular, loss-of-function mutations in key components of PRC2, such as EZH2, EED, and SUZ12, are commonly observed in various cancers, including T-cell acute lymphoblastic leukemia (T-ALL) (40,41). These inactivating mutations or chromosomal translocations can compromise PRC2's ability to repress genes, leading to the inappropriate expression of genes that promote cellular proliferation and inhibit differentiation, thus contributing to tumorigenesis (32). These studies underscore the importance of PRC2 in transcription regulation and its complex relationship with tumorigenesis. To better understand these roles, further exploration of the mechanisms by which PRC2 contributes to gene regulation and cancer development is necessary. These efforts could eventually lead to the development of targeted therapies to modulate PRC2 activity for cancer treatment.

1.3.2 EZH2:structure and function

The structure of EZH2 is characterized by several functional domains, including a SET domain responsible for its methyltransferase activity. The SET domain of EZH2 is flanked by various other domains and motifs that facilitate interaction with DNA, other components of PRC2, and histone tails: a WD-40 binding (WDB) domain, the I-II domains, two SWI3-ADA2-N-CoR-TFIIIB (SANT) domains, a cysteine-rich CXC domain, and the evolutionarily conserved Su(var) 3-9, enhancer of Zeste, trithorax (SET) methyltransferase domain (42,43).
The structural and functional complexities of EZH2, particularly its SET domain located in the C-terminal region, underline its pivotal role in the catalytic activity of the Polycomb Repressive Complex 2 (PRC2). The SET domain is distinguished by two significant pockets: a highly hydrophobic channel composed of amino acids Y641, F667, F724, Y726, and Y728, rich in tyrosine and phenylalanine that accommodates the long aliphatic chain of histone lysine; and a second pocket that specifically binds the cofactor S-adenosylmethionine (SAM), essential for its methyltransferase function (44). This arrangement underscores the precision with which EZH2 mediates the transfer of methyl groups to histone tails, a process central to epigenetic regulation. Adjacent to the SET domain are the pre-SET and post-SET regions, each contributing uniquely to the enzyme's functionality. The pre-SET domain, rich in cysteines (CXC motif), is indispensable for methyltransferase activity, suggesting a role in maintaining the structural integrity necessary for catalysis (45).

Conversely, the post-SET domain, situated C-terminally to the SET domain, acts as a regulatory gate, preventing substrate access to the active site and thereby controlling the enzyme's activity (46). This regulatory mechanism highlights the control EZH2 exercises over its methyltransferase function. Central to EZH2's interaction with chromatin are the SANT1 and SANT2 domains, which facilitate the binding of this chromatin-remodeling protein to histones, thus stabilizing the active site and promoting the enzymatic activity of PRC2 on chromatin substrates. This interaction is crucial for the repression of gene expression through histone modification, illustrating the direct link between EZH2 structure and its epigenetic regulatory functions (47).Further other regulatory complexity is the WD (tryptophan-aspartic acid) repeat domain at the N-terminal region, which contains the SRM. This motif is responsible for the allosteric activation of PRC2 through interaction with the SET domain, highlighting an intricate network of intramolecular interactions that govern the complex's activity (48). The presence of such a motif suggests a sophisticated mechanism by which

EZH2's activity is modulated in response to cellular signals, further emphasizing the enzyme's role as a dynamic regulator of gene expression (49). Together, these structural components of EZH2 not only facilitate its primary function as a methyltransferase within the PRC2 complex but also integrate various regulatory mechanisms that ensure precise control over gene expression.

1.3.3 Role of EZH2 in hematopoiesis and signaling pathways

EZH2 contributes significantly to the maintenance of pluripotency and self-renewal in adult stem cells, highlighting its critical role in normal hematopoiesis (50). This function of EZH2 ensures the continuous production and proper differentiation of blood cells, which is essential for the overall health of the organism.EZH2's expression patterns are crucial in cell proliferation and the immune system. It was observed that EZH2 is highly expressed in proliferating cells during lymphopoiesis, including human germinal center B cells and cycling T and B lymphocytes. This expression pattern suggests a significant role for EZH2 in cell cycle regulation and lymphocyte division, critical processes for immune response and maintenance (50).

The Human Protein Atlas database corroborates the critical role of EZH2 by showing its high expression levels in bone marrow and lymphoid tissue. This expression profile is particularly relevant given the bone marrow's role as the primary site of hematopoiesis and the lymphoid tissue's involvement in generating immune responses. Such high levels of EZH2 expression in these tissues further highlight its indispensable role in the differentiation and proliferation of hematopoietic stem and progenitor cells (HSPCs), ensuring the efficient functioning of the body's hematopoietic and immune systems. Research has shown that alterations in the components of the PRC2 complex can disrupt T-cell development and contribute to immunodeficiency and leukemogenesis. For instance, studies have demonstrated that EZH2 subunit of PRC2 is essential for the early stages of thymocyte development, influencing the transition from double-negative to double-positive T cells (51). Furthermore, the dynamic regulation of PRC2 activity is necessary for the activation and repression of lineage-specific genes at different stages of T cell differentiation, highlighting its importance in T cell fate decisions and immune function (52).

1.4 Homeobox HOX genes

The homeobox (HOX) genes encode a family of transcription factors that are key components of substantial metabolic processes such as lipidic metabolism, and their roles in organogenesis and tumorigenesis have been studied in detail over the past decade. These genes are highly conserved and contain homologous domains in almost all eukaryotic cells. In the human genome, according to the sequence similarity and position correlation in the chromosome, HOX family genes can be divided into 4 clusters, namely HOXA, HOXB, HOXC, and HOXD. Each cluster has between 9 and 11 genes in a row on a homologous strand of DNA, which contains duplication and divergence of ancestral HOX genes. The expression patterns of HOX genes from clusters A and B in hematopoietic stem cells (HSCs) and progenitors mirror their activation during early developmental stages, displaying lineage-specific and differentiation stage-constrained behaviors. For instance, HOXB3, HOXB4, and HOXA9 exhibit pronounced expression levels in undifferentiated hematopoietic cells (53).

As HSPCs differentiate and become fully mature, the *HOXA* 5-10 genes are downregulated and epigenetically silenced (54). Patterns of deregulated HOX expression have been identified in cancer. In some tissues, certain HOX genes that normally have tumour suppressive effects are silenced; in other tissues, particular HOX genes are expressed in an aberrant temporospatial pattern with oncogenic effects (55). This coordinated regulation of *HOXA* gene expression is mediated by various epigenetic factors modulating histone methylation, acetylation, and DNA methylation. In general, two master regulators of HOXA9 expression, the mixed-lineage leukemia proteins and the polycomb group histone methyltransferases, activate and repress *HOXA9* transcription, respectively. Thereby it is considered as one of the driving forces in leukemogenesis (56). *HOXA9* deregulation often coincides with genetic alterations, such as *EZH2* loss-of-function mutation. Interestingly, overexpression of HOXA9 is found in myeloid malignancies with decreased EZH2 expression. Knockdown of EZH2 in ALL cells results in elevated HOXA9 level (57). The studies in mouse models also confirmed that HOXA9 was depressed by EZH2 loss at the myelodysplastic syndrome stage (58).

1.5 LYL1

Transcription factors are crucial in the normal developmental process of hematopoietic cells. Yet, the expression of these transcription factors and their roles in different types of human leukemia remain poorly understood. The LYL1 gene (lymphoblastic leukemia-derived sequence 1) encodes a transcription factor with a basic helix-loop-helix (bHLH) motif and is critically involved in the homeostasis of immature hematopoietic cells (59). Initially, it was discovered in certain instances of T-cell acute lymphoblastic leukemia (T-ALL) at the breakpoint region of the chromosomal translocation t(7;19)(q35;p13) (60). The translocation is in head-to-head juxtaposition with the T-cell antigen receptor beta (TCR-β) gene, resulting in truncation of the LYL1 gene and production of abnormal-sized RNAs.

Bone marrow's multipotent progenitor cells, including MPPs and LMPPs, along with ETPs, demonstrate significant levels of LYL1 expression. Conversely, HSCs and DN2 thymocytes show reduced levels of LYL1 expression (61). Elevated levels of LYL1 are linked to a stem cell-like phenotype in leukemia (characterized by CD34+ markers), indicative of an especially adverse outcome. Previous studies revealed the cancer-causing capacity of LYL1 in mouse model. These mice were genetically modified to produce excessive amounts of LYL1, a result of the gene being under the control of the elongation factor 1α (EF1A) promoter. This modification led to the pronounced LYL1 expression. A significant portion, specifically 30%, of these genetically modified mice, developed leukemia affecting T cells. This condition was further characterized by the widespread invasion of leukemia cells into various organs, underlining the aggressive nature of the disease triggered by LYL1 overexpression (62).

1.6 T-Cell Acute Lymphoblastic Leukemia

T-cell Acute Lymphoblastic Leukemia (T-ALL) is a particular type of ALL, which is characterized by an expansion of early T cell progenitors. It represents a significant proportion of acute lymphoblastic leukemia (ALL) diagnoses, accounting for about 15% of pediatric and 25% of adult ALL cases (63). Although less common, the disease affects adolescents and adults, a group of patients that usually presents a poorer prognosis, with a current 5-year survival rate of 40-50% (64). T-ALL is primarily caused by genetic and epigenetic alterations that lead to the uncontrolled proliferation and impaired differentiation of T-cell precursors. Chromosomal translocations that result in the overexpression of oncogenes or the formation of fusion proteins also contribute to the disease. Moreover, epigenetic modifications, including

DNA methylation and histone modification, play a role in silencing tumor suppressor genes or activating oncogenes, further driving the leukemic transformation of T-cells (14) . Clinically, patients present diffuse infiltration of blasts expressing immature T-cell markers, accompanied by neutropenia, anemia, and/or thrombocytopenia (63). In the last few years, the development of intensive chemotherapy treatments has improved the prognosis of T-ALL patients. Despite the progress made, about ~20% of children and ~50% of adult patients relapse after chemotherapy and the outcome remains very poor. Further, remission is associated with toxicities and a significant toll on long-term health and quality of life (65). Recently, a high-risk subtype of ALL derived from the malignant transformation of early T-cell precursor (ETP) differentiation stage has been described in both human and mouse models of T-ALL (67,68) Notably, early T-cell leukemias share similarities with hematopoietic stem and early progenitor cells and carry oncogenic alterations driving disease progression (14,69). Nonetheless, the genetic or signaling programs, that modulate the early T-ALLs, are still under intense investigation.

1.6.1 The cellular origins of T-ALL

Acute Lymphoblastic Leukemia (ALL) could potentially originate from an isolated normal progenitor cell that encounters an initial oncogenic trigger, which results in the origin of a reprogrammed, premalignant Cancer Stem Cell (CSC) (70). This cell, now susceptible to further genetic alterations, marks the inception of a malignant transformation process. The developmental stage of lymphoid differentiation offers a fertile ground for Leukemia-Initiating Cells (LIC) to emerge. These cells are distinguished by their inherent stem-like qualities, remarkable plasticity, and a unique ability to cultivate an aberrant proinflammatory niche within the bone marrow environment. This niche, in turn, accelerates the progression of ALL by disrupting the balance of normal hematopoiesis (63,70). Within the specific cellular background of the bone marrow, characterized by the CD34+/CD7+/CD1α- phenotype, a distinct subgroup of cells demonstrates the capacity for leukemia initiation and proliferation. This subgroup exhibits similarities to the CD1α- Early T-cell Precursor (ETP) population, which is linked to the particularly aggressive form of leukemia known as ETP-ALL (71). Furthermore, the role of stromal cells in this complex ecosystem cannot be overstated; they are pivotal in the secretion of IL-7, a cytokine essential for the survival and proliferation of ETP-ALL cells. These cells necessitate of IL7R signaling for their growth and survival, highlighting

the critical interplay between ETP-ALL cells and their supporting stromal niches that express IL-7(14).

1.6.2 T-ALL: genetic alterations and transcription factor deregulation

Neoplastic transformation is a multistep process in which various alterations cooperate to deregulate the normal mechanisms controlling cell proliferation, survival, growth, and differentiation. In the case of T-cell acute lymphoblastic leukemia (T-ALL), frequently mutated gene classes include regulators of lymphocyte differentiation, cell cycle genes, tumor suppressors, or molecules mediating the typical cell signaling of lymphoid cells. The close interplay between the key regulators of early T-cell development and the oncogenic signals in T-ALL is exemplified by the prominent role of NOTCH1, a critical factor for T-cell fate specification and thymocyte development.

Activating mutations in NOTCH1 often coincide with the loss of the cyclin-dependent kinase inhibitor 2A (CDKN2A) locus, which encodes the tumor suppressors p16INK4A and p14ARF, and with chromosomal rearrangements, leading to the aberrant expression of a heterogeneous group of T-cell-specific transcription factors that can function as oncogenes. These include genes encoding basic helix–loop–helix (bHLH) factors such as T-Cell Acute Lymphoblastic Leukemia 1 (TAL1), TAL2, the leukemia-associated hematopoietic regulator 1 (LYL1), BHLHB1 (also known as OLIG2), the LIM domain-only (LMO) genes like LMO1 and LMO2; the T cell leukemia homeobox 1 (TLX1), TLX3, NK2 homeobox 1 (NKX2-1), NKX2-2, and the homeobox A (HOXA) genes; MYC, and MYB (67,71). Genetic alterations activating NOTCH1 signaling are characteristic of T-ALL. Other frequently altered TFs include *SPI* and *MEF2C*. Epigenetic regulators are recurrently mutated across all T-ALL subgroups, including EED, EZH2, and SUZ12, which encode core components of the Polycomb Repressive Complex 2 (PRC2)(14,67,71) **Figure 4.**

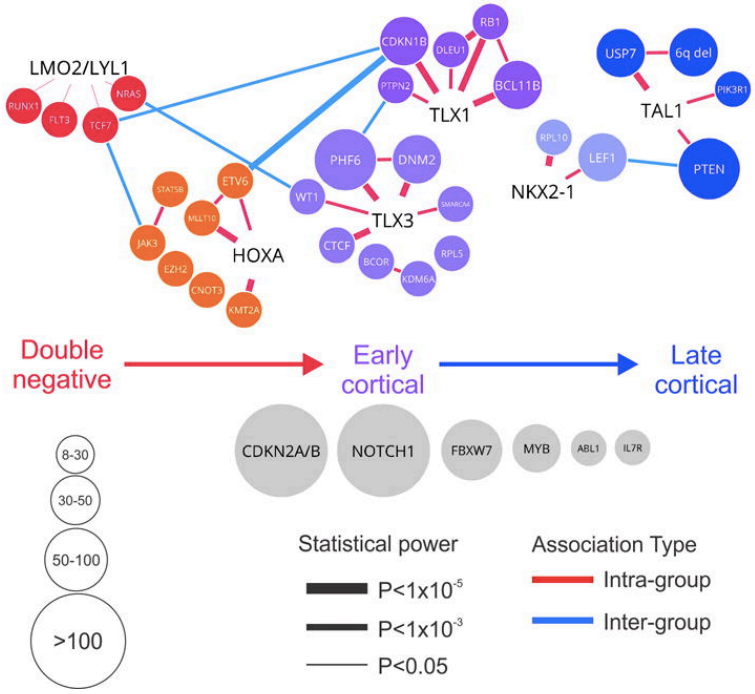


Figure 4 Associations between genetic alterations of T-ALL subgroups and the stages of T-cell development.

Nodes represent genetic alterations and are color-coded according to the T-ALL subtype in which they are significantly enriched (p-value by two-sided Fisher's exact test). The size of each node represents the frequency of the genetic alteration across the entire cohort (see the key in the lower left corner). The thickness of lines connecting the nodes indicates the statistical significance of co-occurrence of a gene–gene or gene–subtype pair (p-value by Spearman correlation analysis). Intra- and intersubtype co-occurrence are indicated by red and blue lines, respectively. Nodes in gray represent genetic alterations that were not specifically enriched in a T-ALL subtype. (Liu, 2017).

This intricate network of genetic and epigenetic alterations highlights the complexity of T-ALL pathogenesis. The deregulation of these key transcription factors and signaling pathways disrupts normal T-cell development, contributing to the leukemic phenotype. The classification of T-ALL into subgroups based on specific genetic profiles not only enhances our understanding of the disease's molecular basis but also opens the door to targeted therapeutic strategies. For instance, the identification of NOTCH1 signaling alterations has led to the exploration of NOTCH inhibitors as a potential treatment strategy for T-ALL. Similarly, the discovery of recurrent mutations in epigenetic regulators such as components of the PRC2 complex suggests that targeting the epigenetic landscape of T-ALL cells may offer new avenues for therapy.

1.6.3 Early T-cell precursor acute lymphoblastic leukemia (ETP-ALL)

A subtype of T-cell Acute Lymphoblastic Leukemia, known as "Early T-cell Precursor" (ETP) ALL, has been identified and accounts for up to 15% of T-ALL cases. It is associated with a high risk of treatment failure (71). Early T-cell precursors represent a subgroup of thymocytes that migrate from the bone marrow to the thymus and retain multilineage differentiation potential, suggesting their direct derivation from hematopoietic stem cells. The definition of ETP-ALL is based on the immunophenotype of the leukemic cells, which are typically CD1α-, CD8-, CD5- (DIM) and express one or more stem cell-related markers (e.g., CD34, CD13, or CD33) (67,71). In the 2017 World Health Organization (WHO) Classification, ETP-ALL falls into the category of early T-ALL. Clinically, patients with ETP-ALL exhibit a poor response to chemotherapy and have a very high risk of relapse. In particular, Campana et al identified a unique biological subtype of childhood leukemia, ETP-ALL, associated with a high risk of remission induction failure or relapse in patients treated with contemporary protocols of intensive chemotherapy for ALL. Despite a complete response (CR) rate of 73% after combination chemotherapy treatment, the median overall survival for patients with ETP-ALL is about 20 months (66) .

Genetically, ETP T-ALLs usually have a lower frequency of NOTCH pathway mutations, rarely have *CDKN2A* deletions, and are characteristically associated with mutations in genes encoding signaling factors and/or epigenetic modifiers. Their deregulated TFs are often either *LYL*/*LMO2* or *HOXA*. Moreover, The molecular characterization of ETP-ALL compared to other T-ALL subgroups has revealed distinct alterations, including some gene mutations that are highly enriched in ETP-ALL (67,71). ETP ALL was characterized by activating mutations in genes regulating cytokine receptor and RAS signaling (67% of cases; NRAS, KRAS, FLT3, IL7R, JAK3, JAK1, SH2B3 and BRAF), inactivating lesions disrupting haematopoietic development (58%; GATA3, ETV6, RUNX1, IKZF1 and EP300) and histone-modifying genes (48%; EZH2, EED, SUZ12, SETD2 and EP300)(14,67,71).

Mutations in epigenetic regulators can change the accessibility of certain chromatin areas to transcription factors, and if this process occurs at incorrect stages of T-cell maturation, it results in aberrant gene expression contributing to the disease's pathogenesis (72). They found that T-ALLs are among the tumors with the highest frequency of mutations in genes encoding proteins involved in epigenetic regulation. Therefore, the disruption of epigenetic homeostasis in normal T cells could act as a central component of hematopoietic T-cell transformation.

Subsequent genomic studies have identified recurrent alterations in genes involved in DNA methylation (DNMT3A, DNMT3B, TET1, IDH1, IDH2), histone methylation (EZH2, SUZ12, EED, JARID2, UTX, JMJD3, NSD2), and histone acetylation (CREBBP, EP300, HDAC7, HDAC5, NCOA3) in T-ALL (14,72).

In particular, patients with ETP-ALL who have mutations in core PRC2 components have unfavorable outcomes, especially in cases mutated in EZH2, with an estimate that 60% of these patients relapse within 5 years (73). Loss-of-function mutations in genes encoding the core components of the Polycomb Repressor Complex 2 (PRC2) are present in 25% of T-ALL cases and 42% of ETP-ALL cases (67). In particular, they reveal the presence of inactivating lesions in 11/64 cases in SUZ12, 8/64 cases in EED, 10/64 cases in EZH2, and any PRC2 alteration in 26/64 cases (73).

Moreover, inactivation of PRC2 components in ETP-ALL results in loss of global cellular H3K27me3, a chromatin mark associated with silent genes. A subset of these genes shows increased transcription. Among all PRC2 factors, EZH2 has the highest rate of mutations, and most mutations in this protein occur in the SET and CXC domains, essential for HMT activity. Liu et al., through genomic analysis in a sample of 264 T-ALL cases, identified 12 missense mutations – P132T, S280C, C289R; P563L, C565Y, C571S, in the CXC domain; Q653K, D657G, R684C, R684H, Y733H in the SET domain; E745K – and 2 frameshifts– V19_S12fs, L739fs – in EZH2— (83). Notably, mutations in the CXC domain of EZH2 lead to mutant isoforms with loss of function that are unable to stabilize the interaction of PRC2 with the nucleosome, resulting in the loss of HMT activity and partially destabilizing the integrity of the complex (83). Furthermore, mutations in other PRC2 factors such as AEBP2, SUZ12, EED, and PHF19 affect their interaction with nucleosomes and the stimulation of PRC2 (71).

1.6.4 Current Challenges in T-ALL Therapy

The treatment approach for T-ALL generally consists of multi-agent chemotherapy regimens, which are designed to target the rapidly proliferating leukemic cells. The chemotherapy regimen for T-ALL usually starts with an intensive induction phase aimed at achieving remission, followed by consolidation and maintenance phases to eliminate residual disease and prevent relapse. Patients showing poor response or high levels of Minimal residual disease (MRD) may undergo risk-intensified treatment regimens, potentially including more

aggressive chemotherapy, targeted therapies, or consideration for hematopoietic stem cell transplantation (74). This personalized approach aims to maximize the chance of cure while minimizing the risk of relapse, underscoring the importance of MRD as a prognostic indicator and a guide for treatment planning in T-ALL.

1.7 CRISPR/Cas9

CRISPR/Cas9 technology represents a revolutionary advancement in the field of genetic engineering, offering unprecedented precision and efficiency in gene editing. The CRISPR/Cas9 system works by leveraging the cell's natural DNA repair mechanisms. Once Cas9 cuts at the targeted DNA loci, the cell attempts to repair the break. There are two primary pathways for this repair: non-homologous end joining (NHEJ) and homology-directed repair (HDR). NHEJ often results in insertions or deletions (indels) at the cut site, which can disrupt gene function (75,76). Alternatively, more precise HDR can be used to insert or replace sequences by providing a DNA template with the desired changes.

The CRISPR-Cas9 system derived from Streptococcus pyogenes utilizes a small non-coding RNA called trans-activating CRISPR RNA (tracrRNA) in the formation of the mature dual-RNA structure of the guide RNA (gRNA). Through complementary pairing with a precursor CRISPR RNA (pre-crRNA), the tracrRNA facilitates RNase III-dependent RNA processing, resulting in the generation of mature crRNA.The gRNA is designed to match a specific DNA sequence in the genome, guiding the Cas9 enzyme to the exact target where a genetic alteration is desired. The crRNA:tracrRNA complex redirects Cas9 toward nearly any target of interest near the PAM sequence by modifying the 20-nucleotide guide sequence within the sgRNA (75). Cas9 nucleases execute precise cutting of the DNA strand by utilizing the HNH and RuvC nuclease domains.

1.7.1 CRISPR/Cas9^{D10A}

To achieve site-specific DNA recognition and cleavage, the protein Cas9 must form a complex with a duplex, crRNA:tracrRNA. The Cas9 enzyme contains two essential nuclease domains, named HNH and RuvC-like that work together to create the DSB (77). Mutations in RuvC-like domain generate a version of the enzyme that lacks cutting ability, known as Cas9 (dCas9). Despite this, dCas9 retains the ability to bind to the crRNA:tracrRNA complex and can still locate and attach to the target DNA sequence. This modified version of Cas9 has been employed in research to mark and visualize repetitive sequences in DNA.

In the mutant form known as Cas9 D10A, which lacks just the RuvC-like nuclease domain activity, the aspartic acid (D) at position 10 of the Cas9 protein is replaced by alanine (A), changing the function of the Cas9 protein: instead of creating double-strand breaks in DNA, it

creates single-strand nicks (78).This modification is often used in a technique called "nickase Cas9," where two Cas9D10A molecules, each with a different guide RNA, are used to create staggered nicks on opposite DNA strands (78). This can improve the precision of genome editing and reduce off-target effects, as the requirement for two guide RNAs to target adjacent sites **Figure 5**.

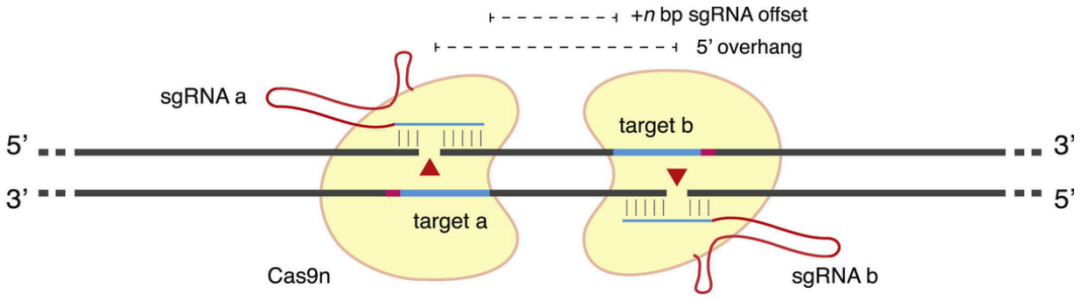


Figure 5 Schematic illustration of double nickase approach.
$Cas9^{D10A}$ is able to cleave only the strand complementary to the sgRNA; a pair of sgRNA-Cas9n complexes can nick both strands simultaneously. sgRNA offset is defined as the distance between the PAM-distal (5′) ends of the guide sequence of a given sgRNA pair (*Ran, F* et all,2013)

2 Aim of the thesis

Pathogenic DNA alterations in the EZH2 gene are frequent in T-ALL patients. Specifically, loss-of-function EZH2 mutations are significantly enriched in T-ALL subgroup at the early stage of T cell development (ETP-ALL). ETP-ALL patients show a very high rate of minimal residual disease response. Thus, more efficient, less toxic targeted therapies are needed to improve patient outcomes. A premise of this PhD thesis is that EZH2 has a tumor suppressor function in T-cell leukemia, leading to transcriptional repression of stem/early progenitor-associated programs driving malignant transformation and disease progression. Based on this premise, we hypothesize that EZH2 inactivation synergistically regulates the gene signatures driven by the stem transcription factor genes to promote the leukemogenesis of human progenitor T-cells and the maintenance and propagation of human T-ALLs. To test this hypothesis, we aim to 1) determine whether CRISPR-Cas9-mediated inactivation of EZH2, alone or in combination with the overexpression of known T-ALL oncogenes, promotes the malignant transformation of human hematopoietic stem/progenitor cells (HSPCs) *in vitro* and *in vivo* assay; 2) characterize the immunophenotypes of *in vitro* derived T cell subsets by multicolor flow cytometry analysis, which allows for the simultaneous detection of several CD markers. 3) combine RNA-Seq and ChIP-Seq to explore how the epigenetic landscape and gene expression are regulated in T-ALLs. We propose to transduce human CD34+ cells from Cord Blood (CB) with lentivectors encoding T-ALL oncogenes alone or in combination with the CRISPR/Cas9-mediated inactivation of EZH2 and to track in *vitro/vivo* phenotypes of cell differentiation and leukemogenesis. T-ALL cell lines will also be genetically manipulated to understand the molecular mechanisms of leukemia maintenance and initiating activity mediated by EZH2 inactivation. These results will elucidate significant aspects of leukemogenesis, and the progression of early T-cell leukemia, thereby identifying new therapeutic targets. Furthermore, these findings will highlight molecular mechanisms by which the epigenetic modifier EZH2 can rewire stem-cell-associated transcription programs, promoting the capacity of tumor cells to maintain their malignant properties and/or resistance to conventional therapeutic approaches. It will also provide important insight for designing novel therapeutic options for the high-risk subgroup of ETP-ALL patients.

.

3 Material and Methods

3.1 Generation of Donor DNA Template and PCR Amplification

For the experimental purpose, a double-strand DNA donor template encoding the human PGK-1 promoter upstream of the GFP coding sequence flanking by approximately 500 base pairs homology arms of Exon 10 of EZH2 was designed. The genomic databases and transcriptome analysis tools (e.g., Ensembl, UCSC Genome Browser) were used to verify that exon 10 is shared across all transcript variants of EZH2. The donor template sequences, already cloned into a pcDNA3.1(+) plasmid (Invitrogen, cat. #V79020), were generated by PCR amplification and subsequently confirmed for accuracy via Sanger sequencing. PCR was performed with Phusion® High-Fidelity DNA Polymerase (NEB) according to the manufacturer's instructions. The double-strand DNA fragments were generated by PCR amplification. PCR was performed with Phusion® High-Fidelity DNA Polymerase (NEB) according to the manufacturer's instructions.The Forward and Reverse primers are listed in Table 1.

PCR reactions were conducted at the following cycling conditions: initial denaturation at 98°C for 3 minutes, followed by 35 cycles of denaturation step at 98°C for 30 seconds, annealing step at 60°C for 30 seconds, extension step at 72°C for 67 seconds. Final extension was conducted at 72°C for 5 minutes. Subsequently, the amplified product was checked on a 1% agarose gel electrophoresis. The PCR product was purified using the QIAquick PCR Purification Kit from QIAGEN. The concentration of the amplified donor DNA was spectrophotometrically quantified using the Nanodrop instrument from Thermofisher.

3.2 Cells and cell culture

3.2.1 Human Hematopoietic Stem Cell Culture (HSPCs)

Human umbilical cord blood samples used for this project were obtained from STEM CELL Technologies (#70008) with informed consent from donors. These samples contain pools of cord blood samples from multiple donors (pool of 100 individuals).

3.2.2 Primary cell culture

Cells were thawed and prestimulated in StemSpan™ CD34+ Expansion Supplement (10X) diluted 1:10 in StemSpan Serum-Free Expansion Medium II (SFEM II, STEMCELL Technologies). For *in vitro* growth cell culture, cord blood samples were resuspended in StemSpan Lymphoid Progenitor Expansion Supplement (10X), diluted 1:10 in StemSpan Medium II (TEM media) or in StemSpan Medium II supplemented with hIGF (10 µg/mL) (PeproTec, cat. #100-12), hFGF (10 µg/mL) (PeproTec, cat. #110-17A), hSCF (10 µg/mL) (PeproTec, cat. #300-07), hTPO (10 µg/mL) (PeproTec, cat. #300-18) and antibiotics according to the manufacturer's protocol. This media will from now on be referred to as Wiekmeijer prestim media (Wiekmeijer et al., 2014). The cells were then maintained in an incubator at 37 °C with 5% CO_2 in a humidified environment.

For Limiting Dilution analysis (LDA) the StemSpan T cell Generation Kit (STEMCELL Technologies) was used. Briefly, non-treated culture dishes were coated with StemSpan Lymphoid Differentiation Coating Material (100X), diluted 1:100 in PBS, and left at room temperature for 2 hours. The coating material was removed and the plates were rinsed with PBS. Then, Cord blood samples were resuspended in TEM media or Wiekmeijer prestim media for 3 weeks.

Human T-ALL cell line LOUCY (CRL-2629) was purchased from American Type Culture Collection (ATCC). Human T-ALL cell line CUTLL-1 (CVCL_4966) was obtained under MTA with INSERM from Columbia University. The cells were cultured in RPMI 1640 media supplemented with 10% FBS (Sigma) with 2 mM Gluta-MAX (Gibco), and antibiotics.

3.3 Limiting dilution analysis (LDA)

CB cells were edited via CRISPR/Cas9 method, then transduced with lentiviral constructs, and finally expanded for 10 days in culture before LDA. Briefly, transduced cells were seeded at variable cell numbers (1, 10, 100, and/or 100 cells) into individual wells of a flat-bottom 96-well plate and cultured in TEM media or Wiekmeijer prestim media for 3 weeks. The colonies were counted and assessed by flow cytometer analysis. The percentage of GFP tag integration was assessed using Droplet Digital PCR and TaqMan technique (Bio-Rad).A threshold yield of 100 cells was used to define positive wells and further calculate “stem cell” frequencies by Extreme limiting dilution analysis (ELDA) (Y. Hu & Smyth, 2009).

3.4 CRISPR/Cas9 Nickase^{D10A}

The CRISPR/Cas9^{D10A} genome editing was performed as described in the "Alt-R CRISPR-Cas9 System: Delivery of ribonucleoprotein complexes into Jurkat & cells using the Neon® Transfection System" (IDT). Two EZH2 crRNAs targeting the coding sequence of exon 10 were chosen based on predicted on-target and off-target scores (IDT gRNA design tool). The components Alt-R CRISPR crRNA sense, Alt-R CRISPR crRNA antisense, and Alt-R tracrRNA (IDT #1073189) were reconstituted in Nuclease-free IDTE buffer pH 7.5 (IDT #330375) to achieve a concentration of 200 µM. The sequences of crRNA sense and antisense are reported in Table 1.The crRNA:tracrRNA complex was prepared from equimolar concentrations of crRNA sense, antisense, and tracrRNA to form the guide RNA mix (36 pmol). Oligonucleotide hybridization was carried out in a thermocycler at 95 °C for 5 minutes, 60 °C for 1 minute, followed by 25 °C for at least 10 minutes. Subsequently, Alt-R S.p. Cas9D10A nickase V3 (IDT #1081062) was diluted in Buffer T (Invitrogen, Thermo Fisher Scientific) to a concentration of 36 µM and then added to the gRNA mix to form the ribonucleoprotein complex (RNP), left at room temperature for 20 minutes.

3.5 Electroporation

Donor DNA previously prepared and the RNP complex were introduced into human CD34+ cord blood cells via electroporation. For each condition, 2.0 × 10^5 cells were centrifuged at 300 × g for 5 minutes, washed with 1 mL of Phosphate Buffered Saline (PBS) (Gibco™ PBS, pH 7.4, cat. #10010023), centrifuged at 300 × g for 5 minutes, and then resuspended in the electroporation mix containing the RNP mix, Electroporation Enhancer oligonucleotide (IDT #1075915) (C_f=10uM), template dsDNA (1ug), and Buffer T to a final volume of 200,000 cells/12 µL. 10 microliters of this solution were used for electroporation, performed at 1600 V, 3 pulses of 10 ms each (Protocol #27182, STEM CELL Technologies), using the Neon™ transfection system (Thermo Fisher Scientific). After electroporation, the cells were transferred to a 48-well plate with StemSpan™ SFEM II medium and StemSpan™ CD34+ Expansion Supplement (1X) for post-editing recovery for approximately 48 hours.

3.6 DNA extraction

DNA was isolated from electroporated cells at different time points using the QIAamp DNA Blood Mini kit from QIAGEN (Cat. No./ID #51104). The quantitative and qualitative analysis of the extracted samples was assessed using a spectrophotometer (Nanodrop).

3.7 Droplet digital PCR

The integration rate of the GFP tag was assessed using the Droplet Digital PCR™ QX200 system (ddPCR Bio-Rad). For experimental purposes, two customized oligonucleotide probes, labeled with FAM for the EZH2 gene and HEX for the GAPDH gene were acquired from Bio-Rad Laboratories. A total of 5-20 ng of dsDNA was generally used as input. The ddPCR was performed according to the manufacturer's protocol on the QX200 Droplet Digital PCR System (Bio-Rad). Primers and Probes are listed in Table 2. Results were analyzed on QuantaSoft™ Analysis Pro software (Bio-Rad). For experimental purposes, the percentage of GFP integration was obtained as a positive fraction of wtExon10 over the internal control (positive signal of the GAPDH amplicon).

3.8 Lentiviral Vectors

HOXA9, LYL1, NOTCH1 (ΔE allele), TLX1, TLX3, TAL1, NKX2-2 were previously cloned into pRRL-cPPT/CTS- MNDU3-PGK-mtagBFP2-WPRE, immediately downstream of the MNDU3 promoter. All constructs were verified by sequencing. Protein overexpression was confirmed via Western blot (WB).

3.8.1 Lentiviral Production and Titration

Lentiviral particles were produced by transient transfection of 293T-HEK cells using polyethylenimine (PEI) with second-generation packaging/envelope vectors, including pCMV-dR8.74 (Addgene #22036), pRSV Rev (Addgene #12253), and pCMV-VSV-G (Addgene #8454). 293T cells were plated in culture at a concentration of 4 × 10^6 cells/dish (10 cm) in DMEM medium (Gibco™, cat. #11965092) with 10% FBS (Gibco™, #16000044) and incubated at 37 °C overnight. The following day, 293T cells were transfected with a DNA mix containing four vectors, as outlined below.

- pCMV-dR8.74 (2.3 μg)
- pRSVRev (0.9 μg)
- pCMV-VSV-G (1.2 μg)
- Vector containing the transcription factor (3.6 μg).

The PEI was added to the DNA mix (at a 2:1 ratio) in a final volume of 1 ml of DMEM. After a 20-minute incubation at room temperature, the mixture was added dropwise to the cells, followed by incubation at 37 °C, 5% CO_2 for 16 hours. A medium change was conducted on the third day. On the final day, the supernatant was collected, filtered through a 0.45 μm membrane (Merk Millipor Ltd, ref. SLHPM33RS), and concentrated at 100,000 × g for 95 minutes at 4 °C using a Beckman Coulter Optima XE Ultracentrifuge with a Type 70i Fixed-Angle Titanium Rotor. Subsequently, the lentiviral particles were resuspended at a 200X concentration in DMEM with 10% DNase 1X and left on a rocker for 45 minutes. Finally, a centrifuge at 12,000 rpm for 10 seconds was performed to remove debris, and the lentiviruses were aliquoted and stored at -80 °C. For lentiviral titering, 1 × 10^5 293T cells were plated per well in a 24-well plate with DMEM 10% FBS. After overnight incubation, Polybrene was added, and the lentivirus was introduced via serial dilution. Cells were incubated for 48 hours at 37 °C with 5% CO2 and on the fourth day, cells were analyzed using flow cytometry (BD FACSCanto II, Becton Dickinson Bioscience) to assess the level of expression of the mTagBFP2 fluorescent tag in live cells by adding the fluorescent DNA dye DRAQ7, in a 1:1000 dilution, to exclude dead cells. Lentiviral titer analysis was performed using FlowJo™ software (BD Bioscience).

3.8.2 Lentiviral Transduction of CD34+ CB cells

CD34+ cord blood cells were seeded into 96-well round bottom plates coated with 5 μg/cm^2 Fibronectin (STEM CELL Technologies #07159) and prestimulated in StemSpan™ CD34+ Expansion Supplement (10X) diluted 1:10 in StemSpan Serum-Free Expansion Medium II (SFEM II, STEMCELL Technologies) for at least 24 hours before virus addition.
The next day, the cells were treated with 8 μM of cyclosporin H (abcam #ab141923) and incubated for 16 hours at 37 °C, 5% CO_2. Subsequently, CB cells were then transduced by direct addition of concentrated viral supernatants and incubated at 37 °C for 6 hours. Transduction efficiency was measured by flow cytometry.

3.9 Flow Cytometry Analysis

3.9.1 Surface staining

Surface antibody staining was performed in 50 µL of antibody mixture (in 0.5% BSA staining buffer) for 20-30 minutes at 4 °C. For surface antibody panels containing more than two or three antibodies, BD Horizon^{TM} Brilliant Stain Buffer was used (BD Biosciences). For all other cases, PBS 3% FBS solution was used as a Staining Buffer. Samples were stained with fluorochrome-conjugated antibodies every 4/5 days with the antibody mastermix. Live/dead cell gating was performed by staining with propidium iodide, DAPI, DRAQ7 Dye (#D15106 Invitrogen), or LIVE/DEAD Fixable Viability Cell 440UV (BD Biosciences). While the analysis was carried out using the FACSCantoTM II flow cytometer or spectral FACSymphonyTM A5 SE (BD Biosciences), the sorting stage was performed on MoFlo Astrios EQ (Beckman Coulter Life Sciences). Flow cytometry data were then analyzed on FlowJo software (BD Biosciences), applying a specific compensation matrix. Data collected from 25 dimensions were reduced to 2 dimensions using the UMAP plugin (version 3.1, FlowJo). An example gating strategy is provided in Appendix A. The early T cell immune panel is reported in Table 3. A complete list of surface antibodies used for spectral flow immune panel can be found in Table 4.

3.9.2 Intracellular staining

T-ALL CUTLL-1 cells were washed with Dulbecco's phosphate-buffered saline (DPBS) and fixed with 2% PFA for 30 minutes at room temperature. Then, cells were washed and resuspended in a permeabilization buffer (0.5% Tween 20 in DPBS/0.5% BSA/3% Rabbit serum, or True-NuclearTM 10X Perm Buffer). Intracellular staining was performed in 50 µL for 1 hour at room temperature, or overnight at 4 °C. After staining, cells were washed with permeabilization buffer and DPBS before flow cytometry.

3.10 Murine Experimental Models

NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG) mice were irradiated with non-lethal doses (200 cGy of X-rays at 150 cGy/min). The injection of 10^5-1- 10^6 human CB cells was performed intravenously by tail vein injection. All experimental manipulations of the mice were carried out by trained personnel operating at the Plaisant animal facility of the Animal Facility in Castel Romano, RM. The mice were checked daily, and their health status was monitored and reported in every aspect. In addition, periodic peripheral blood samples were taken from each mouse, then sacrificed upon reaching predefined human clinical morbidity endpoints or time. All experimental procedures were approved by the Ministry of Health and comply with Italian bioethical requirements: Animal Research (3Rs principles).

3.11 RNA sequencing

3.11.1 Cell Cultures and Cell sorting

ETP-ALL-like LOUCY cells were edited by the described CRISPR/Cas9 method for EZH2 inactivation/GFP+ insertion and subsequently, transduced with lentiviruses as previously described. Following the genetic modification, Loucy cells were cultured in a RPMI medium with 10% FBS and then, FACS-sorted with MoFlo Astrios Cell Sorter (Beckman Coulter, Life Sciences) to achieve a high level of cell purity. The integration efficiency of treated cells was tested by digital droplet PCR (ddPCR). Flow cytometry and immunoblotting assays have confirmed the absence of intracellular EZH2 protein levels.
EZH2-knocked-out/ GFP+ Loucy cells were transduced by spinoculation (800 ×g for 1 h at 35 °C) with viral supernatants in 4 μg mL−1 polybrene.
The transduction was obtained using the condition reported below:

- Empty Vector (EV)/ mTagBFP2;
- HOXA9/mTagBFP2;
- LYL1/mTagBFP2;
- EZH2 wild-type (isoform C)/mTagBFP2 + HOXA9 mcherry
- EZH2 wild-type (isoform C)/mTagBFP2 + LYL1 mcherry

After 48 hours, the efficiency of transfection and expression levels of GFP, mTagBFP and mCherry were assessed using flow cytometry (BD FACSCanto II, Becton Dickinson

Bioscience). Post-transfection, these cells were further expanded and then sorted again to establish pure lines exhibiting the desired selection markers.

3.11.2 RNA extraction

$5x10^6$ cells from three independent experiments of each condition were harvested and the total RNA extraction was performed using total RNA extraction kit (Qiagen). The concentration and the purity of the RNA were spectrophotometrically quantified by QubitTM RNA HS Assay Kit (Thermofisher) and TapeStation High sensitivity RNA ScreenTape (Agilent).

3.11.3 Bioinformatic Analysis

The RNA Preparation, Library Construction and sequencing was performed in service at BMR Genomics Srl, Italy. Libraries were sequenced on an Illumina platform (150PE - 2x20 million reads sequencing). The RNA-sequencing data processing involved two main steps. Initially, alignment of the reads was performed using the STAR 2.7.10a aligner (10.1093/bioinformatics/bts635), targeting the GRCh38/hg38 human genome assembly as the reference genome. Subsequently, quantification of the reads was carried out using Salmon, a tool optimized for accurate transcript quantification. For the integration of gene-level expression estimates within a statistical computing framework Tximport (v1.22.0) was used to load the gene-level expression estimates in R(4.3.1). The matrix containing raw gene counts was exported for analysis in the DEBrowser (v1.29.3) application doi.org/10.1186/s12864-018-5362-x. DEBrowser based on Shiny was utilized for interactive analysis, providing a platform for intuitive data exploration and visualization. Only genes with a row sum equal or greater than 5 were loaded and used as filter cutoff. The data was normalized by MRN (Median Ratio Normalization). Differential expression genes were found by DESeq2 using a parametric method (default) for the type of fitting of dispersions and Wald test used to get p-value.

3.11.4 Genesets for GSEA

TheGenesets
“BILD_HRAS_ONCOGENIC_SIGNA TURE”, and “KRAS.600_UP.V1_UP”.were downloaded from the MSigDB database(http://www.broadinstitute.org/gsea/msigdb/index.jsp) and (Subramanian et al., 2005).

UP_IN_HUMAN_ETP-ALL was constructed from (Zhang et al., 2012a) (GEO Accession: GSE28703). The Dataset was collapsed and the top 500 genes up in ETP compared to Typical T-ALL with a $p<0.05$ were selected.

HSC_human_up was constructed from (Novershtern et al., 2011)(GEO Accession: GSE24759). The Dataset was collapsed and the top 500 genes up in HSC compared to all other tested cell types with a $p<0.05$ were selected.

DN1vsDN2/3_up was constructed by selecting the top 500 genes upregulated in DN1 cells compared to DP cells in the RNA-seq data described in Table S2a of (Zhang et al., 2012b).

3.12 Antibodies

For Immunoblot and Intracellular staining assay the following antibodies were used:

- anti-EZH2 monoclonal antibody (Cod. #5246, Cell Signaling)
- anti-EZH2 polyclonal antibody (Cod. #49-1043, Thermofisher)

3.13 Statistical analysis

Quantitative data were analyzed using GraphPad Prism 9 software. Statistical data were presented as mean ± SEM. NS indicates a p value > 0.05, *p, < 0.05, **p, < 0.01, ***p, < 0.001, and ****p, <0.0001. Statistical significance was determined using a Two-way ANOVA test.

Table 1 List of primers and crRNAs used in CRISPR/Cas9 experiment

Primers and Oligo	
Name	**Sequence (5'-3')**
EZH2_Left Arm forward	AAAGCTAGGTTTACAAATTCACAATATAGAA
EZH2_Right Arm reverse	TTTCTCGAGAGATGAGGACAACTCAAATCCA
crRNA sense CD. Cas9.MJGB5841.AG #1	CATTAATGTGCTGGAATCAA
crRNA antisense CD. Cas9.MJGB5841.AK#2	CTGGGCCTGCTACTGTTATTG
crRNA sense CD. Cas9.TBJK2366.AI #3	GGCTTCCCAATAACAGTAGC
crRNA sense CD. Cas9.TBJK2366.AN #4	TGGGAAGCCGTCCTCTTCTG
pcDNA3.1_Sequencing forward	GGGATTTCCAAGTCTCCAC
pcDNA3.1_Sequencing reverse	CCCAGAATAGAATGACACCTACTC

Table 2 List of primers and probes used in digital droplet PCR (ddPCR)

Digital droplet PCR Primers and probes	
Name	**Sequence (5'-3')**
EZH2_Amplicon forward #1	AGGAGTTTGCTGCTGCTCTC
EZH2_amplicon reverse #2	TCCCTGCTTCCCTATCACTG
EZH2_Amplicon forward #3	GATTATTTGTGATAAATGGATAATGTG
EZH2_amplicon reverse #4	CAGAGGAGCTCGAAGTTTCA
GAPDH_amplicon forward #1	CTACTAGCGGTTTTACGGGCG
GAPDH_amplicon reverse #2	TCGAACAGGAGGAGCAGAGAGCGA
EZH2 (5'-FAM) Probe	AGTAGCAGGCCCAGCACCCCCAC
GAPDH (5'-HEX) Probe	CCGCCGCGCCCCCGGTTTCTATA

Table 3 Early T cell immunopanel used in immunophenotyping experiment

Laser	Detector	Fluorophore	Marker	Clone	Vendor
405	448/59	mTagBFP2			
	710/45				
	755 LP				
488	513/26	GFP			
561	579/16	PE	CD34	581	BD Pharmigen™
	614/20				
	692/75				
640	671/30	APC	CD1a	HI149	BD Pharmigen™
	722/44	AF700, DRAQ7	CD45, DNA (viability)	HI30	BD Pharmigen™
	795/70	PE-Cy7	CD38	HIT2	Biolegend®

Table 4 List of antibodies used in spectral flow cytometry analysis

Detector	Antigen	Cells detected	Clone	Fluorophore	Vendor
B510	GFP				
B537	CD45RA	Lymphoid progenitors, näive T and B cells	HI100	SparkBlue550	BioLegend
B675	TCR γδ	γδ T cells	B1	PerCP-Cy5.5	BioLegend
B810	HLA-DR	APCs	L243	PE-Fire 810	BioLegend
R660	CD8	DP T cells, mature T helper cells, monocytes	RPA-T8	APC	BD
R680	CD33	Myeloid progenitors, HSCs, activated T cells	WM53	AF647	BioLegend
R730	CD45	Hematopoietic cells	HI30	AF700	BioLegend
R780	CD7	T cell progenitors, T cells	M-T701	APC-H7	BioLegend
V450	mtagBFP2				
V470	CD3	T cells	UCHT1	BV480	BD Horizon
V540	CD15	Myeloid cells	W6D3	BV510	BD Horizon
V615	CD10	Lymphoid progenitors	HI10a	BV605	BD Horizon
V660	CD127 (IL-7Rα)	stem cells/precursor	A019D5	BV650	BioLegend
V710	CD38	Lymphoid progenitors	HIT2	BV711	BD Horizon
V785	CD44	T cell subsets	IM7	BV786	BD Horizon
UV379	CD13	Granulocytes, monocytes	WM15	BUV395	BD OptiBuild
UV444	LIVE DEAD				
UV515	CD117	stem cells/precursor	104D2	BUV496	BD OptiBuild
UV585	CD34	HSCs/progenitor cells	581	BUV563	BD OptiBuild
UV660	CD4	DP T cells, mature T helper cells, monocytes	SK3	BUV661	BD Horizon
UV736	CD5	T cells, B cells	UCHT2	BUV737	BD Horizon
UV809	CD11b	Myeloid cells, NK cells, T cell subsets	ICRF44	BUV805	BD OptiBuild
Y585	TCR αβ	αβ T cells	IP26	PE-A	BioLegend
Y660	CD1α	T cells (post-commitment), DCs	H149	PE-eFluor610	ThermoFisher
Y730	CD244	T cells	C1.7	PE-Cy5.5	ThermoFisher
Y825	CD223 (Lag-3)	T cells	7H2C65	PE-Cy7	BioLegend

4 Results

4.1 Optimization of a protocol for the inactivation of EZH2 in human hematopoietic stem/progenitor cells (HSPCs) using the CRISPR/Cas9 system

Frequent loss-of-function mutations in EZH2 were identified in T- ALL patients (14). Moreover, several recurrent mutations in EZH2 are associated with poor prognosis in the aggressive early T-ALL (ETP-ALL), which is characterized by a deregulated expression of stem/early progenitor-associated transcription factors (14,71). Interestingly, in mice, homozygous inactivation of EZH2 induces T-cell leukemia, suggesting a tumor suppressor role of EZH2 in T-ALL (68). However, the precise role of EZH2 in the initiation and development of human T-cell leukemias remains incompletely understood.

To determine how EZH2 inactivation promotes the malignant transformation of hematopoietic stem and progenitor cells (HSPCs), a CRISPR/Cas9-based approach was designed and optimized using both CUTLL-1 cell line and human HSPCs as reported in method (section 3.3) (**Fig. 6A**). In particular, a Cas9 nickase approach was used to enhance the specificity and to reduce the potential off-target effects of the CRISPR/Cas9 system (78). The Cas9n, guided by two gRNAs, creates nicks on both DNA strands. The proximity of these nicks leads to double-strand breaks, which is essential for activating the homology direct Homology-Directed Repair (HDR) mechanism (78,79). For this purpose a pair of sense (S) and antisense (AS) gRNAs were designed, using IDT software (80): In particular, the two designed guides display a 28 bp offset generating 5' overhangs upon nicking (scheme reported in Appendix B)

Additionally, to achieve a precise Homology-Directed Repair, a DNA fragment containing the constitutive human PGK-1 promoter followed by the GFP marker gene and flanked by homology arms of exon 10 of EZH2 was designed according to published protocol instructions (79). To directly quantify and compare the efficiency of CRISPR/Cas9-induced HDR-mediated DNA integration the digital droplet PCR technique was performed. Specifically, two hydrolysis probes were used for HDR detection (FAM, EZH2) and multiplexed with a standard reference gene hydrolysis probe (GAPDH, HEX) (**Fig. 6B**).

A significant decrease in the fraction of wildtype exon 10 of EZH2 was found within GFP+ population when compared to internal control (**Fig. 6C**). To validate digital droplet PCR data,

absolute quantification results were compared with the endogenous levels of EZH2 obtained using intracellular flow cytometric immunostaining. As depicted in (**Fig. 6D**), GFP+ population showed a lower level of EZH2 expression than unmanipulated cells. These data suggest that the CRISPR/Cas9 system successfully achieved the knockout of the EZH2 gene, effectively silencing its expression.

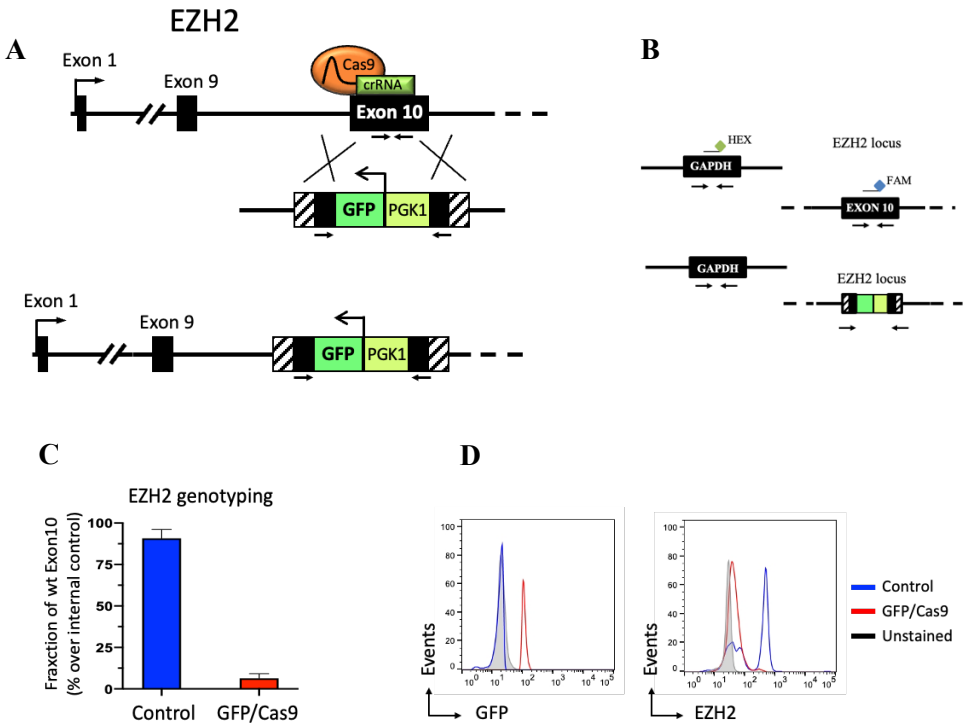


Figure 6 Validation of CRISPR/Cas9-mediated editing of EZH2 gene in human T-ALL cells.
A) Schematic overview of experimental approach. A DNA fragment, including the GFP gene under the constitutive hPGK-1 promoter and flanked with two homologous arms, was integrated into the exon 10 of EZH2 human gene by CRISPR/Cas9-mediated homologous recombination. The insertion of PGK1/GFP fragment resulted in the disruption of exon 10 and loss of EZH2 expression. Primer pairs used for the ddPCR-based genotyping of exon 10 are indicated by small arrows. **B**) General experimental design of digital droplet PCR (ddPCR). **C**) Genotyping of PGK1/GFP insertion into EZH2 locus by digital droplet PCR (ddPCR). In CUTLL-1 T-ALL cell line, the EZH2 gene was edited through direct delivery of PGK1/GFP DNA fragment together with the ribonucleoprotein (RNP) complex consisting of Cas9 protein and single guide RNA (sgRNA) by electroporation. Genomic DNA was purified from CRISPR/Cas9-treated CUTLL-1 cell line after 3 days of *in vitro* growth. In the graph it is reported the fraction of PCR-amplified region in Exon 10 over an internal untreated region located in the GAPDH locus (chr12:6534373-6534538). (**D**) Expression level of GFP and intracellular EZH2 protein in the CRISPR/Cas9-treated CUTLL-1 cell line by flow cytometry.

4.2 EZH2 knockout combined with overexpressing LYL1 and HOXA9 drives expansion of human progenitors T cells *in vitro* cell culture

Given the observation that EZH2 inactivation, in mouse model, leads to the development of T-cell leukemias (68,81), and the pathogenic mutations are common in patients with T-ALL, one of the aims of this research was to determine whether CRISPR-Cas9-mediated inactivation of EZH2, either individually or in combination with overexpression of known T-ALL oncogenes (71) could lead to the malignant transformation of hematopoietic stem/progenitor cells HSPCs, thereby recapitulating the origin of T-ALL. To test this hypothesis, a recently developed, efficient, and reproducible human model of T-ALL from primary human CD34+ cord blood was used (82). This model involves lentiviral transduction of CD34+ cord blood (CB) cells obtained from healthy donors with different combinations of T- ALL oncogenes.

In this context, CD34+ cells were pre-stimulated and then edited by the described CRISPR/Cas9 method to inactivate EZH2. Subsequently, EZH2-edited cells were transduced with lentiviruses containing HOXA9, LYL1, TLX1, TLX3, a dominant active form of NOTCH1 (NOTCH1ΔE allele) (82), TAL1 and NKX2.1 and cultured in the StemSpan SFEM II serum-free media (Stemcell Technologies) with StemSpan T Cell Progenitor Expansion Supplement to evaluate their *in vitro* long-term expansion and selection up to 31 days (**Fig. 7A**). To establish a baseline for cell growth and behavior in the absence of the gene of interest, an empty vector as control was included.

By flow cytometric tracking, the mTagBFP2+/GFP+ (referred to as B+G+) cell fraction in the EZH2 edited overexpressing LYL1 or HOXA9-transduced HSPCs, progressively expanded within 31 days when compared to $EZH2^{KO}$ + Empty. V (p-value <0.0001, Two-way ANOVA test). This effect was less pronounced in EZH2 knockout condition, alone or in combination with the overexpression of NOTCH1ΔE, TLX1, and TLX3 oncogenes. Moreover, overexpression of TAL1 or NKX2-1 showed a strong reduction in transduced cells with any of the three independent experiments (**Fig. 7B**).

Although less informative, the absolute growth curve of EZH2 edited overexpressing HOXA9 or LYL1 cells reached a 10^7-fold expansion by day 31, especially when compared to HOXA9 or LYL1 overexpression in a context of $EZH2^{WT}$ allele (p-value <0.0001, Two-way ANOVA

test). Conversely, HSPCs that were EZH2-edited and/or TAL1/NKX2-1-transduced displayed a significant detrimental impact on cell fitness over time (**Fig. 7C**).

To further support these results, a digital droplet PCR assay was utilized to accurately track the levels of integration of the fluorescent marker GFP at exon 10 of EZH2 from the same samples on days 10 and 31 of cell culture. Strong positive concordance between integration rate and GFP expression was found, suggesting that the effect observed was on-target (Pearson corr. = 0.73, p-value 5.4E-05) (**Fig. 7D**).

Taken together, these results suggest that the knockout or knockdown of EZH2 could synergistically cooperate with LYL1 or HOXA9 transcription factors, making stem and progenitor cells more responsive at the earliest stages of leukemic transformation. This conclusion is supported by several lines of evidence. EZH2 has been previously shown to bind to HOXA-cluster genes through chromatin reorganization and regulate positively its expression in T-ALL (83). Additionally, mouse T-ALLs resulting from loss of EZH2 exhibited an increased expression of two genes associated with stem cells and early progenitors, HOXA9 and LMO2/LYL1, thereby promoting early T-cell leukemia (84). Finally, in models of human T-ALL, HOXA9 or LYL1 induced by the loss of PRC2 activity promotes a stem cell-like or progenitor state similar to the immature T-ALL subtype (41). However, it is important to note that these transcription factors have distinct roles in hematopoietic differentiation and early T-cell development, making the comparison less informative. Despite these differences, evidence shows that epigenetic changes could significantly reshape the chromatin accessibility influencing gene expression of stem cell patterns associated with T-cell development and leukemic transformation (85).

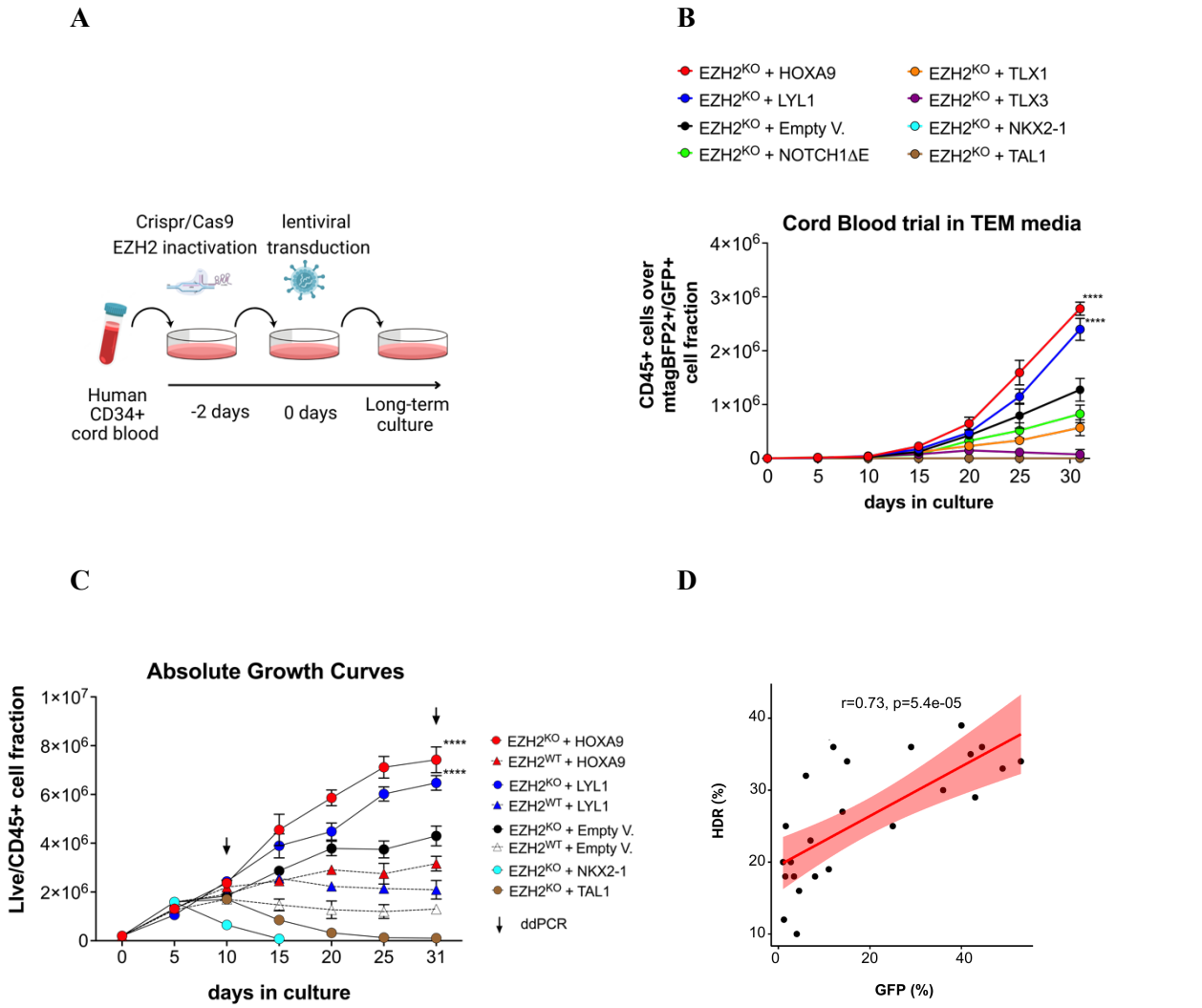


Figure 7 CRISPR/CAS9-mediated EZH2 inactivation in combination with HOXA9 or LYL1 overexpression, promotes cell expansion in TEM media

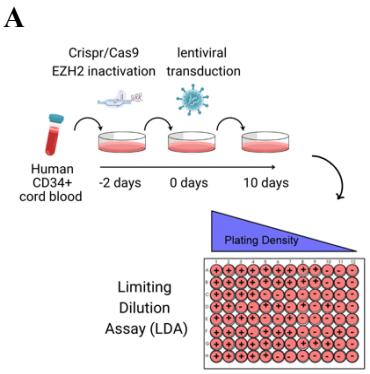
A) Experimental workflow. CD34+ HSPCs from cord blood were isolated, edited by the described CRISPR/CAS9 method and/or transduced with HOXA9, LYL1, TLX1, TLX3, TAL1, NKX2-1, NOTCH1ΔE or Empty vector control on day 0. Treated cells were cultured in a serum-free StemSpan SFEM II medium supplemented with StemSpan T Cell Progenitor Expansion Supplement (StemCell Technologies) for 31 days. **B**) Growth curves of hCD45+/live cells, normalized over the percentage of mtagBFP2+/GFP+ cell fraction. Plots show values from three biological replicates. **** p-value <0.0001. Statistical analysis was performed by Two-way ANOVA. An example of cytometer gating strategy is presented in (Appendix A). **C**) Absolute growth of treated cells. The total number of cells is estimated by calculating the multiplicative increase in cell yield at every passage. **** p-value <0.0001 by Two-way ANOVA. **D**) Pearson correlation plot. ddPCR was performed to detect the integration of the PGK-GFP cassette in exon 10 of the EZH2 locus. ddPCR is a direct measure of the percentage of cells that have integrated a correct PGK-GFP cassette in the EZH2 locus (HDR%). Analysis of GFP expression by flow cytometry confirmed the results obtained by ddPCR with a Pearson correlation of 0.73. Statistical analysis was performed by GraphPad Software.

4.3 Limiting dilution analysis (LDA) shows increased leukemia clonogenic capacity for EZH2 KO overexpressing LYL1 and HOXA9 oncogenes

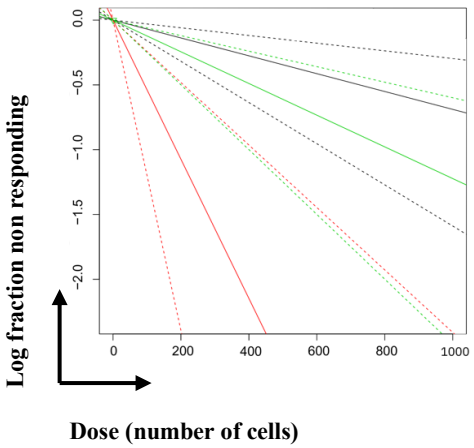
To further explore the relevance of the potential positive interaction between the knockout of EZH2 and the overexpression of HOXA9 or LYL1, a Limiting Dilution Analysis (LDA) was conducted. Specifically, EZH2-edited and/or LYL1/HOXA9-transduced HSPCs were placed into individual wells of a 96-well plate on bare plastic with decreasing number of cells in each row of 12 wells (1000, 100, 10, and 1 cells/well) after 10 days of lentiviral transduction and cultured in a StemSpan T-cell expansion serum-free media for almost 3 weeks. Subsequently, each well was assessed for the formation of tumor colonies. (**Fig. 8A**)

The number of wells containing colonies in each group was visually inspected under a microscope, and wells were classified as positive or negative upon the basis of cell growth. To verify that was accurate, all wells identified as positive were harvested and checked by flow cytometric for B+/G+ tracking. Afterwards, the integration efficiency of treated colony was tested by digital droplet PCR (ddPCR). Finally, the data were analyzed using the extreme limiting dilution analysis (ELDA) software (86). The repopulating ability displayed an approximately 10-fold and 6-fold increase in stem cell frequency for $EZH2^{KO}$ + LYL1 and $EZH2^{KO}$ + HOXA9, respectively (**Fig. 8B-C**).

Although LYL1 condition showed the most pronounced effect in terms of culture-initiating cell frequencies significant (1/186) for EZH2 ^{KO} + LYL1 compared to $EZH2^{WT}$ + LYL1 (1/1751) and $EZH2^{KO}$ + Empty V. (1/818), EZH2 edited overexpressing HOXA9 cells also showed a statistically significant difference when compared to the HOXA9 condition in $EZH2^{WT}$ context. (1/308) in EZH2 ^{KO} + HOXA9 compared to $EZH2^{WT}$ + HOXA9 (1/1863) and $EZH2^{KO}$ + Empty V. (1/818).

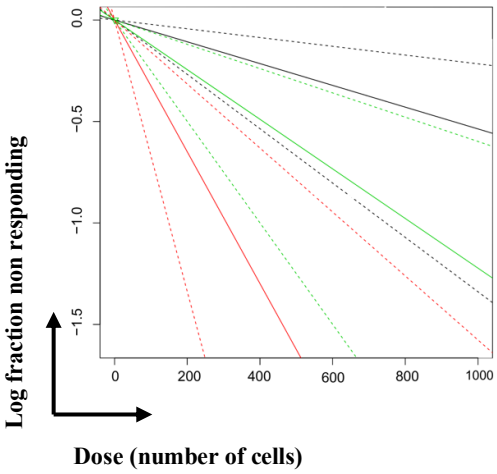


B



Group	stem cell frequency	95% CI
EZH2 ^{KO} + LYL1***	1/186	1/415-83.6
EZH2^{WT} + LYL1	1/1751	1/3354-628.4
EZH2 ^{KO} + Empty V.	1/818	1/1671-400

C



Group	stem cell frequency	95% CI
EZH2 ^{KO} + HOXA9***	1/308	1/634-150
EZH2^{WT} + HOXA9	1/1863	1/4644-747
EZH2 ^{KO} + Empty V.	1/818	1/1617-400

Figure 8 EZH2 knock-down or knock-out increases leukemic cell fitness in different contexts. Log-fraction plot and calculated stem cell frequency of LDA results are presented. Unsorted cells were plated at limiting dilutions 10 days after lentiviral transduction and cultured for 3 days in TEM media. The stem cell frequency was then calculated by ELDA. Mean ± 95% CI of three biological replicates are shown.*** $p<0.001$. Statistical analysis was performed by Two-way ANOVA (GraphPad Software)

4.4 Immunophenotypes of *in vitro* CB-derived early T-cell progenitors

To characterize the immunophenotypic profiles of human manipulated T-cell progenitors, three independent samples were collected and analyzed by flow cytometry. A panel of markers was chosen to resolve various early T-cell populations at 10, 20, and 31 days (Table 3). Included among the markers analyzed were CD34 and CD38, which are sequentially expressed during the early development of T cells, and CD1a, expressed in the later stages of T cell differentiation (18,20). When compared to the control condition, EZH2-edited overexpressing LYL1 cell fraction (B+/G+) resulted in a more pronounced enrichment of early T cell population (CD34+CD38-CD1a-), specifically classified as Double Negative 1 (DN1) thymocytes. Increased number and proportion of DN1 thymocytes were observed starting from the 20th day (mean 9.2 ± 1.0 vs 3.55 ± 1.2), culminating in a more distinct peak on the 31 days (mean 19.33 ± 2.1 vs 7 ± 1.23) (**Fig. 9A**).

The overexpression of HOXA9 in EZH2 edited cells showed an expansion of the DN1 stage at the same time points, but no significant difference was observed on days 10 and 20. Though the effect was less marked, a significant difference was observed on day 31(mean 14 ± 2.5 vs 9.66 ± 1.2) (**Fig. 9B**). Additionally, neither of the two conditions exhibited a clear positivity for the CD1a marker suggesting that these subpopulations not yet be committed to the T-lineage.

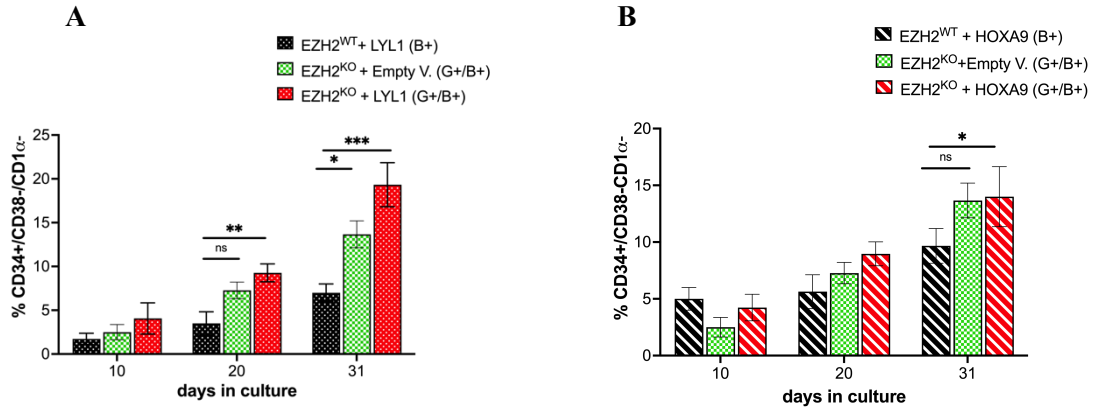


Figure 9 Flow cytometric immunophenotypes of manipulated cells cultured in TEM media
Immunophenotypic assessment of CD34+ CD38-CD1a- early T-cells at the double negative 1 (DN1) development stage in double modified cell populations by flow cytometry.**A**) $EZH2^{KO}$ +LYL1, vs $EZH2^{KO}$ + Empty V. and $EZH2^{WT}$ +LYL1 **B**) $EZH2^{KO}$ +HOXA9 vs $EZH2^{KO}$ + Empty V. and $EZH2^{WT}$ +HOXA9. * p value < 0.05. ** p value< 0.01. *** p value <0. 001.ns, non significant. Statistical analysis was performed by Two-way ANOVA (GraphPad Software).

Subsequently, to further corroborate the interpretation of the data and deeper understand the immunophenotype changes of subpopulations within the transduced CB cells, a spectral flow cytometry analysis was performed. Specifically, the phenotypes of LYL1 or HOXA9-transduced HSPCs, in an EZH2 knockout or wild type context, were compared at the same time points of bulk *in vitro* cell cultures (only 31 days showed). As 23 markers were used including (myeloid and T-cell differentiation markers, as well as markers of hematopoietic progenitor and stem cells), a Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction was performed. FlowSom clustering algorithm was also used to classify cells based on phenotype and each cell was grouped into major clusters (**Fig. 10A**). The expression patterns of CD markers (proportion and intensity of positive fraction) were displayed across each cluster, providing a general overview into the distinct stages of manipulated HSPCs cell differentiation (**Fig. 10B**).

This analysis showed that all clusters displayed a variety of CD marker expressions, suggesting that the cellular populations within each cluster are not homogeneous but rather exhibit a broad spectrum of phenotypic characteristics that were challenging to identify. However, when comparing the cluster composition from each population grouped by conditions ($EZH2^{WT}$ + LYL1, $EZH2^{KO}$ + Empty V. and $EZH2^{KO}$ + LYL1), a significant difference in percent of cells in each cluster was observed. Specifically, clusters #2, and #3, mainly included cells from both $EZH2^{KO}$ + Empty V. and $EZH2^{KO}$ + LYL1, with clusters #4 and #5 composed of $EZH2^{KO}$ + LYL1 cells (64.3% and 100%, respectively) (**Fig. 10C, D**). These clusters were characterized by the enrichment of early T-cell progenitors (CD7+CD5-) and the almost complete absence of committed T-cell progenitors (CD7+CD5+CD1a+). Moreover, an aberrant expression of CD4 was observed in clusters #2 and #3. A deeper examination of these populations revealed that myeloid/stem cell receptor c-Kit (CD117) exhibited high levels of expression (proportion and intensity) exclusively within clusters #4 and #5 (**Fig. 10B-E**).

Of note, both clusters also exhibited the highest expression level of CD44 and CD7, primarily expressed in early T-cell progenitors, accompanied by a lack of CD1a and CD4/CD8 expression markers, suggesting that these populations represented very primitive stages of T-cell development. On the other hand, the co-expression of myeloid/stem markers (such as CD33, CD15, or CD13) may indicate an aberrant phenotype with signatures similar to T/Myeloid Mixed phenotype.

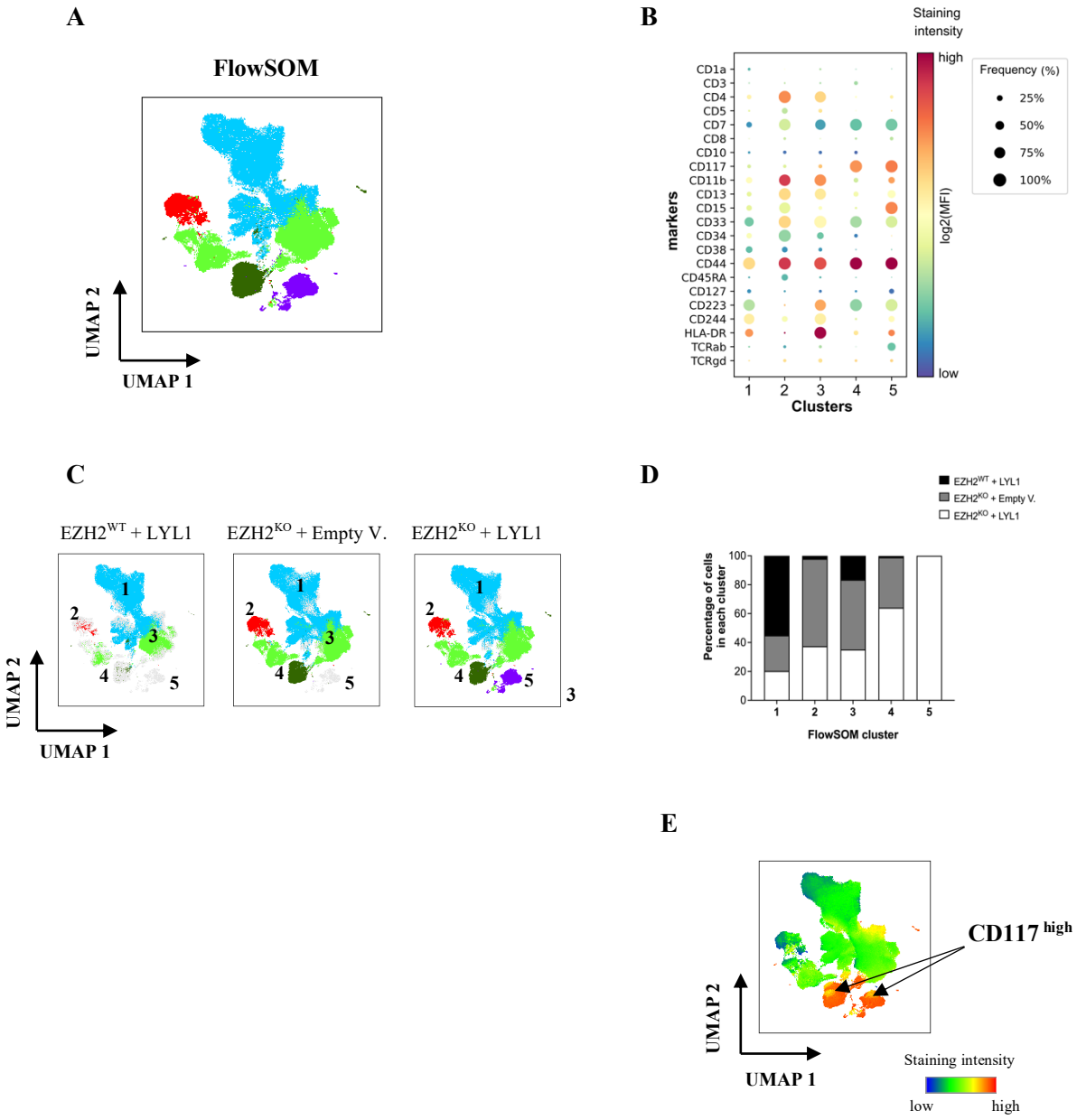


Figure 10 Multiparameter analysis of EZH2-edited and/or LYL1 overexpressed HPSCs by spectral flow cytometry.

A) 23-parameter UMAP distribution of flow cytometric immunophenotypes of manipulated cells cultured in early T-cell expansion media for 31 days. FlowSOM clustering was performed and resulted in 5 phenotypic clusters (1-5). **B**) Dot plot indicating the expression of CD markers for each UMAP cluster. Frequency indicates the percent of cells positive for each CD marker based on manual gating and is displayed by the size of each dot (red = high staining intensity blue = low staining intensity). **C**) Edited and/or transduced HSPCs projected into UMAP space and segregated according to the oncogenes overexpressing. **D**) Bar plot indicating the percent of cells in each FlowSOM cluster. **E**) UMAP heatmap showing the levels of CD117 expression intensity.

Similar results were observed in the EZH2 edited overexpressing HOXA9 cells (**Fig. 11A-E**). Specifically, Cluster #5 exhibited the highest expression of CD117 and CD44 followed by CD33 and CD7. Once again, the immunophenotype showed enrichment of cells resembling early T-cell progenitors (CD7+CD5-) or myeloid-skewed progenitors (CD33+CD7+) and an almost complete absence of both committed T-cell progenitors(CD7+CD5+CD1a+) and CD4/CD8 expression. However, a spectrum of CD marker expression was observed, making it difficult to precisely distinguish differentiation stages within individual clusters.

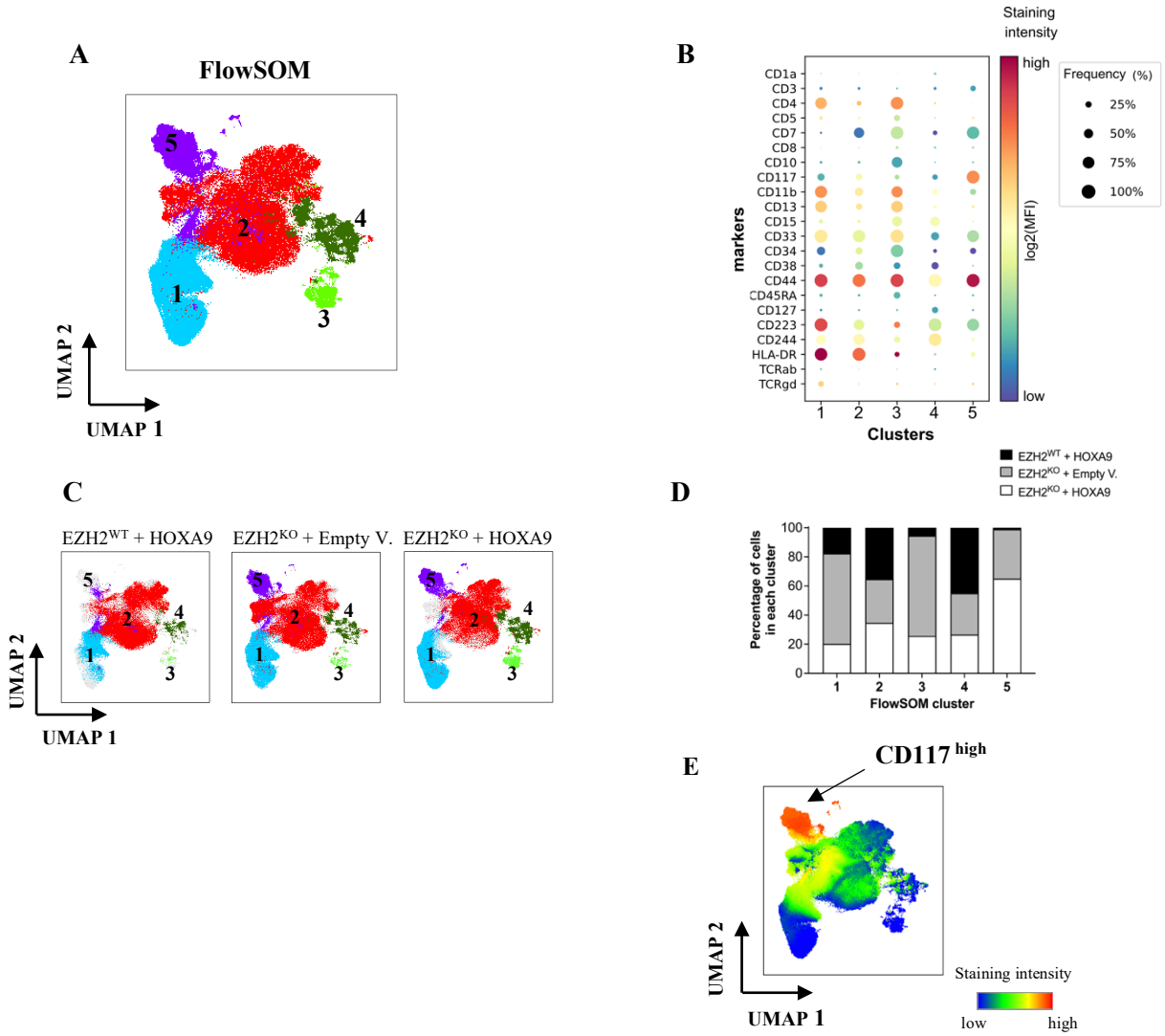


Figure 11 Multiparameter analysis of EZH2-edited and/or HOXA9 overexpressed HPSCs by spectral flow cytometry.

A) 23-parameter UMAP distribution of flow cytometric immunophenotypes of manipulated cells cultured in early T-cell expansion media for 31 days. FlowSOM clustering was performed and resulted in 5 phenotypic clusters (1-5). **B**) Dot plot indicating the expression of CD markers for each UMAP cluster. Frequency indicates the percent of cells positive for each CD marker based on manual gating and is displayed by the size of each dot (red = high staining intensity blue = low staining intensity). **C**) Edited and /or transduced HSPCs projected into UMAP space and segregated according to the oncogenes overexpressing. **D**) Bar plot indicating the percent of cells in each FlowSOM cluster. **E**) UMAP heatmap showing the levels of CD117 expression intensity

The c-Kit receptor, also known as CD117, is normally expressed in hematopoietic stem cells, multipotent progenitors, common lymphoid progenitors, and early-stage thymocytes. Recently, it has been suggested that increased expression of CD117 is associated with stem-like phenotype and worse clinical outcomes in ETP T-ALL (88,89). In mouse T-ALL models, EZH2 inactivation promotes *in vivo* immature immunophenotypes with CD117 stem-like features, suggesting that EZH2 loss induces the differentiation block at early stages in mouse T-cell leukemia (84). Increased expression of CD117 is associated with stem-like phenotype and worse clinical outcomes in human ETP T-ALL (67). Furthermore, our laboratory recently demonstrated that some human primary T-ALLs characterized by CD117+CD82+ yielded highly penetrant and serially transplantable leukemias when compared to negative cell fractions (90).

Based on these observations, it was hypothesized that CD117 can promote the leukemogenesis of human progenitor T- cells and the maintenance and propagation of human T-ALLs. To test this hypothesis CD117 was included as a CD marker for immunophenotyping transformed HSPCs. By tracking through flow cytometry, the positive population of CD117 increased progressively, reaching a peak of 25 % by day 31 for the EZH2 knockout overexpressing LYL1 cells, while in the HOXA9 condition, the increase was up to 22%, and for EZH2 knockout, it was 18%. Conversely, the only transduced HSPCs failed to show any significant increase, with an average percentage drop ranging from 20% to about 2% by the end of cell culture expansion. (**Fig. 12,13A**).

Furthermore, CD117 positive population in $EZH2^{KO}$ + LYL1 condition exhibited a bright expression that peaked on the thirty-first day. Even though the effect is reduced, the same trend was observed for the knockout of EZH2 overexpressing HOXA9 cells. Interestingly, a significant decrease in intensity up to a complete loss of CD117 was found in un-edited HSPCs, suggesting that the expression of CD117 might be a distinctive characteristic of EZH2 knockout (**Fig. 12, 13B**).

Finally, the immunophenotype of CD117 positive population showed for both conditions the highest level of CD7+, and then CD44+, CD1a- (**Fig.12C, 13C**). These markers were chosen as previous studies have correlated T-cell commitment in human thymocytes with expression of CD7 and CD1a, and more recently, the loss of CD44 expression, which precedes gain of CD1a (7,87). Taken together these results suggest that EZH2 knockout can block the expansion of mature T-cells resulting in a skewing toward the myeloid lineage.

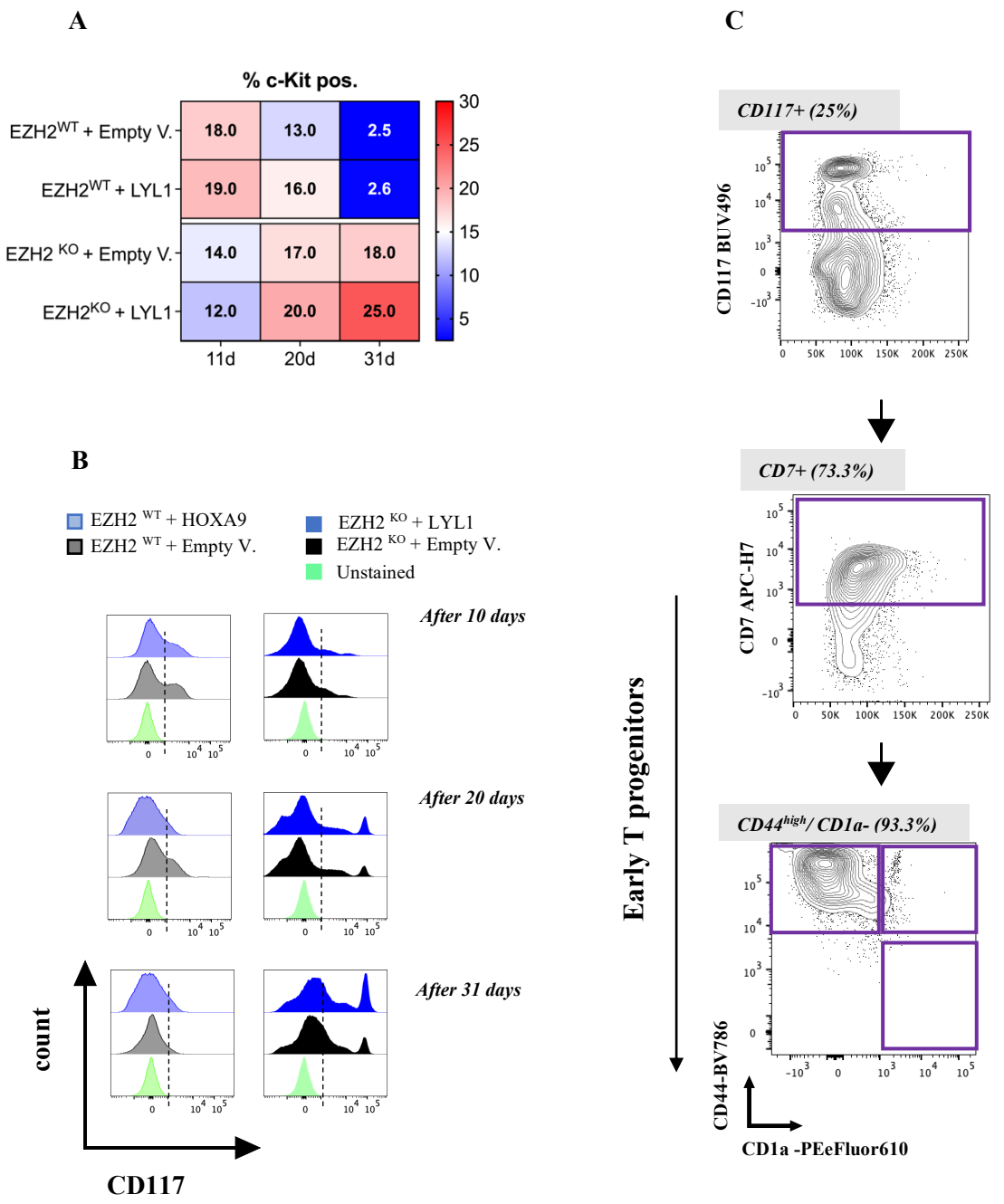


Figure 12 CD117 is expressed in EZH2 knockout overexpressing LYL1 within early T cell progenitors

A) Percentage of positive cells for CD117 marker on EZH2-edited and/or LYL1 transduced human CB cells at 10,20- and 31 days post-transduction. **B**) Overlay histograms showing CD117 positive population in CD45/mtagBFP2 positive cells. **C**) Contour plots indicating the proportion of early T progenitor cells (CD7+CD44+CD1a-) gated on CD117 positive population as indicated. Human CD34+ CB cells were cultured in the StemSpan SFEM II serum-free media with StemSpan T Cell Progenitor Expansion Supplement (Stemcell Technologies

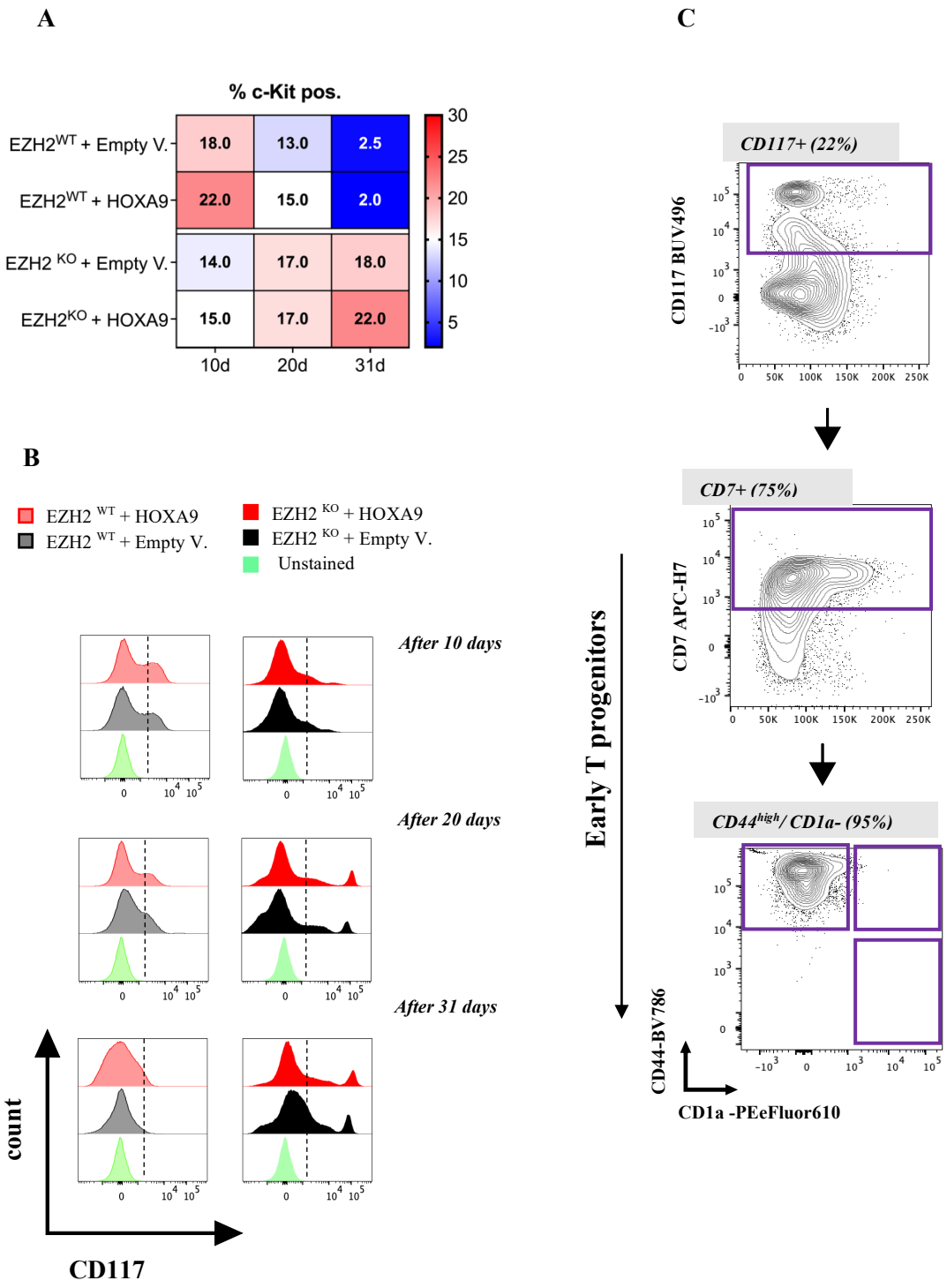


Figure 13 CD117 is expressed in EZH2 knockout overexpressing HOXA9 within early T cell progenitors

A) Percentage of positive cells for CD117 marker on EZH2-edited and/or HOXA9 transduced human CB cells at 10,20- and 31 days post-transduction. **B**) Overlay histograms showing CD117 expression in CD45/mtagBFP2 positive cells. **C**) Contour plots indicating the proportion of early T progenitor cells (CD7+CD44+CD1a-) gated on CD117 positive population as indicated. Human CD34+ CB cells were cultured in the StemSpan SFEM II serum-free media with StemSpan T Cell Progenitor Expansion Supplement (Stemcell Technologies)

4.5 EZH2 Knock-out cells overexpressing HOXA9 or LYL1 generate Leukemia *in vivo*

Given the apparent synergy observed in the transformation of human early progenitor T cells *in vitro* assays, the leukemogenic capacity of EZH2 knock-out in the HOXA9 or LYL1 transduced HSPCs was assessed *in vivo* environment. Human CD34+ CB cells were EZH2 edited and transduced with lentivectors containing HOXA9 or LYL1 or Empty Vector as control. After 10 days of cell culture in TEM media, 10^5 engineered treated cells were injected into five immunocompromised mice for each condition. Progenitor T cells with normal EZH2 expression and transduced with lentiviruses expressing HOXA9, LYL1, and Empty Vector were also injected as a control. Of note, the injected cells included a mixture of non-transduced (B-G−), singly transduced (B+G− and B- G+), and doubly transduced (B+G+) populations. Around 250-300 days post-injection, EZH2 knockout overexpressing LYL1 or HOXA9 conditions showed a higher penetrance when compared to HOXA9-only and LYL1-only leukemia. Of note, in the cohort of EZH2^{KO} + HOXA9 and EZH2^{KO}+ LYL1 injected mice, the median survival differed significantly, from 116 (EZH2^{KO} + HOXA9) to 209 (EZH2^{KO} + LYL1) days, as detailed in (**Fig. 14)**. Collectively, these findings suggest that HOXA9 or LYL1 overexpression may lead to enhanced fitness in a deficient EZH2 context while the only EZH2 knockout, EZH2 wildtype overexpressing HOXA9 or LYL1 are not sufficient to generate penetrant T-cell leukemias *in vivo*.

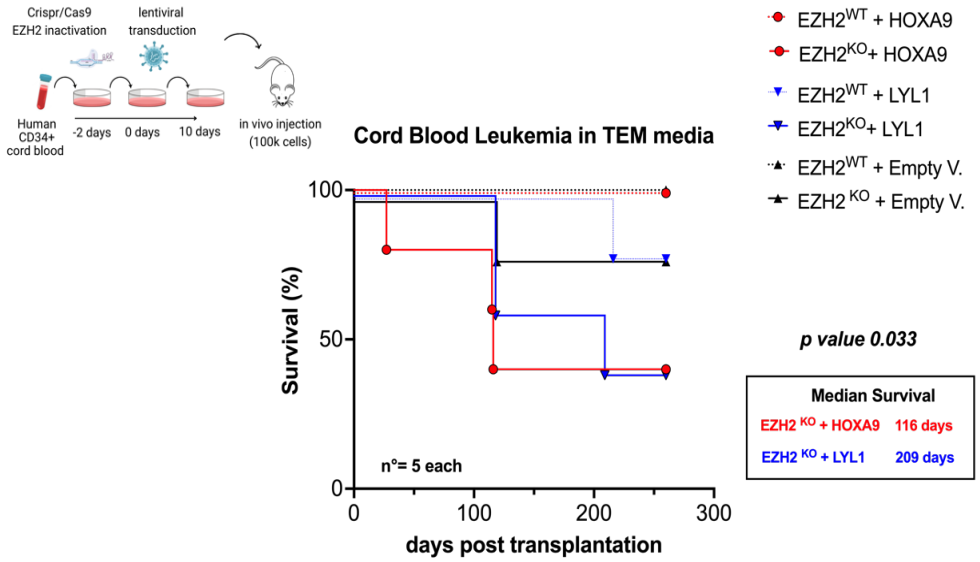


Figure 14 EZH2 knockout CD34+ cord blood cells overexpressing HOXA9 or LYL1 develop leukemia at high penetrance

Leukemia-free survival of mice transplanted with 100,000 EZH2 ^{KO}+HOXA9 cells, EZH2 ^{KO}+LYL1, EZH2 ^{KO} +Empty V., EZH2 ^{WT}+ HOXA9, EZH2 ^{WT}+ LYL1 or EZH2 ^{WT} +Empty V., cells. Median of survival is shown. $p<0.05$. Statistical analysis was performed by Log-rank (Mantel-Cox) test.

4.6 EZH2 Knockout is required for cell growth of transduced HSPCs in a media for hematopoietic stem cell expansion

Since the knockout of EZH2, in combination with HOXA9 or LYL1 overexpression, led to the high expression of CD117 in lymphoid progenitor media (TEM media), it was hypothesized that EZH2 could potentially delay or even block the commitment of T-cell lineage, stimulating the proliferation of hematopoietic stem cells. This idea was motivated by three main observations: first, the inhibition of histone methylation, specifically via EZH2 which catalyzes the trimethylation of histone H3 on lysine 27 (H3K27me3), impacts the lineage commitment in progenitor cells, potentially leading to alterations in the balance between different hematopoietic lineages (91). Second, changes in the epigenetic landscape, particularly influenced by EZH2 activity, have been associated with shifts in cell fate decisions (92), and third, EZH2 is frequently mutated or dysregulated in human T-ALL leukemia, including a subtype of T-acute lymphoblastic leukemia (T-ALL) that exhibits myeloid characteristics, suggesting a tumor suppressor role of EZH2 of early T-ALL (67,71).

Given that, the effect of EZH2 silencing on the early transformation stages of CB-derived T-ALL model was tested using a StemSpan^{TM} SFEM II medium which is specifically formulated to support the expansion of hematopoietic stem cells (HSPCs). It was also enriched with a cocktail of cytokines to stimulate cell growth and differentiation (Wiekmeijer prestim media). Specifically, EZH2-edited cells were transduced with HOXA9, LYL1, TLX1, TLX3, NOTCH1ΔE, lentiviruses or empty vectors as control, and cultured up to 31 days (**Fig. 15A**). EZH2 edited cells overexpressing LYL1 or HOXA9 rapidly transformed, got outcompeted by BFP-/GFP- cell fraction, and expanded over 31 days when compared to EZH2^{KO} + Empty.V (p-value <0.0001, Two-way ANOVA test) (**Fig. 15B**). On the other hand, other combinations of T ALL oncogenes showed a lower rate of expansion and selection over time suggesting that EZH2 inactivation, combined with deregulated expression of LYL1 or HOXA9, can promote the leukemogenesis of human progenitor T- cells (**Fig. 15B, C**). Finally, a strong positive concordance between integration rate and GFP expression was found, suggesting that the effect observed was on-target (Pearson corr. = 0.79, p-value 1.7e-05) (**Fig. 15D**).

To further support these data, the repopulating ability of the transduced cells was assessed by a 3-week. Limiting Dilution analysis (LDA), following the method reported in section 3.2, showed a stem cell frequency statistically significant (1/86.1) for EZH2 ^{KO} + LYL1, compared to EZH2^{WT} + LYL1 (1/413.5) and EZH2^{KO} + Empty V. (1/3336.6) (**Fig16A**).

Although less intense, the same trend was observed for the HOXA9 condition. The stem cell frequency was (1/724) in EZH2 ^{KO} + HOXA9 condition compared to EZH2^{WT} + HOXA9 (1/1875) and EZH2^{KO} + Empty V. (1/3336.6), showing less than 5-fold difference in both conditions (**Fig.16B**).

To verify if the increased cycling capacity and stem cell frequency seen *in vitro* would translate into more penetrant leukemias *in vivo* assay, 10^5 EZH2-edited and/or LYL1 or HOXA9 transduced HSPCs and an Empty Vector as control were injected into 5 NGS mice for each condition after 10 days of cell culture in Wiekmeijer prestim media, as previously reported. EZH2^{KO}+ LYL1 condition showed very high penetrance, while the effect in EZH2^{KO} + HOXA9 was less pronounced (100% and 60%, respectively). Also, the median survival differed significantly, from 120 for EZH2^{KO} + HOXA9 to 51 for EZH2^{KO} + LYL1 days, as detailed in (**Fig.17)**. These data suggest that the inactivation of EZH2 in combination with LYL1 oncogene increased fitness of transformed cells in vivo, generating highly penetrant leukemias.

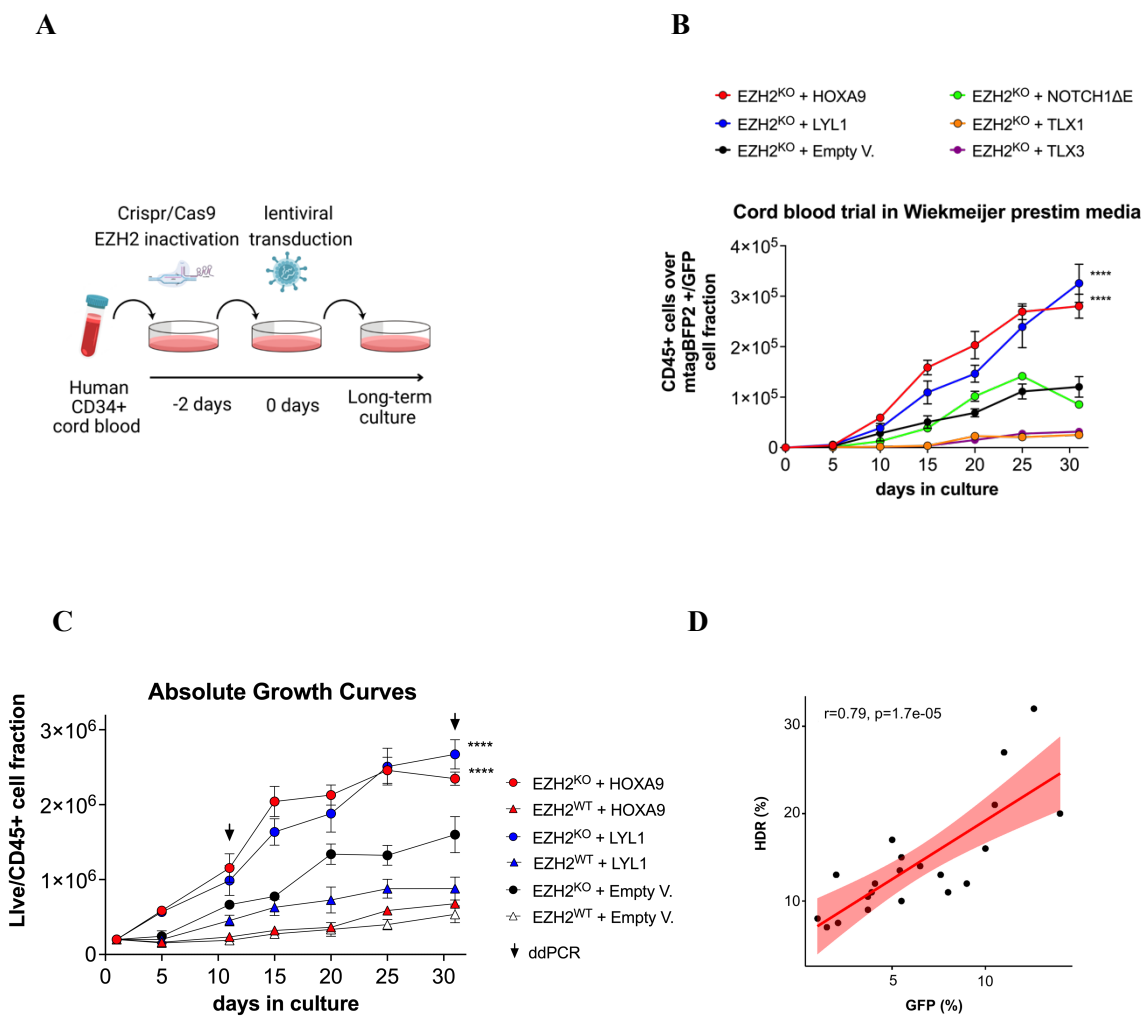
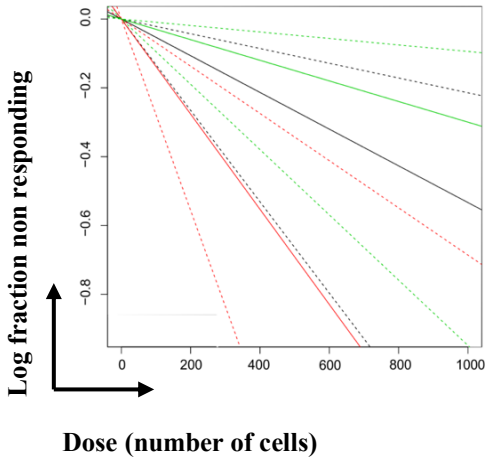


Figure 15 EZH2 edited and/or transduced LYL1/ HOXA9 HSPCs promotes cell growth in Stem media

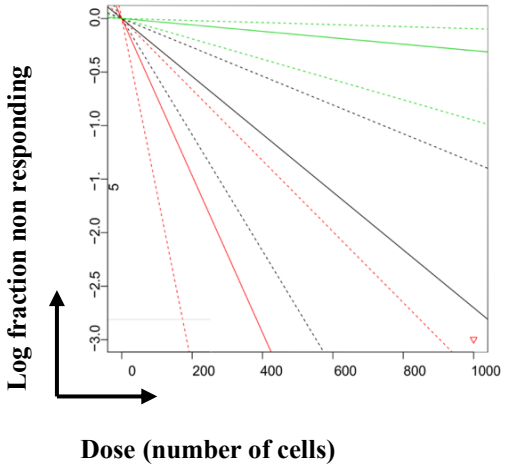
A) Experimental workflow. CD34+ HSPCs from cord blood were isolated, edited by the described CRISPR/CAS9 method and/or transduced with HOXA9, LYL1, TLX1, TLX3, N1dE or Empty vector control on day 0. Treated cells were cultured in Wiekmeijer prestim media for 31 days. **B**) Growth curves of hCD45+/live cells, normalized over the percentage of mtagBFP2+/GFP+ cell fraction. Plots show values from three biological replicates. **** p-value <0.0001. Statistical analysis was performed by Two-way ANOVA (GraphPad Software). **C**) Absolute growth of treated cells. The total number of cells is estimated by calculating the multiplicative increase in cell yield at every passage. **** p-value <0.0001 by Two-way ANOVA. **D**) Pearson correlation plot. ddPCR was performed to detect the integration of the PGK-GFP cassette in exon 10 of the EZH2 locus. ddPCR is a direct measure of the percentage of cells that have integrated a correct PGK-GFP cassette in the EZH2 locus (HDR%). Analysis of GFP expression by flow cytometry confirmed the results obtained by ddPCR with a Pearson correlation of 0.79 p-value 1.7e-05. Statistical analysis was performed by GraphPad Software

A



Group	stem cell frequency	95% CI
EZH2 ^{KO} + LYL1***	1/86.1***	1/177-41.8
EZH2^{WT} + LYL1	1/413.5	1/823-207.5
EZH2 ^{KO} + Empty V.	1/3336.6	1/10551-1055.2

B



Group	stem cell frequency	95% CI
EZH2 ^{KO} + HOXA9*	1/724*	1/1457-359
EZH2^{WT} + HOXA9	1/1875	1/4670-753
EZH2 ^{KO} + Empty V.	1/3336.6	1/105551-1055

Figure 16 *EZH2* knock-down or knock-out increases leukemic cell fitness in different contexts
A) Log-fraction plot and calculated stem cell frequency of LDA results for **A**) EZH2 ^{KO}+LYL1 vs EZH2 ^{WT}+LYL1 and EZH2 ^{KO}+ Empty V. **B**) EZH2 ^{KO}+HOXA9 vs EZH2 ^{WT}+HOXA9 and EZH2 ^{KO}+ Empty V. Unsorted cells were plated at limiting dilutions after 10 days of lentiviral transduction and allowed to grow for 3 weeks in Wiekmeijer prestim media. The stem cell frequency was then calculated by ELDA software. Mean ± 95% CI of three biological replicates are shown. * $p<0.05$, *** $p<0.001$

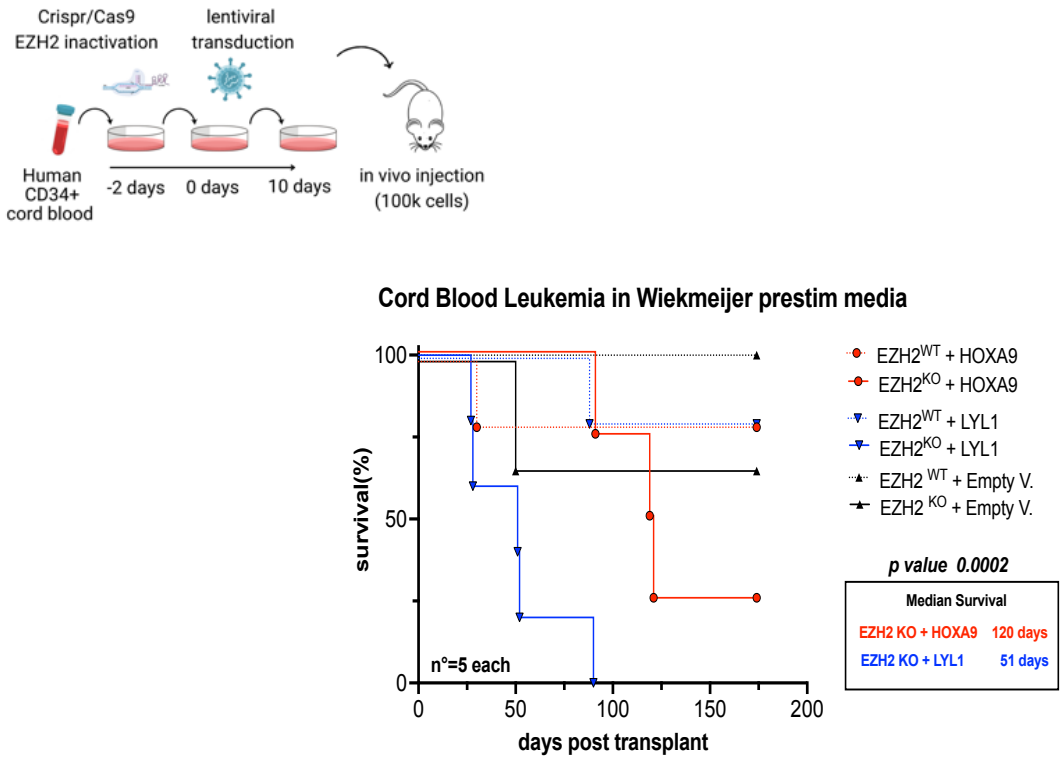


Figure 17 Kaplan-Meier survival curve of EZH2^{WT} or EZH2^{KO} in different context
Leukemia-free survival of mice transplanted with 100,000 EZH2 ^{KO}+ LYL1, EZH2 ^{KO}+ HOXA9, EZH2 ^{KO}+ Empty V., EZH2 ^{WT}+ LYL1, EZH2 ^{WT} + HOXA9 or Empty Vector cells. Median of survival is shown. Statistical analysis was performed by Log-rank (Mantel-Cox) test.

4.7 Immunophenotypes of *in vitro* CB-derived population in different conditions

To precisely track the significant immunophenotypic caused by the inactivation of EZH2 of oncogene-transduced CB cells, a multi parameter flow cytometric analysis was conducted. Specifically, an early T-cell immunopanel including, CD34, CD38, and CD1a markers, which are essential for tracking the initial stages of differentiation in the T-cell lineage, were included. Data collected from three independent experiments illustrate the immunophenotypic levels at three distinct time points (10, 20, and 31 of *in vitro* cell culture).

When compared to control conditions, the cell fraction of EZH2 knockout overexpressing LYL1 cells (marked by mtagBFP2/GFP+) was enriched with early T cell populations, specifically identified as Double Negative 1 (DN1) thymocytes. This suggests that T-cell differentiation can be delayed or blocked under these conditions.

For these experiments, EZH2 knockout overexpressing LYL1 cell fraction (B+/G+) revealed significant enrichment of the early T cell population characterized by CD34+CD38-CD1a- markers) when compared to control groups. The increased number and proportion of DN1 thymocytes were evident starting from day 20 for EZH2^{KO} + LYL1 vs EZH2 ^{WT}+ LYL1 (mean 9.35 ± 1.01 vs 1.68± 0.63), with a significant increment observed at the end of the cell culture period (mean 28.5 ± 2.5 vs 1.55± 0.81) (**Fig.18A**).

For EZH2 knockout overexpressing HOXA9 cell fraction no significant difference was observed on days 10 and 20. Though the effect was less marked, a significant difference was observed on day 31 for EZH2^{KO} + HOXA9 vs EZH2^{WT}+ HOXA9 (mean 14.80 ± 1.89 vs 5.75 ± 1.22) (**Fig.18B**).

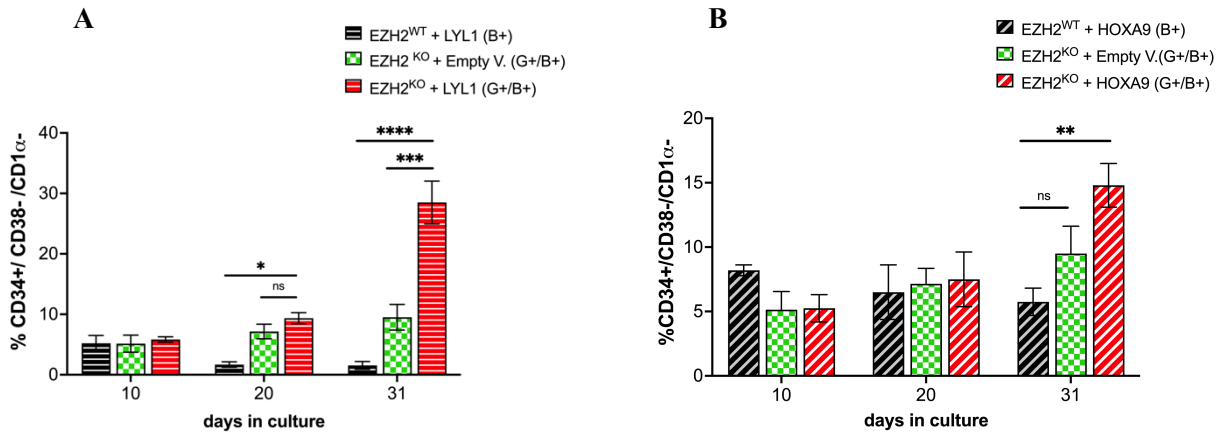


Figure 18 Flow cytometric immunophenotypes of manipulated cells cultured in Stem media
Immunophenotypic assessment of CD34+ CD38-CD1a- early T-cells at the double negative 1 (DN1) development stage in double modified cell populations by flow cytometry.**A**) EZH2^{WT} +LYL1, vs EZH2^{KO} + Empty V. and EZH2^{KO} +LYL1 **B**) EZH2^{WT} +HOXA9 vs EZH2^{KO} + Empty V. and EZH2 +HOXA9. * p value < 0.05. ** p value< 0.01. *** p value <0. 001.ns, non significant. Statistical analysis was performed by Two-ways ANOVA (GraphPad Software).

Then, a multi parameter spectral flow cytometry was used to investigate the immunophenotypic changes for all conditions. The immunophenotype was assessed using a 23-colour panel containing T, myeloid-lineage stem cell markers, as described previously. UMAP analysis and FlowSOM algorithm revealed eight major clusters (**Fig.19A,B**). Specifically, clusters #2, #5, #6, #8, mainly enriched from both $EZH2^{KO}$ + Empty V. and $EZH2^{KO}$ + LYL1 (47,10%, 56.40%, 29.10%, 23.8% and 38.3%, 34.2%, 70.8%, 72.7% respectively), while clusters #3 and #4 were composed of $EZH2^{KO}$ + LYL1 cells (100%) (**Fig.19 C, D**).

Clusters #4, #6, and #8 exhibit the highest expression of CD117 in terms of both proportion and intensity, as also cluster #3, even though it represents only 1.20% of total cells (**Fig. 19E**). These subpopulations were characterized by the highest expression of CD44 and enriched for CD7+CD5- markers. Also, a complete absence of committed T-cell progenitors CD7+CD5+CD1a+ was observed. Moreover, an aberrant expression of CD4 and CD8 was detected in clusters #2 and #5 and clusters #3 and #8, respectively. Nonetheless, these populations still expressed a spectrum of stem and myeloid markers, including CD33, CD15, and CD13 making it difficult to understand the immunophenotype for these clusters.

In the $EZH2^{KO}$ + HOXA9 condition, Clusters #4 and #5 were mainly enriched for edited cells with the highest expression of CD117. Again, the data showed that CD117 expression was associated with the highest levels of CD44 and CD7, indicating an early T cell progenitor commitment (**Fig.20A-E**).

Although this analysis does not definitively establish that EZH2-edited overexpressing LYL1 and HOXA9 cells induce a differentiation delay or block towards the T lineage commitment, the enrichment of T progenitor cells CD7+CD44+CD1a- associated with an aberrant myeloid T/Myeloid mixed phenotypes could recapitulate stages of immature T-ALL as observed in the earliest forms of ETP-ALL leukemia (14).

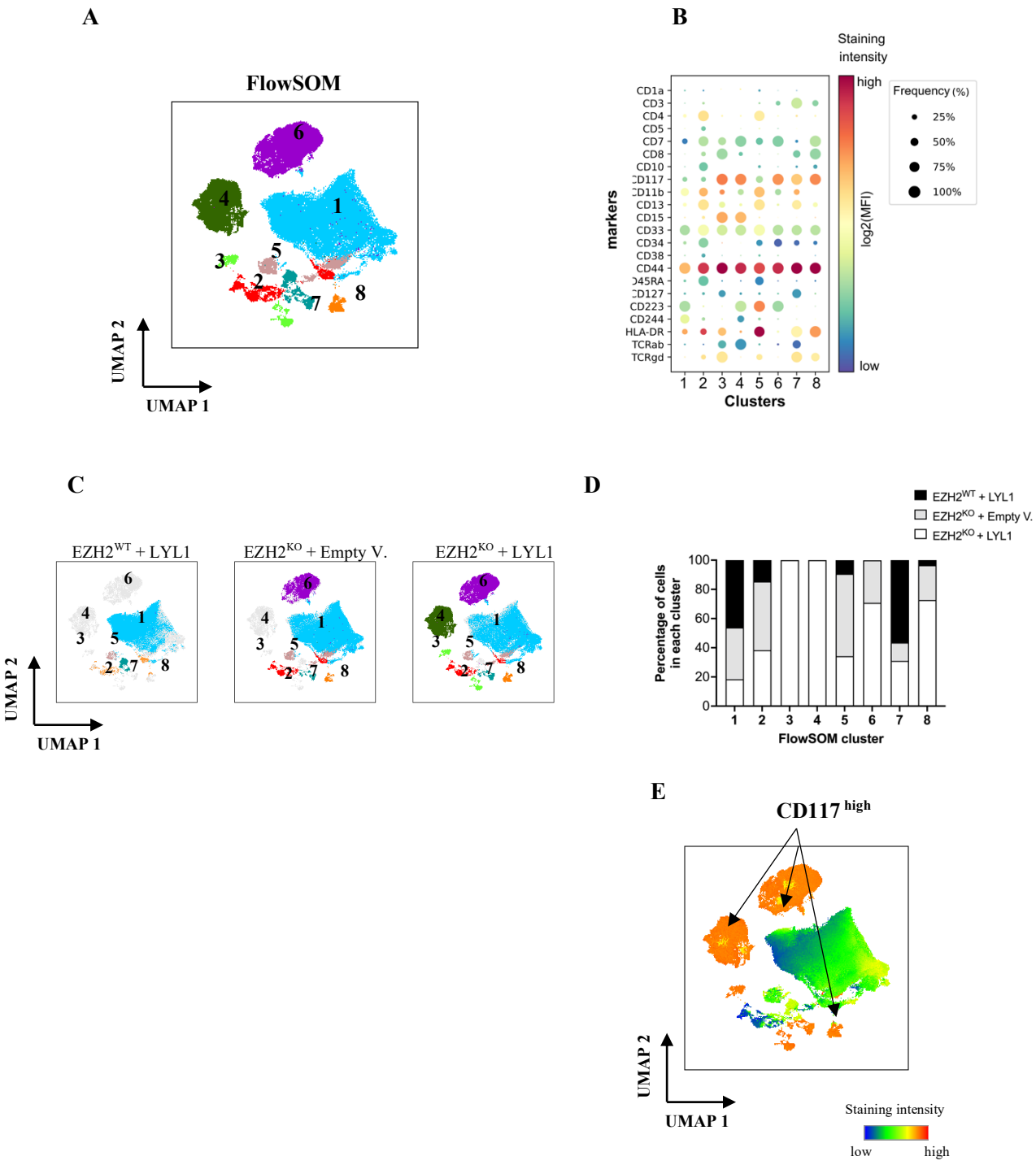


Figure 19 23-parameter UMAP distribution of flow cytometric immunophenotypes of EZH2^{KO} or EZH2^{WT} overexpressing LYL1

A) 23-parameter UMAP distribution of flow cytometric immunophenotypes of manipulated cells cultured in Wiekmeijer prestim media for 31 days. FlowSOM clustering was performed and resulted in 8 phenotypic clusters (1-8). **B**) Dot plot indicating the expression of CD markers for each UMAP cluster. Frequency indicates the percent of cells positive for each CD marker based on manual gating and is displayed by the size of each dot (red = high staining intensity blue = low staining intensity). **C**) Edited and /or transduced HSPCs projected into UMAP space and segregated according to the oncogenes overexpressing. **D**) Bar plot indicating the percent of cells in each FlowSOM cluster. **E**) UMAP heatmap showing the levels of CD117 expression intensity

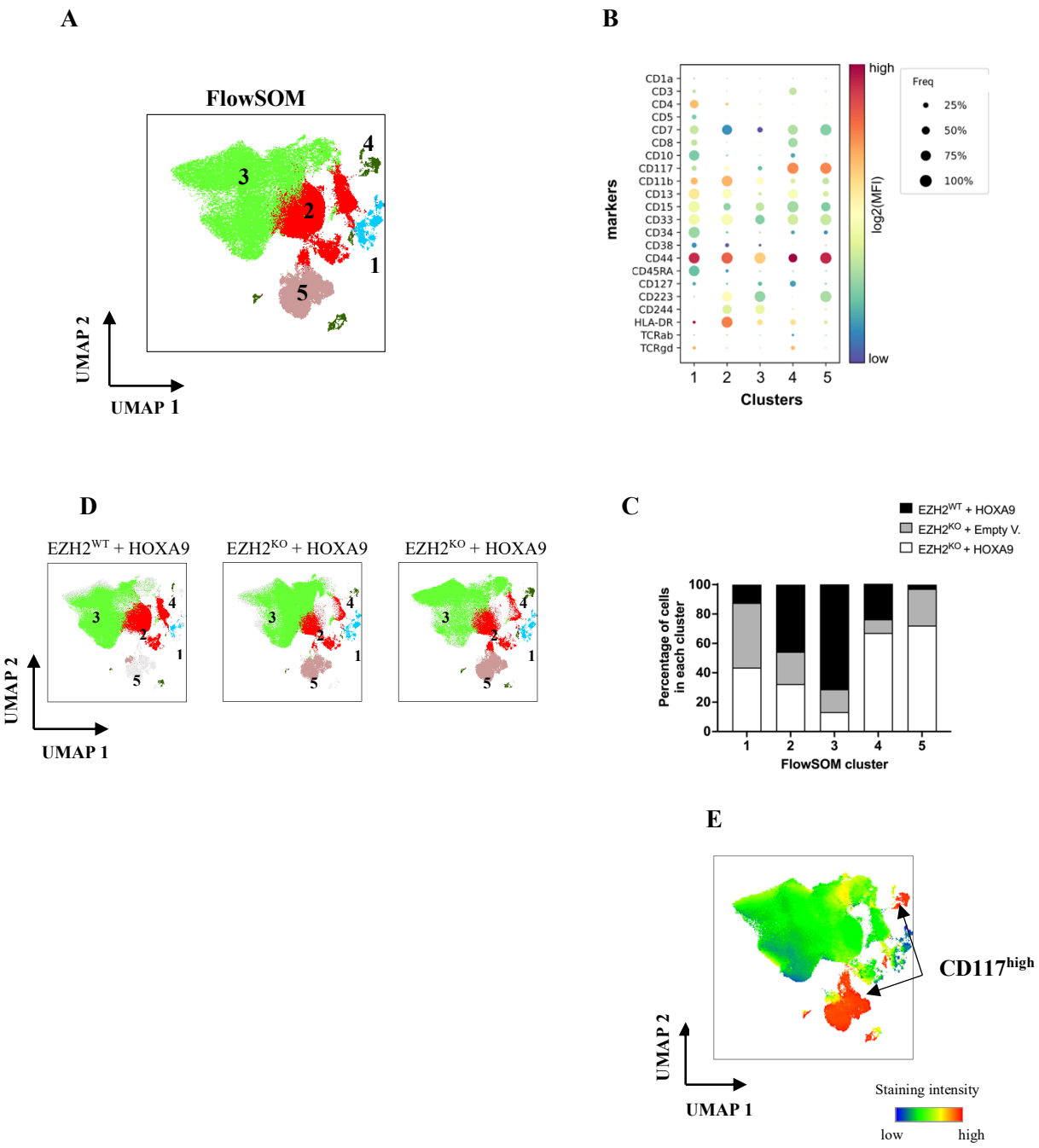


Figure 20 23-parameter UMAP distribution of flow cytometric immunophenotypes of EZH2^{KO} or EZH2^{WT} overexpressing HOXA9

A) 23-parameter UMAP distribution of flow cytometric immunophenotypes of manipulated cells cultured in Wiekmeijer prestim media for 31 days. FlowSOM clustering was performed and resulted in 5 phenotypic clusters (1-5). **B**) Dot plot indicating the expression of CD markers for each UMAP cluster. Frequency indicates the percent of cells positive for each CD marker based on manual gating and is displayed by the size of each dot (red = high staining intensity blue = low staining intensity). **C**) Edited and /or transduced HSPCs projected into UMAP space and segregated according to the oncogenes overexpressing. **D**) Bar plot indicating the percent of cells in each FlowSOM cluster. **E**) UMAP heatmap showing the levels of CD117 expression intensity

It was observed that using StemSpan SFEM II serum-free media with StemSpan T Cell Progenitor Expansion Supplement (TEM media) the positive population of CD117 increased progressively, by day 31 both for the EZH2 knockout overexpressing LYL1 or HOXA9 conditions. Based on the previously developed hypothesis, the presence of the myeloid stem cell marker,CD117, was evaluated under the same conditions but in Wiekmeijer prestim medium. The CD117 positive population in the $EZH2^{KO}$ + LYL1 condition exhibited the brightest expression that peaked on 31 days (50%), compared to the control (2.5%) (**Fig**.**21A,B**). Interestingly, in the $EZH2^{WT}$ + LYL1 condition, CD117 positive population disappears over time.

An effect was also observed in the $EZH2^{KO}$ + HOXA9 only condition, though it was more contained (35%) (**Fig**.**22A,B**). Subsequently, the immunophenotype of CD117 positive cell fraction was observed **(Fig**.**21,22C).** The data showed for both $EZH2^{KO}$ + LYL1 or HOXA9 conditions an enrichment of CD7+, CD44+, CD1a- indicating the presence of early T cell progenitors and the complete absence of the later stages of T-cell lineage commitment.

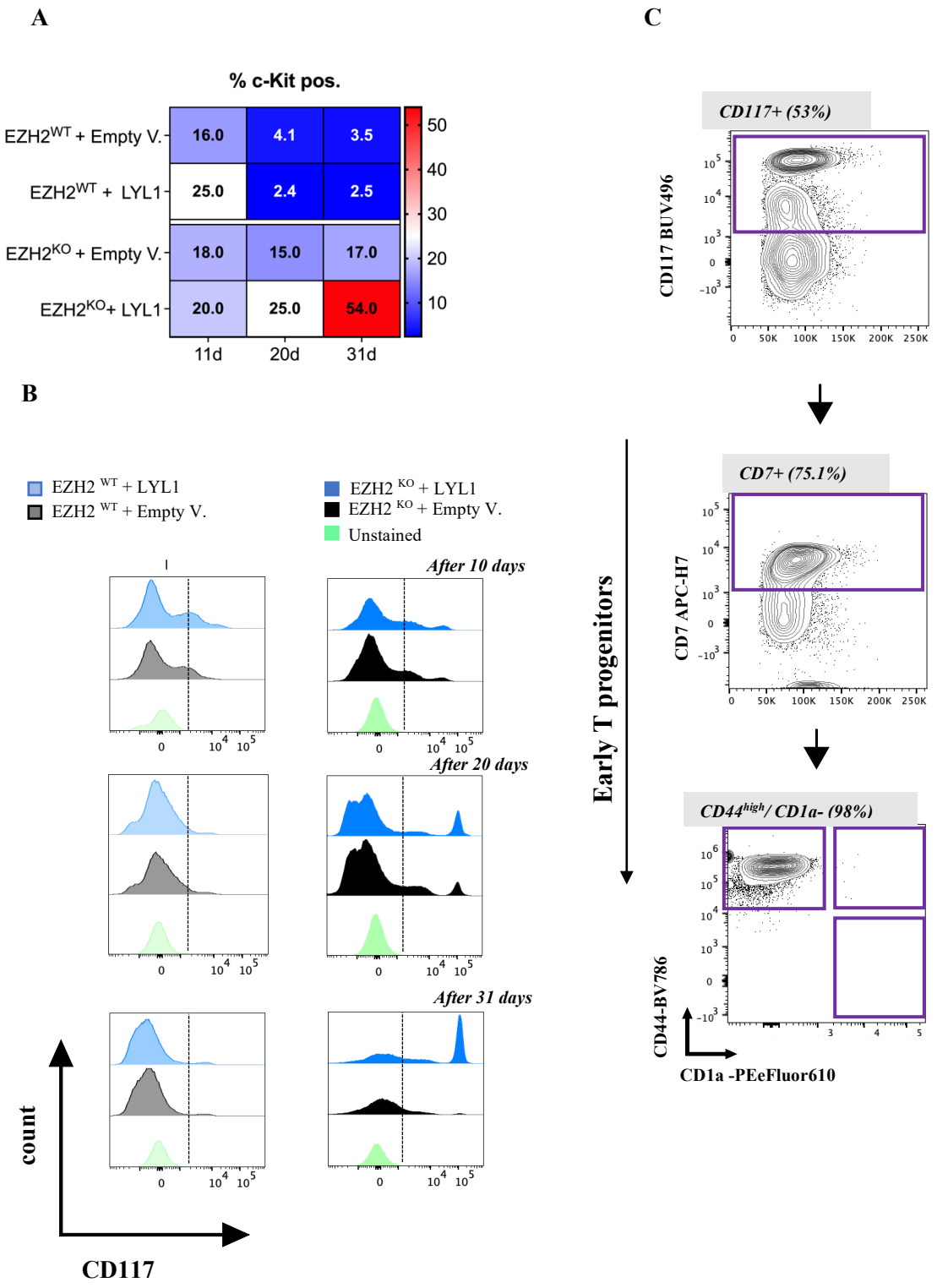


Figure 21 EZH2-edited and/or LYL1 transduced HSPCs are enriched in CD117 within early T cell progenitors

A) Percentage of positive cells for CD117 marker on EZH2-edited and/or LYL1 transduced human CB cells at 10,20- and 31 days post-transduction. **B**) Overlay histograms showing CD117 expression in CD45/mtagBFP2 positive cells. **C**) Contour plots indicating the proportion of early T progenitor cells (CD7+CD44+CD1a) gated on CD117 positive population as indicated. Human CD34+ CB cells were cultured in Wiekmeijer prestim medium.

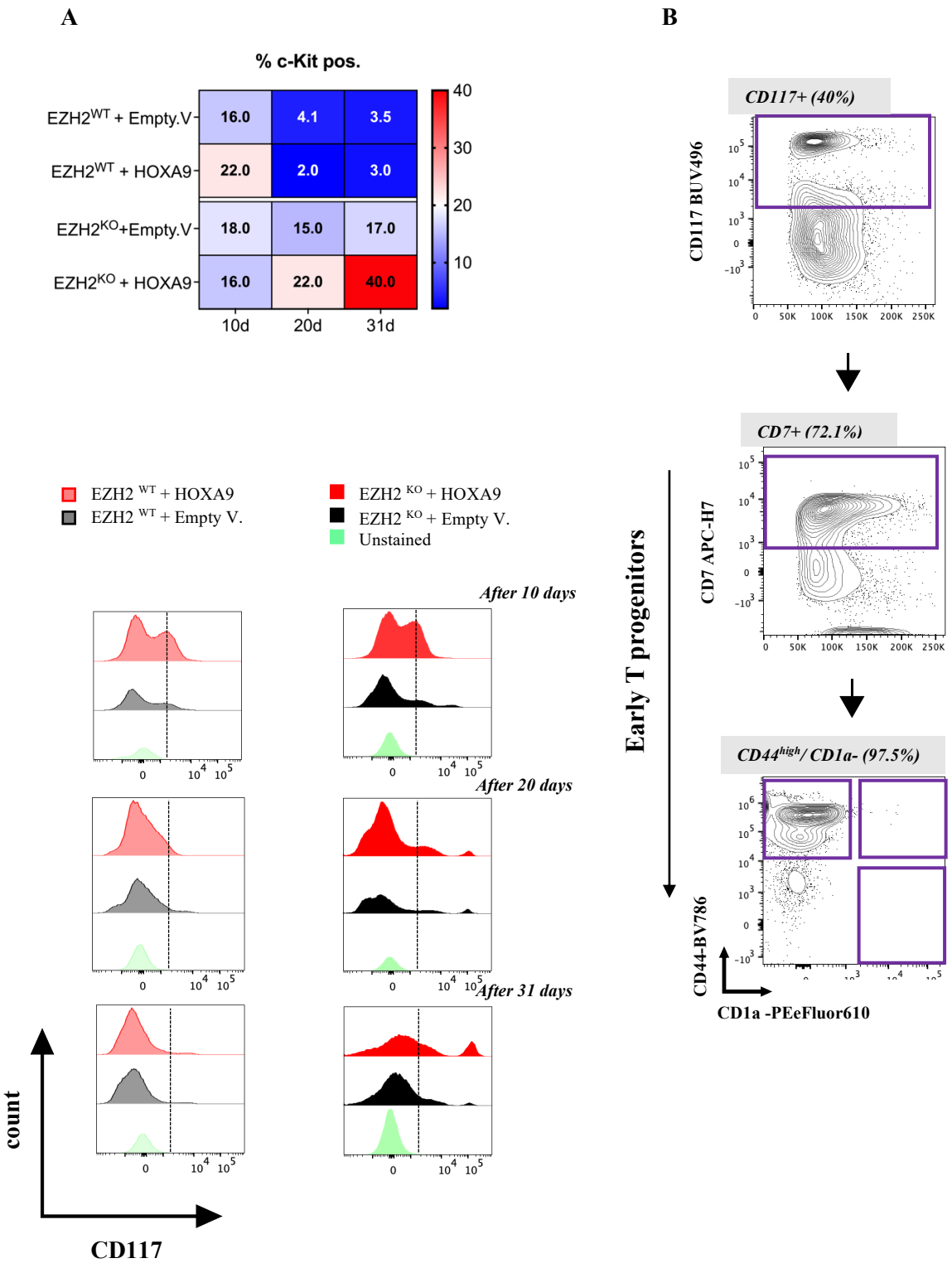


Figure 22 EZH2-edited and /or HOXA9 transduced HSPCs are enriched in CD117 within early T cell progenitors

A) Percentage of positive cells for CD117 marker on EZH2-edited and/or HOXA9 transduced human CB cells at 10,20- and 31 days post-transduction. **B**) Overlay histograms showing CD117expression in CD45/mtagBFP2 positive cells. **C**) Contour plots indicating the proportion of early T progenitor cells (CD7+CD44+CD1a) gated on CD117 positive population as indicated. Human CD34+ CB cells were cultured in Wiekmeijer prestim medium

4.8 EZH2 KO with LYL1 or HOXA9 reveals an early gene signature in T-ALL cell line, Loucy

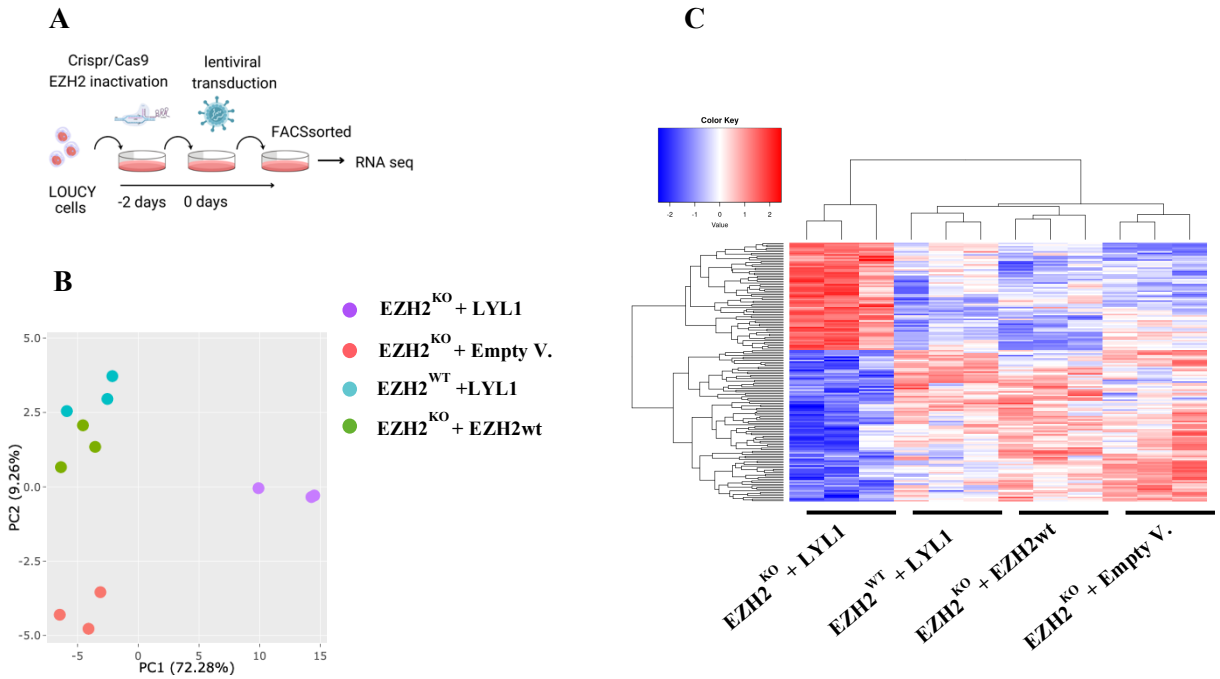
To understand the molecular mechanisms driving the oncogenic potential of EZH2 knockout, in combination with HOXA9 or LYL1, a whole transcription RNA sequencing analysis was conducted on an established T-ALL Loucy cell line (93). It was reported that Loucy cells show an ETP-ALL gene expression signature and are likely derived from an ETP-ALL patient (94). Loss-of-function EZH2 mutations are significantly enriched in the ETP-ALL subgroup which is characterized by a deregulated stem cell-associated transcriptional program (67,71). Also, the EZH2-mutated patients are closely associated with the upregulation of HOXA cluster genes and LYL1 oncogenes (41), making Loucy cell line a perfect model to study the molecular biology of ETP-ALL Leukemia.

RNA sequencing (RNA-Seq) analysis revealed 351 differentially expressed genes (DEGs) in $EZH2^{KO}$ + LYL1 cells, with 219 upregulated and 132 downregulated (fold change ≥1.2 or ≤0.7, p-value < 0.05, Wald's test). Additionally, this condition resulted in a distinct gene expression profile, segregating the RNA-Seq samples into a separate group according to both principal component analysis (PCA) and hierarchical clustering (**Fig.23A-C**). Of interest, molecular pathways associated with cell cycle progression, proliferation, and differentiation of early T cells, including "G2-M Checkpoint", and "KRAS Signaling Up" were identified by the Enrichment analysis tool (https://maayanlab.cloud/Enrichr*)* (**Fig.23D**). Concordant with existing literature NF-κB by TNF-alpha plays a fundamental role in promoting cell survival and proliferation, as well as in orchestrating the inflammatory response within the leukemic microenvironment, making this pathway intriguing for further studies (95,96). Moreover, Gene Set Enrichment Analysis (GSEA) showed genes enriched in the human DN1 stage compared to DN2/3 stages "DN1vsDN2/3_up", suggesting that EZH2 KO overexpressing LYL1 accentuates transcriptional programs characteristic of early T progenitor cells. These results are in line with previous studies, in which EZH2 knockout is involved in the maturation block of thymocytes by regulating the expression of the oncogenic transcription factors and stem cell-like gene signature (14,84).

A significant enrichment in the HRAS and KRAS gene pathways was found in these samples. Specifically, the "Bild_HRAS_Oncogenic signature" and "KRAS.600_UP. V1_UP" gene pathways were observed enriched in EZH2 knockout cells overexpressing LYL1(**Fig**. **23E**). It is known that HRAS and KRAS pathways play significant roles in the pathogenesis and

progression of T-ALL (97). While CD117 is not universally expressed in T-ALL, cases with aberrant CD117 expression or activity have been identified (73). In this scenario, the potential activation of HRAS and KRAS, key components of the RAS/MAPK signaling pathway, by CD117 activation could enhance downstream signaling cascades, thereby promoting leukemic transformation and progression as previously reported (97).

In the EZH2 knockout overexpressing HOXA9 condition, cells clustered in a distinct group compared to control cells, as demonstrated by both principal component analysis (PCA) and hierarchical clustering. (**Fig.24A, B**). Differentially expressed genes (DEGs) in $EZH2^{KO}$ + HOXA9 revealed 455 upregulated and 189 downregulated (fold change ≥1.2 or ≤0.7, p-value < 0.05, Wald's test). This cluster was enriched for the MSIgDB HallMark gene set "G2-M Checkpoint", suggesting that the EZH2 KO overexpressing HOXA9 cells not only influences but also significantly impacts cell cycle regulation, particularly at the G2-M transition **(Fig.24C)**. Moreover, this analysis showed that the inflammatory stimulation TGF beta signaling pathway was enriched in this condition, compared to other clusters. Of note, Gene Set Enrichment Analysis (GSEA) from DEG genes revealed enrichment for "UP_IN_HUMAN_ETP-ALL", "DN1VSDN2/DN3_UP" and "HSC_HUMAN_UP" gene signatures indicating that leukemic cells acquire early T-ALL phenotypes and stem cell-like properties, contributing to their proliferative and survival advantages. **(Fig.24D)**.



D

MSigDB HallMark 2020	P-value	adjusted P-value
G2-M Checkpoint	6.4E-03	5E -02
TNF-alpha Signaling via NF-kB	3.0E-04	5.0E-02
KRAS Signaling Up	2.2E-02	1.2E-01

E

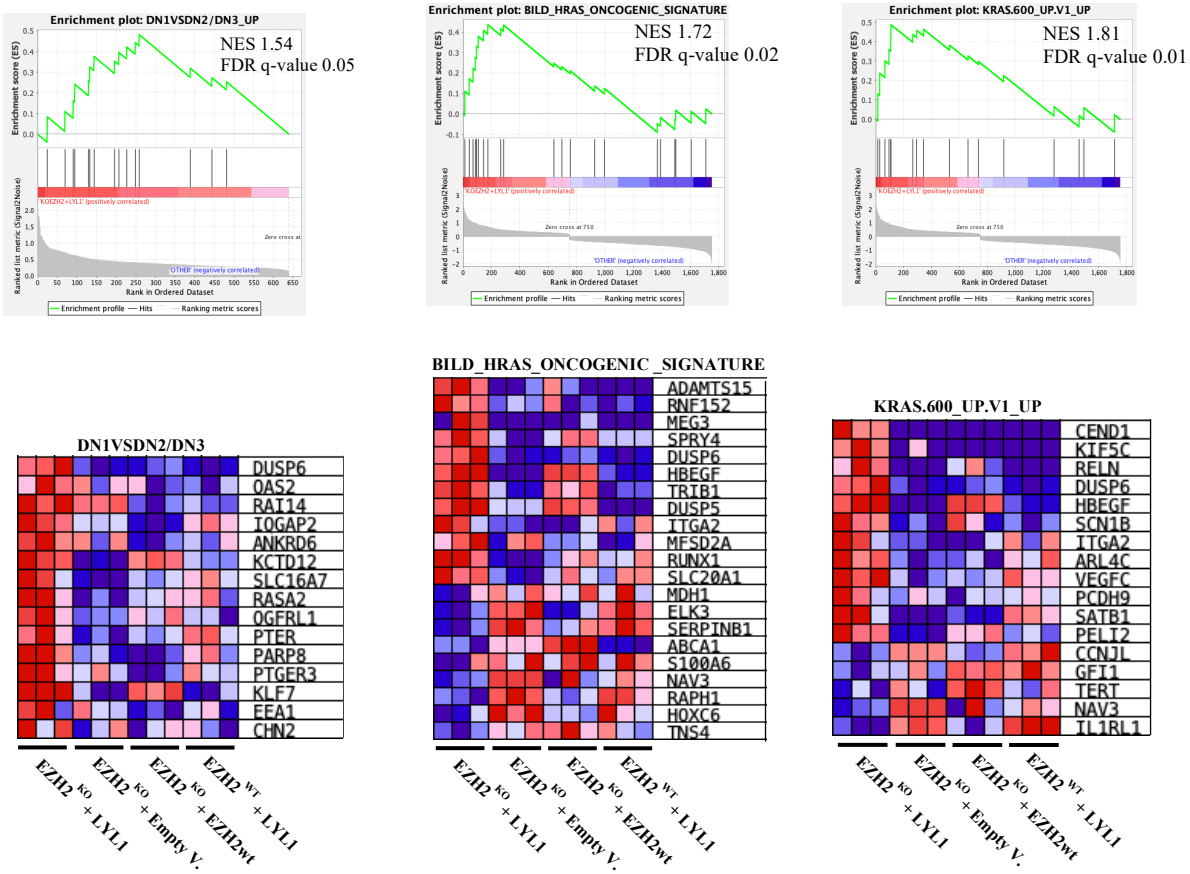
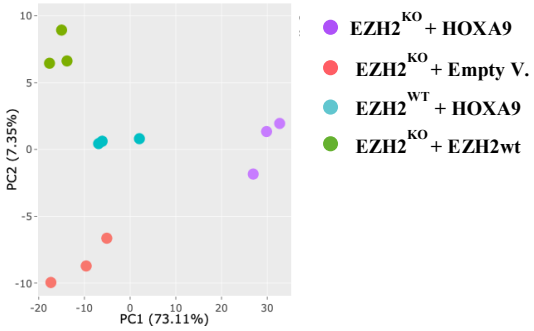
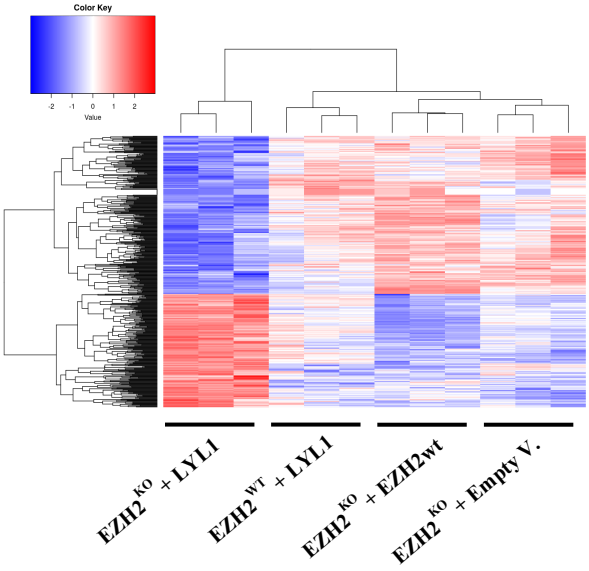


Figure 23 Transcriptomic analysis of EZH2 KO overexpressing LYL1 in Loucy, T-ALL cell line
General Experimental design used for generation of manipulated Loucy cells. Briefly, cells were edited via CRISPR/CAS9 for EZH2 knockout/GFP knock-in expression, transduced with lentiviral constructs, FACS sorted, and for performing RNA extraction and RNA Sequencing (RNA-Seq). **B**) Principal component analysis (PCA) of differentially expressed genes (DEGs) from RNA-seq data by DEbrowser tool. EZH2 knockout/GFP knock-in Loucy cells, were transduced with lentiviruses for expression of Empty V., LYL1, $EZH2^{WT}$ (rescue condition) or $EZH2^{WT}$ + LYL1 as indicated in A.**C**) Heatmap and hierarchical clustering of differentially expressed genes (DEGs) (p-value ≤ 0.05 and LogFC ≥1.2 or ≤ 0.7 by DEbrowser tool. Gene expression values were normalized by a regular logarithmic transformation (MRN) a scaled by row. **D**) Table of the hallmark gene sets significantly enriched in $EZH2^{KO}$ +LYL1 cells, compared to other cell conditions by Enrichr online tool. **E**) GSEA Enrichment plot (score curves) in $EZH2^{KO}$ + LYL1 vs $EZH2^{KO}$ +Empty V., $EZH2^{KO}$ + $EZH2^{WT}$ and $EZH2^{WT}$ + LYL1 clusters. Signatures were obtained from the Molecular Signatures Database (MSigDB) (method section 3.10.4). GSEA was then run with 1000 permutations. NES normalized enrichment score, FDR false discovery rate

A



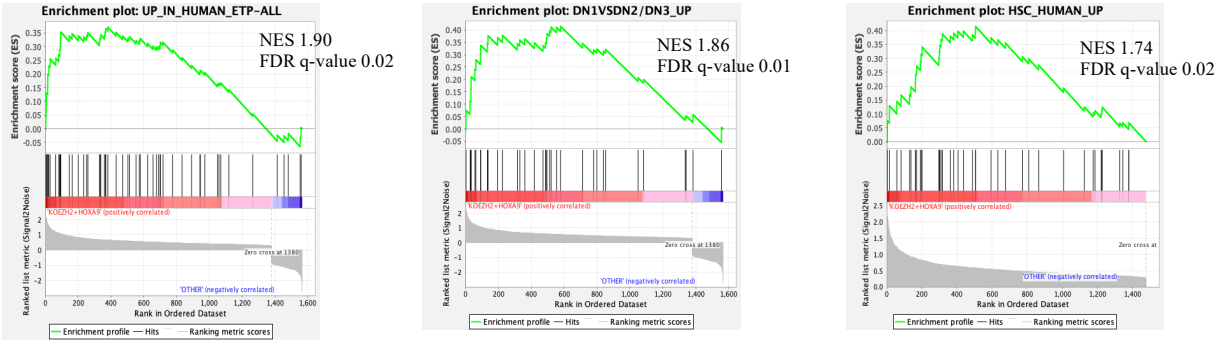
B



C

MSigDB HallMark 2020	P-value	Adjusted P-value
G2-M Checkpoint	8.534E-07	4.1E-05
TGF beta signaling pathway	3.4E-03	2.11E-01

D



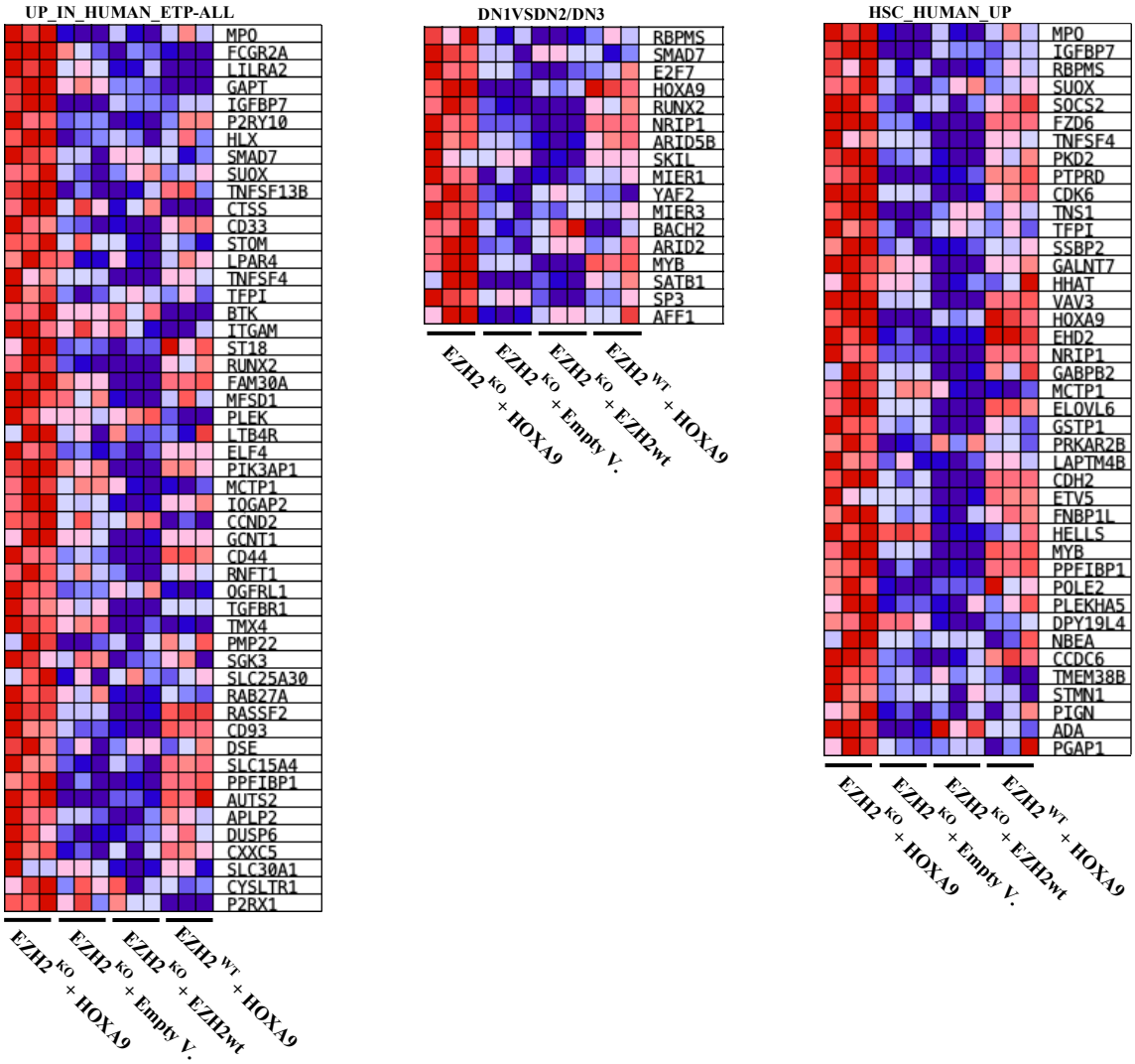


Figure 24 Transcriptomic analysis of EZH2 KO overexpressing HOXA9 in Loucy, T-ALL cell line
A) Principal component analysis (PCA) of differentially expressed genes (DEGs) from RNA-seq data by DEbrowser tool. EZH2 knockout/GFP knock-in Loucy cells were transduced with lentiviruses to express Empty V., HOXA9, $EZH2^{WT}$ (rescue condition) or $EZH2^{WT}$ + HOXA9 as indicated in A. **B**) Heatmap and hierarchical clustering of differentially expressed genes (DEGs) (p-value ≤ 0.05 and LogFC ≥1.2 or ≤ 0.7 by DEbrowser tool). Gene expression values were normalized by a regular logarithmic transformation (MRN) and scaled by row. **C**)Table of the hallmark gene sets significantly enriched in $EZH2^{KO}$ +HOXA9 cells, compared to other cell conditions by Enrich online tool. **D**) GSEA Enrichment plot (score curves) in $EZH2^{KO}$ +HOXA9 vs $EZH2^{KO}$ +Empty V., $EZH2^{KO}$ + $EZH2^{WT}$ and $EZH2^{WT}$ +HOXA9 clusters. Signatures were obtained from the Molecular Signatures Database (MSigDB) (method section 3.10.4). GSEA was then run with 1000 permutations. NES normalized enrichment score, FDR false discovery rate.

4.9 EZH2 LoF mutations are significantly associated with stem-cell gene signature of early T-ALL cases

Previous works have reported that pathogenic DNA alterations in the EZH2 gene are frequent in T-ALL patients (41,71). In particular, loss-of-function EZH2 mutations are notably frequent in pediatric cases and significantly enriched in the T-ALL subgroup at the early stage of T cell development (14). Notably, EZH2 inactivation has been associated with poor prognosis in the aggressive ETP-ALL, a subgroup characterized by aberrant expression of myeloid and stem cell markers (67).

To further investigate differences between EZH2-mutated and PRC2wt T-ALLs, a publicly available patient dataset that contains both clinical and sequencing data was used (71). In this cohort, the prevalence of PRC2 alterations was 9% (24 altered patients out of a total of 240 PRC2 wild-type cases). Among these alterations, EZH2, a core member of the PRC2 complex, was found to be altered in the majority of cases 58%, while genetic anomalies in EED and SUZ12 were less frequent, 33% and 8%, respectively. Of note, cases carrying alterations in more than one PRC2 member were relatively rare. The most predominant alterations of the EZH2 gene were point mutations (missense, 80%) and out-of-frame mutations (INDEL, 20%). Genetic alterations of EED and SUZ12 were found to be mutually exclusive (**Fig. 25A**).

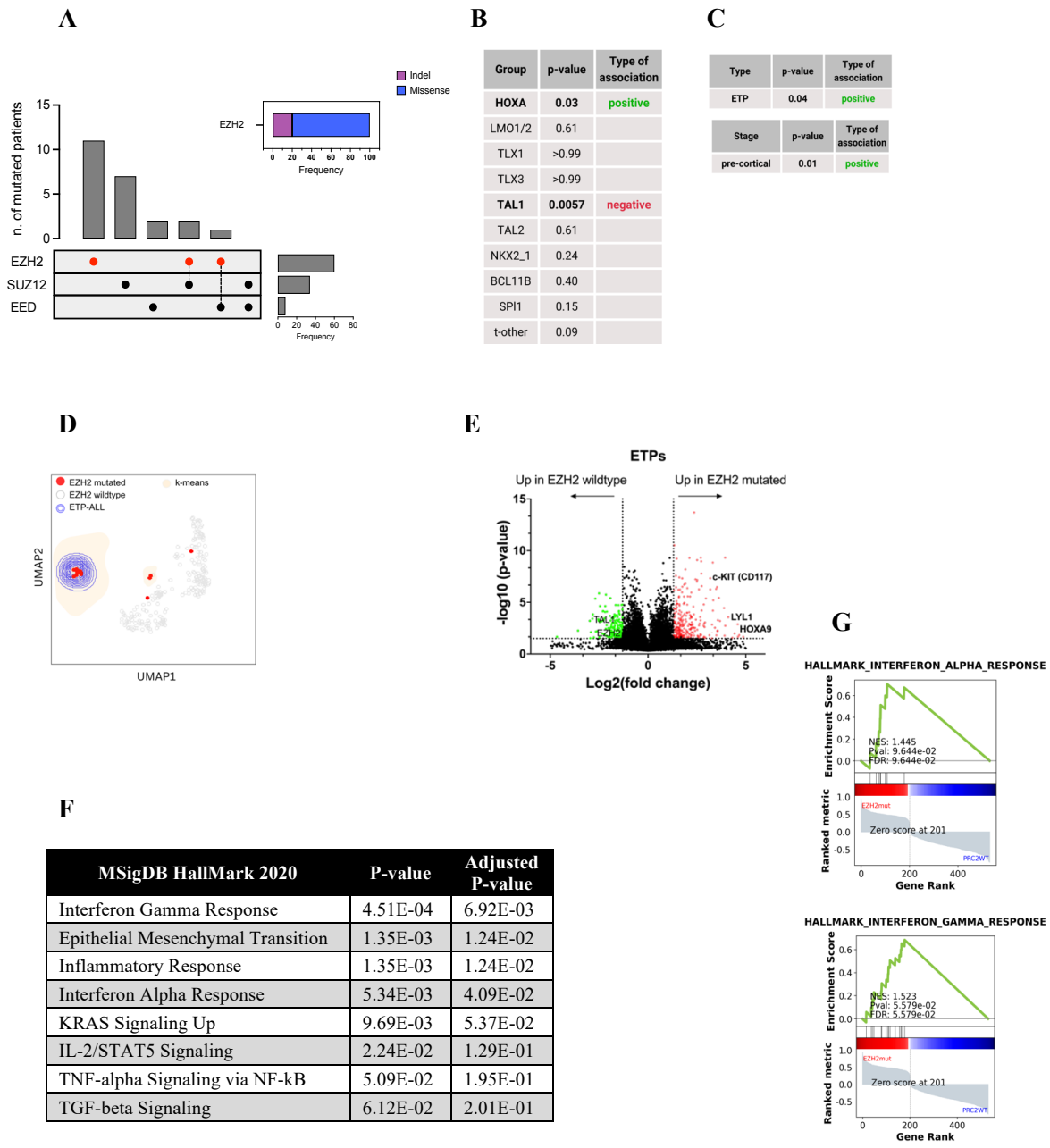
Given the key role of the EZH2 in the PRC2 complex, statistical analysis across the different T-ALL subtypes was conducted. Through RNA sequencing (RNA-seq) data, it was observed that EZH2-mutated patients were statically enriched in the HOXA subgroup, known for its immature immunophenotype in T-ALL. Interestingly, these cases were frequently observed in a stem/immature immunophenotypic stage such as ETP-ALL, and in the genetic subtype of immature (pre-cortical stage) leukemias. Conversely, EZH2 loss of function was negatively associated with the late-cortical stage of thymocyte differentiation overexpressing TAL1 gene function (**Fig. 25B-C**).

Then, differential gene expression analysis between two of these groups was performed. 1001 DEGs (264 up in EZH2 mutated, 155 up in PRC2wt, log2FC$\geq$1.2 or $\leq$-1.2, p-value <0.05) were identified. Based on UMAP visualization and k-means clustering analysis EZH2-mutated patients clustered together with ETPs, demonstrating transcriptional similarities (**Fig. 25D**). Of note, the EZH2-mutated patients shared a gene expression signature characterized by elevated

expression levels of HOXA9 and LYL1 and were statistically associated with stem cell markers of early T-ALL, such as CD117 receptor (**Fig. 25E**).

Pathway analysis of the differentially expressed genes (DEG) showed that EZH2 mutated patients were characterized by enrichment for pathways related to "K-RAS signaling up", which is somehow linked to the activation of CD117 (27). Of note, inflammation pathways, such as "Interferon Gamma Response", "Inflammatory Response", "Interferon Alpha Response" ,TNF-"alpha Signaling via N F-kB" and "TGF-beta Signaling" were found also enriched by enrichR analysis (**Fig. 25F**). Also, Gene Set Enrichment Analysis (GSEA) from DEG genes revealed enrichment for Interferon "Alpha Response" and "Interferon Gamma Response" gene signature (**Fig. 25G).** In line with previous studies chronic inflammatory stimulation can delay T-cell differentiation and alter the growth and differentiation trajectory of transformed T cells, which favors selection of more immature T-ALL subtypes (95,96). This observation would require further scientific investigation.

Additionally, unlike what was observed in T-ALL Loucy cell line, this preliminary analysis did not reveal a positive correlation with the same stemness-related gene signature. The observed divergence could be due to a broad range of genetic states that might affect the comparison between the primary samples and the established T-ALL line. Another limitation is that a wide variety of genetic alterations in T-ALL primary samples can contribute significantly to the complexity of the disease. Indeed, genetic alterations in T-ALL can include chromosomal abnormalities, such as translocations, deletions, and gene fusions as well as mutations in genes that regulate cell cycle progression, apoptosis, T-cell development, and signaling pathways (67,70). This divergence emphasizes the need for a cautious approach in extrapolating findings from cell line models to the patient setting. Future research should focus on expanding the genetic and epigenetic characterization of both primary T-ALL samples and cell lines.



F

MSigDB HallMark 2020	P-value	Adjusted P-value
Interferon Gamma Response	4.51E-04	6.92E-03
Epithelial Mesenchymal Transition	1.35E-03	1.24E-02
Inflammatory Response	1.35E-03	1.24E-02
Interferon Alpha Response	5.34E-03	4.09E-02
KRAS Signaling Up	9.69E-03	5.37E-02
IL-2/STAT5 Signaling	2.24E-02	1.29E-01
TNF-alpha Signaling via NF-kB	5.09E-02	1.95E-01
TGF-beta Signaling	6.12E-02	2.01E-01

Figure 25 EZH2-mutated leukemias are enriched in the HOXA-overexpressing subgroup of early T-ALL

A) The upset plot indicates the number of events and the frequency observed for EZH2, SUZ12 and EED core component of PRC2 complex in T-ALL patients from the T-ALL TARGET study (study; dbGaP phs00018/000464). 264 of those cases have been reported in Liu et al (2017). **B**) Fisher's exact test analysis between the indicated groups of EZH2 mutated vs. wild type leukemias to the different molecular/genetic T-ALL subtypes, **C)** ETP-ALL or stages of T cell development (Liu dataset).**D**) UMAP visualization of the transcriptomes from 254 patient T-ALLs by SeqGeq software (FlowJo). The top 1,000 most variable genes were selected by Welch's t-test (p-value<0.05). The red and blue circles represent EZH2-mutated and ETP-ALL patients, respectively. **E**) Volcano plot depicts the differentially expressed genes between EZH2-mutated vs. wild-type ETP-ALLs. A few select genes that passed the significance threshold (FC > 1.3 or < 0.7 and adjusted p-value <0.05) are highlighted. **F**) Table of the hallmark gene sets significantly enriched in EZH2 mutated vs PRC2wt by Enrich online tool**. G)** GSEA Enrichment plot (score curves) in EZH2-mutated vs. wild-type ETP-ALL

5 Discussion and future prospective

The role of Enhancer of Zeste Homolog 2 (EZH2) in the origin of T-cell acute lymphoblastic leukemia (T-ALL) is a concept of significant interest due to its critical function within the Polycomb Repressive Complex 2 (PRC2) in regulating gene expression through epigenetic mechanisms. Specifically, EZH2 catalyzes the trimethylation of histone H3 lysine 27 (H3K27me3), leading to the repression of target gene expression regulating differentiation processes in normal development (3,12).

A premise of this project is that EZH2 has a tumor suppressor function in T-cell leukemia leading to transcriptional repression of stem/early progenitor-associated programs driving malignant transformation and disease progression (3,8,15). Based on this observation, it was hypothesized that EZH2 inactivation could synergistically regulate the gene signatures driven by the stem/cell transcription factors to promote the leukemogenesis of human progenitor T-cells and the maintenance and propagation of human T-ALLs.

Consistent with this idea, we intended to define the molecular proprieties of EZH2 silencing in early T cell leukemias. For this reason, we transduced human HSPCs from cord blood (CB) with lentivectors encoding known T-ALL transcription factors alone or in combination with the CRISPR/Cas9-mediated inactivation of EZH2 gene and tracked in vitro/vivo phenotypes of cell differentiation and leukemogenesis.

By employing a human model of T-ALL from CD34+ Cord Blood cells (9), we found that EZH2 inactivation via CRISPR/Cas9 synergically cooperates with the overexpression of stem/early progenitor-associated transcription factors such as LYL1 and HOXA9 promoting the leukemogenesis of human progenitor T-cells and the propagation of penetrance leukemias *in vivo* experiments. The data collected from *in vitro* conditions, positively correlated with the maintenance of stem/progenitor features over time. This effect was found more pronounced in Wiekmeijer prestim media which is specifically formulated to support the expansion of hematopoietic stem/progenitor cells (HSPCs). These results fit a model in which LoF mutations in EZH2 lead to inadequate activation of differentiation genes, resulting in a block of differentiation at early stages of T-cell development (1,8).

Previous works showed that the block of EZH2 expression promotes the transcriptional activation of stem cell-associated genes and causes the expansion of early T-cell phenotypes during the malignant transformation of mouse HSCPs (84). Of note, in mouse model, homozygous inactivation of EZH2 leads to T-cell leukemia (68), and promotes the reactivation of stem/early progenitor-associated transcription factors such as HOXA-cluster genes, cooperating with NRAS (8), RUNX, and interferon-gamma (IFNγ) response to accelerate leukemia onset (84).

On the other hand, the inactivation of EZH2 in human T-ALL cell line Jurkat led to marked changes in gene-expression pattern, with downregulation of expression of T-cell differentiation-associated genes and concomitant upregulation of ETP signature genes such as LYL1, HHEX, KIT, and MYCN (40). Additionally, in human T-ALL samples, HOXA9 and LYL1 genes were found to be repressed by PRC2 activation as well as induced by the loss of PRC2 activity (41).

Consistent with these data, bioinformatic analysis of RNAseq data from a large pediatric cohort consisting of 264 diagnostic T-ALL samples (71), has highlighted a significant association between EZH2 mutations and specific leukemia subtypes. Particularly, EZH2 mutated patients are statistically enriched in immature HOXA leukemias and in a subgroup of T-ALL, at the early stage of T cell development (ETP-ALL), which is characterized by aberrant expression of myeloid and stem cell markers and/or stem cell transcription factors such as LYL1.

Conversely, EZH2 mutations have also been identified in mature T ALL (14). Although in our model, the overexpression of TAL1 (generally overexpressed in mature cortical T-ALL) exhibited a significant detrimental impact on cell fitness over time, previous works support the hypothesis that this oncogenic transcription factor can drive immature leukemias when associated with high overexpression of stem cell-associated transcriptional program or permissive epigenetic state occurring in a specific stage of T cell development (8).

Our in vitro cord blood (CB) trial demonstrated that knocking out EZH2, in combination with the overexpression of NOTCH1ΔE, resulted in a reduced capacity for expansion and selection over time, compared to the overexpression of stem/early transcription factors. This outcome does not align with the canonical model where NOTCH1 binding leads to the loss of H3K27me3 and the eviction of PRC2 (98). However, recent studies highlight the low incidence of activating mutations in NOTCH1 in early ETP-ALL, suggesting a distinct oncogenic

mechanism in this leukemia subtype, making the leukemic cells less dependent on NOTCH1 signaling for their growth and survival (99).

Based on these observations, we speculate that EZH2 loss-of-function (LoF) mutations, occurring in T-cell lineage cells at a more advanced stage of thymic development, could result in epigenetic deregulation, thereby inducing the transformation and progression of T-ALL. Such anomalies may lead to the inappropriate expression of markers associated with HSPCs, which are typical of less differentiated subtypes.

Although the exact mechanism of cellular reprogramming that occurs with EZH2 loss will require further study, in this project we observed significant phenotypic differences that are consistent with more aggressive and immature ETP-ALL. Specifically, flow cytometry analysis of LYL1 or HOXA9-transduced HSPCs, in an EZH2 knockout context, exhibited a more pronounced enrichment of early T cell population (CD34+CD38-CD1a-) specifically classified as Double Negative 1 (DN1) thymocytes when compared to control.

This raises the question of whether leukemia transformation and progression are due to transcription factor dysregulation or serve as a secondary consequence of developmental arrest mediated by epigenetic changes.

Moreover, a deeper analysis revealed that EZH2 knockout, when combined with LYL1 or HOXA9, led to the high-level expression of CD117 cell receptor, along with aberrant expression of cell markers characteristic of less-differentiated phenotypes. Surprisingly, the CD117 positive population was characterized by an accumulation of early T-cell progenitors CD7+CD5- or myeloid-skewed progenitors CD33+CD7+, and an almost complete absence of committed T-cell progenitors CD7+CD5+CD1a+. These characteristics are similar to ETP-ALL and Mixed-Phenotype acute leukemias which suggest that these leukemias are at the Myeloid to T-cell crossroad and could have higher accessibility to myeloid genes and signaling pathways (14,67)

Of note, the interpretation of a broad spectrum of phenotypic landscapes observed in our analysis may be attributed to the use of CD34+ Cord Blood cells. CD34+ cells are known to be a heterogeneous population (1). The generation of distinct subpopulations from the same oncogene might be a result of differences in the cell of origin and their permissivity to transformation. In addition, this model preserves a great level of heterogeneity and polyclonality during the in vitro stages of leukemic transformation suggesting that another

reason for this variability could be related to microenvironmental factors and clonal selection. This implies that EZH2-edited and/or transduced LYL1 or HOXA9 HSPCs may show unique lineage-fate propensities during cell culture that may alter the direct comparison between the different cell conditions. In this context an innovative strategy might involve identifying a population of progenitor cells that mimic the characteristics of HSPC cells migrating into the thymus, thereby triggering T-cell ontogeny. This approach could precisely replicate the processes of T-lineage commitment.

ETP-ALL LOUCY cell line was also genetically manipulated to understand the molecular mechanisms of leukemia maintenance through RNA and ChIP sequencing experiments. Although the RNA sequencing data have already been collected and the analyses completed, our team is currently engaged in experiments with ChIP sequencing. Specifically, Transcription Factor ChIP-Seq data for both HOXA9 and LYL1 along with Histone marks (H3K27me3 and H3K27ac) are still in progress.

However, Gene Set Enrichment Analysis (GSEA) showed enrichment for terms such as "Bild_HRAS_Oncogenic signature" and "KRAS.600_UP", as well as "DN1vsDN2/3_up" gene signatures for EZH2 KO overexpressing LYL1 cells. Conversely, "UP_IN_HUMAN_ETP-ALL", "DN1vsDN2/3_up" and "HSC_Human_up" terms were enriched for EZH2 KO overexpressing HOXA9 cells when compared to control. This suggests that overexpression of both LYL1 or HOXA9 oncogenes may accentuate transcriptional programs characteristic of early T progenitor cells, aggressive early T-ALL (ETP-ALL), or immature immunophenotypes with stem-like features. However, more experiments are necessary to further test this hypothesis.

By analyzing a publicly available T-ALL dataset, we found a striking enrichment of several inflammatory signatures in EZH2 mutated T-ALL patients. Specifically, early T-cell leukemias showed enrichment of Gene Ontology (GO) terms associated with inflammatory signaling pathways. Moreover, Gene Set Enrichment Analysis (GSEA) revealed enrichment for the terms "Hallmark_interferon_apha_response" and "Hallmark_interferon_gamma_response", suggesting that chronic inflammatory stimulation may alter the growth and differentiation trajectory of transformed cells, which ultimately favors the selection of more immature T-ALL subtypes.

Consistent with this view, pathway analysis of the differentially expressed genes by gProfiler showed enrichment for “TNF-alpha Signaling via NF-kB” and “TGF beta signaling” inflammatory pathways in EZH2 KO overexpressing LYL1 or HOXA9 Loucy cells, respectively. It is known that inflammation plays a crucial role in the immune system, and dysregulation of inflammatory signaling pathways is a significant risk factor for multiple diseases, including T-ALLs(69). I hypothesize inflammaging may skew T-cell differentiation towards more immature and aggressive phenotypes of T-ALLs, especially in the context of EZH2 loss with HOXA9 or LYL1 overexpression.

Thus, exploring the impact of inflammaging on immature and EZH2-dependent T-cell leukemias could reveal specific strengths or dependencies of different T-ALL genetic subtypes. Ultimately, this strategy could prove useful in identifying novel therapeutic targets against inflammaging-related dependencies in early T-ALL patients.

Looking forward, the integration of RNA-Seq and ChIP-Seq data will provide a comprehensive view of the transcriptional and epigenetic mechanisms that could be involved in T-ALL development. This approach leads to identifying specific genome regions and their roles in gene regulation, with the aim of uncovering new complex regulatory networks and/or potential therapeutic targets.

6 Bibliography

1. Doulatov S, Notta F, Laurenti E, Dick JE. Hematopoiesis: A Human Perspective. Cell Stem Cell. 2012 Feb;10(2):120–36.

2. Becker AJ, McCULLOCH EA, Till JE. Cytological Demonstration of the Clonal Nature of Spleen Colonies Derived from Transplanted Mouse Marrow Cells. Nature. 1963 Feb;197(4866):452–4.

3. Notta F, Doulatov S, Laurenti E, Poeppl A, Jurisica I, Dick JE. Isolation of Single Human Hematopoietic Stem Cells Capable of Long-Term Multilineage Engraftment. Science. 2011 Jul 8;333(6039):218–21.

4. Rodrigues CP, Shvedunova M, Akhtar A. Epigenetic Regulators as the Gatekeepers of Hematopoiesis. Trends Genet. 2021 Feb;37(2):125–42.

5. Weissman IL. Translating Stem and Progenitor Cell Biology to the Clinic: Barriers and Opportunities. Science. 2000 Feb 25;287(5457):1442–6.

6. Petrie HT, Zúñiga-Pflücker JC. Zoned Out: Functional Mapping of Stromal Signaling Microenvironments in the Thymus. Annu Rev Immunol. 2007 Apr 1;25(1):649–79.

7. Rossi FMV, Corbel SY, Merzaban JS, Carlow DA, Gossens K, Duenas J, et al. Recruitment of adult thymic progenitors is regulated by P-selectin and its ligand PSGL-1. Nat Immunol. 2005 Jun;6(6):626–34.

8. Hosokawa H, Rothenberg EV. How transcription factors drive choice of the T cell fate. Nat Rev Immunol. 2021 Mar;21(3):162–76.

9. Dik WA, Pike-Overzet K, Weerkamp F, De Ridder D, De Haas EFE, Baert MRM, et al. New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling. J Exp Med. 2005 Jun 6;201(11):1715–23.

10. Ho L, Crabtree GR. Chromatin remodelling during development. Nature. 2010 Jan;463(7280):474–84.

11. Ducasse M, Brown MA. Epigenetic aberrations and cancer. Mol Cancer. 2006;5(1):60.

12. Kanwal R, Gupta S. Epigenetic modifications in cancer. Clin Genet. 2012 Apr;81(4):303–11.

13. Tollervey JR, Lunyak VV. Epigenetics: Judge, jury and executioner of stem cell fate. Epigenetics. 2012 Aug 18;7(8):823–40.

14. Brady SW, Roberts KG, Gu Z, Shi L, Pounds S, Pei D, et al. The genomic landscape of pediatric acute lymphoblastic leukemia. Nat Genet. 2022 Sep;54(9):1376–89.

15. Lawrence M, Daujat S, Schneider R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends Genet. 2016 Jan;32(1):42–56.

16. Du J, Johnson LM, Jacobsen SE, Patel DJ. DNA methylation pathways and their crosstalk with histone methylation. Nat Rev Mol Cell Biol. 2015 Sep;16(9):519–32.

17. Shen H, Laird PW. Interplay between the Cancer Genome and Epigenome. Cell. 2013 Mar;153(1):38–55.

18. Seto E, Yoshida M. Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. Cold Spring Harb Perspect Biol. 2014 Apr 1;6(4):a018713–a018713.

19. Berger SL. The complex language of chromatin regulation during transcription. Nature. 2007 May;447(7143):407–12.

20. Ding J, Liu L, Chiang YL, Zhao M, Liu H, Yang F, et al. Discovery and Structure-Based Design of Inhibitors of the WD Repeat-Containing Protein 5 (WDR5)–MYC Interaction. J Med Chem. 2023 Jun 22;66(12):8310–23.

21. Golbabapour S, Majid NA, Hassandarvish P, Hajrezaie M, Abdulla MA, Hadi AHA. Gene Silencing and Polycomb Group Proteins: An Overview of their Structure, Mechanisms and Phylogenetics. OMICS J Integr Biol. 2013 Jun;17(6):283–96.

22. Cao R, Zhang Y. The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr Opin Genet Dev. 2004 Apr;14(2):155–64.

23. Klymenko T, Papp B, Fischle W, Köcher T, Schelder M, Fritsch C, et al. A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. Genes Dev. 2006 May 1;20(9):1110–22.

24. Cutucache CE, Iqbal J, Bierman PJ, Bociek RG, Weisenburger DD, Joshi SS. Polycomb response element-binding sites in the MDR of CLL: Potential tumor suppressor regulation. Adv Biosci Biotechnol. 2013;04(01):129–35.

25. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, et al. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. Science. 2002 Nov;298(5595):1039–43.

26. Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. Drosophila Enhancer of Zeste/ESC Complexes Have a Histone H3 Methyltransferase Activity that Marks Chromosomal Polycomb Sites. Cell. 2002 Oct;111(2):185–96.

27. Müller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, et al. Histone Methyltransferase Activity of a Drosophila Polycomb Group Repressor Complex. Cell. 2002 Oct;111(2):197–208.

28. Laugesen A, Højfeldt JW, Helin K. Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. Mol Cell. 2019 Apr;74(1):8–18.

29. Ferrari KJ, Scelfo A, Jammula S, Cuomo A, Barozzi I, Stützer A, et al. Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. Mol Cell. 2014 Jan;53(1):49–62.

30. Højfeldt JW, Laugesen A, Willumsen BM, Damhofer H, Hedehus L, Tvardovskiy A, et al. Accurate H3K27 methylation can be established de novo by SUZ12-directed PRC2. Nat Struct Mol Biol. 2018 Mar;25(3):225–32.

31. Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, et al. CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. Development. 2009 Sep 15;136(18):3131–41.

32. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. Nature. 2011 Jan 20;469(7330):343–9.

33. Smits AH, Jansen PWTC, Poser I, Hyman AA, Vermeulen M. Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Res. 2013 Jan 1;41(1):e28–e28.

34. Cao R, Zhang Y. SUZ12 Is Required for Both the Histone Methyltransferase Activity and the Silencing Function of the EED-EZH2 Complex. Mol Cell. 2004 Jul;15(1):57–67.

35. Pasini D, Bracken AP, Jensen MR, Denchi EL, Helin K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. 2004 Oct 13;23(20):4061–71.

36. Kim J, Lee Y, Lu X, Song B, Fong KW, Cao Q, et al. Polycomb- and Methylation-Independent Roles of EZH2 as a Transcription Activator. Cell Rep. 2018 Dec;25(10):2808-2820.e4.

37. Ciferri C, Lander GC, Maiolica A, Herzog F, Aebersold R, Nogales E. Molecular architecture of human polycomb repressive complex 2. eLife. 2012 Oct 30;1:e00005.

38. Simon JA, Lange CA. Roles of the EZH2 histone methyltransferase in cancer epigenetics. Mutat Res Mol Mech Mutagen. 2008 Dec;647(1–2):21–9.

39. Neri F, Zippo A, Krepelova A, Cherubini A, Rocchigiani M, Oliviero S. Myc Regulates the Transcription of the PRC2 Gene To Control the Expression of Developmental Genes in Embryonic Stem Cells. Mol Cell Biol. 2012 Feb 1;32(4):840–51.

40. León TE, Rapoz-D'Silva T, Bertoli C, Rahman S, Magnussen M, Philip B, et al. *EZH2* -Deficient T-cell Acute Lymphoblastic Leukemia Is Sensitized to CHK1 Inhibition through Enhanced Replication Stress. Cancer Discov. 2020 Jul 1;10(7):998–1017.

41. Andrieu GP, Kohn M, Simonin M, Smith CL, Cieslak A, Dourthe MÉ, et al. PRC2 loss of function confers a targetable vulnerability to BET proteins in T-ALL. Blood. 2021 Nov 11;138(19):1855–69.

42. Li LY. EZH2: novel therapeutic target for human cancer. BioMedicine. 2014 Mar;4(1):1.

43. Tan J zhi, Yan Y, Wang X xi, Jiang Y, Xu HE. EZH2: biology, disease, and structure-based drug discovery. Acta Pharmacol Sin. 2014 Feb;35(2):161–74.

44. Antonysamy S, Condon B, Druzina Z, Bonanno JB, Gheyi T, Zhang F, et al. Structural Context of Disease-Associated Mutations and Putative Mechanism of Autoinhibition Revealed by X-Ray Crystallographic Analysis of the EZH2-SET Domain. Rocha S, editor. PLoS ONE. 2013 Dec 19;8(12):e84147.

45. Xiao B, Wilson JR, Gamblin SJ. SET domains and histone methylation. Curr Opin Struct Biol. 2003 Dec;13(6):699–705.

46. Xiao B, Jing C, Kelly G, Walker PA, Muskett FW, Frenkiel TA, et al. Specificity and mechanism of the histone methyltransferase Pr-Set7. Genes Dev. 2005 Jun 15;19(12):1444–54.

47. Mesa AM, Rosenfeld CS, Tuteja G, Medrano TI, Cooke PS. The Roles of the Histone Protein Modifier EZH2 in the Uterus and Placenta. Epigenomes. 2020 Sep 2;4(3):20.

48. Chammas P, Mocavini I, Di Croce L. Engaging chromatin: PRC2 structure meets function. Br J Cancer. 2020 Feb 4;122(3):315–28.

49. Wang J, Wang GG. No Easy Way Out for EZH2: Its Pleiotropic, Noncanonical Effects on Gene Regulation and Cellular Function. Int J Mol Sci. 2020 Dec 14;21(24):9501.

50. Batool A. Role of EZH2 in cell lineage determination and relative signaling pathways. Front Biosci. 2019;24(5):947–60.

51. Chan YS, Göke J, Lu X, Venkatesan N, Feng B, Su IH, et al. A PRC2-Dependent Repressive Role of PRDM14 in Human Embryonic Stem Cells and Induced Pluripotent Stem Cell Reprogramming. Stem Cells. 2013 Apr 1;31(4):682–92.

52. Schuettengruber B, Bourbon HM, Di Croce L, Cavalli G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell. 2017 Sep;171(1):34–57.

53. Alharbi RA, Pettengell R, Pandha HS, Morgan R. The role of HOX genes in normal hematopoiesis and acute leukemia. Leukemia. 2013 May;27(5):1000–8.

54. Aryal S, Zhang Y, Wren S, Li C, Lu R. Molecular regulators of HOXA9 in acute myeloid leukemia. FEBS J. 2023 Jan;290(2):321–39.

55. Shah N, Sukumar S. The Hox genes and their roles in oncogenesis. Nat Rev Cancer. 2010 May;10(5):361–71.

56. Sun Y, Zhou B, Mao F, Xu J, Miao H, Zou Z, et al. HOXA9 Reprograms the Enhancer Landscape to Promote Leukemogenesis. Cancer Cell. 2018 Oct;34(4):643-658.e5.

57. Göllner S, Oellerich T, Agrawal-Singh S, Schenk T, Klein HU, Rohde C, et al. Loss of the histone methyltransferase EZH2 induces resistance to multiple drugs in acute myeloid leukemia. Nat Med. 2017 Jan;23(1):69–78.

58. Khan SN, Jankowska AM, Mahfouz R, Dunbar AJ, Sugimoto Y, Hosono N, et al. Multiple mechanisms deregulate EZH2 and histone H3 lysine 27 epigenetic changes in myeloid malignancies. Leukemia. 2013 Jun;27(6):1301–9.

59. Zohren F, Souroullas GP, Luo M, Gerdemann U, Imperato MR, Wilson NK, et al. The transcription factor Lyl-1 regulates lymphoid specification and the maintenance of early T lineage progenitors. Nat Immunol. 2012 Aug;13(8):761–9.

60. Cleary ML, Mellentin JD, Spies J, Smith SD. Chromosomal translocation involving the beta T cell receptor gene in acute leukemia. J Exp Med. 1988 Feb 1;167(2):682–7.

61. Lukov GL, Rossi L, Souroullas GP, Mao R, Goodell MA. The expansion of T-cells and hematopoietic progenitors as a result of overexpression of the lymphoblastic leukemia gene, Lyl1 can support leukemia formation. Leuk Res. 2011 Mar;35(3):405–12.

62. Zhong Y, Jiang L, Hiai H, Toyokuni S, Yamada Y. Overexpression of a transcription factor LYL1 induces T- and B-cell lymphoma in mice. Oncogene. 2007 Oct 18;26(48):6937–47.

63. Belver L, Ferrando A. The genetics and mechanisms of T cell acute lymphoblastic leukaemia. Nat Rev Cancer. 2016 Aug;16(8):494–507.

64. Paul S, Kantarjian H, Jabbour EJ. Adult Acute Lymphoblastic Leukemia. Mayo Clin Proc. 2016 Nov;91(11):1645–66.

65. Pui CH, Evans WE. Treatment of Acute Lymphoblastic Leukemia. N Engl J Med. 2006 Jan 12;354(2):166–78.

66. Pui CH, Pei D, Campana D, Bowman WP, Sandlund JT, Kaste SC, et al. Improved Prognosis for Older Adolescents With Acute Lymphoblastic Leukemia. J Clin Oncol. 2011 Feb 1;29(4):386–91.

67. Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, Payne-Turner D, et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature. 2012 Jan;481(7380):157–63.

68. Simon C, Chagraoui J, Krosl J, Gendron P, Wilhelm B, Lemieux S, et al. A key role for *EZH2* and associated genes in mouse and human adult T-cell acute leukemia. Genes Dev. 2012 Apr 1;26(7):651–6.

69. Fattizzo B, Rosa J, Giannotta JA, Baldini L, Fracchiolla NS. The Physiopathology of T- Cell Acute Lymphoblastic Leukemia: Focus on Molecular Aspects. Front Oncol. 2020 Feb 28;10:273.

70. Girardi T, Vicente C, Cools J, De Keersmaecker K. The genetics and molecular biology of T-ALL. Blood. 2017 Mar 2;129(9):1113–23.

71. Liu Y, Easton J, Shao Y, Maciaszek J, Wang Z, Wilkinson MR, et al. The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. Nat Genet. 2017 Aug;49(8):1211–8.

72. Huether R, Dong L, Chen X, Wu G, Parker M, Wei L, et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. Nat Commun. 2014 Apr 8;5(1):3630.

73. Sin C fung, Man P hei M. Early T-Cell Precursor Acute Lymphoblastic Leukemia: Diagnosis, Updates in Molecular Pathogenesis, Management, and Novel Therapies. Front Oncol. 2021 Nov 29;11:750789.

74. Meyer LK, Roy RP, Pölönen P, Elsayed A, Huang BJ, Kimura S, et al. A Transcriptional Classifier Identifies Pediatric T-Cell Acute Lymphoblastic Leukemias at High Risk for End of Induction Minimal Residual Disease. Blood. 2022 Nov 15;140(Supplement 1):1722–3.

75. Xue C, Greene EC. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. Trends Genet. 2021 Jul;37(7):639–56.

76. Zaboikin M, Zaboikina T, Freter C, Srinivasakumar N. Non-Homologous End Joining and Homology Directed DNA Repair Frequency of Double-Stranded Breaks Introduced by Genome Editing Reagents. Hu W, editor. PLOS ONE. 2017 Jan 17;12(1):e0169931.

77. Song X, Huang H, Xiong Z, Ai L, Yang S. CRISPR-Cas9 ^{D10A} Nickase-Assisted Genome Editing in Lactobacillus casei. Atomi H, editor. Appl Environ Microbiol. 2017 Nov 15;83(22):e01259-17.

78. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell. 2013 Sep;154(6):1380–9.

79. Koch B, Nijmeijer B, Kueblbeck M, Cai Y, Walther N, Ellenberg J. Generation and validation of homozygous fluorescent knock-in cells using CRISPR–Cas9 genome editing. Nat Protoc. 2018 Jun;13(6):1465–87.

80. IDT Custom Alt-R CRISPR-Cas9 gRNA [Internet]. Available from: https://eu.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM.

81. Danis E, Yamauchi T, Echanique K, Zhang X, Haladyna JN, Riedel SS, et al. Ezh2 Controls an Early Hematopoietic Program and Growth and Survival Signaling in Early T Cell Precursor Acute Lymphoblastic Leukemia. Cell Rep. 2016 Mar;14(8):1953–65.

82. Kusakabe M, Sun AC, Tyshchenko K, Wong R, Nanda A, Shanna C, et al. Synthetic modeling reveals HOXB genes are critical for the initiation and maintenance of human leukemia. Nat Commun. 2019 Jul 2;10(1):2913.

83. Gentile C, Berlivet S, Mayran A, Paquette D, Guerard-Millet F, Bajon E, et al. PRC2-Associated Chromatin Contacts in the Developing Limb Reveal a Possible Mechanism for the Atypical Role of PRC2 in HoxA Gene Expression. Dev Cell. 2019 Jul;50(2):184-196.e4.

84. Booth CAG, Barkas N, Neo WH, Boukarabila H, Soilleux EJ, Giotopoulos G, et al. Ezh2 and Runx1 Mutations Collaborate to Initiate Lympho-Myeloid Leukemia in Early Thymic Progenitors. Cancer Cell. 2018 Feb;33(2):274-291.e8.

85. Yang L, Chen F, Zhu H, Chen Y, Dong B, Shi M, et al. 3D genome alterations associated with dysregulated HOXA13 expression in high-risk T-lineage acute lymphoblastic leukemia. Nat Commun. 2021 Jun 17;12(1):3708.

86. Hu Y, Smyth GK. ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J Immunol Methods. 2009 Aug;347(1–2):70–8.

87. Canté-Barrett K, Mendes RD, Li Y, Vroegindeweij E, Pike-Overzet K, Wabeke T, et al. Loss of CD44dim Expression from Early Progenitor Cells Marks T-Cell Lineage Commitment in the Human Thymus. Front Immunol [Internet]. 2017 Jan 20 [cited 2024 Feb 16];8. Available from: http://journal.frontiersin.org/article/10.3389/fimmu.2017.00032/full

88. Jain N, Lamb AV, O'Brien S, Ravandi F, Konopleva M, Jabbour E, et al. Early T-cell precursor acute lymphoblastic leukemia/lymphoma (ETP-ALL/LBL) in adolescents and adults: a high-risk subtype. Blood. 2016 Apr 14;127(15):1863–9.

89. Wang C, Oshima M, Sato D, Matsui H, Kubota S, Aoyama K, et al. Ezh2 loss propagates hypermethylation at T cell differentiation–regulating genes to promote leukemic transformation. J Clin Invest. 2018 Aug 31;128(9):3872–86.

90. Panelli P, De Santis E, Colucci M, Tamiro F, Sansico F, Miroballo M, et al. Noncanonical β-catenin interactions promote leukemia-initiating activity in early T-cell acute lymphoblastic leukemia. Blood. 2023 Mar 30;141(13):1597–609.

91. Lee SH, Li Y, Kim H, Eum S, Park K, Lee CH. The role of EZH1 and EZH2 in development and cancer. BMB Rep. 2022 Dec 31;55(12):595–601.

92. Sun S, Yu F, Xu D, Zheng H, Li M. EZH2, a prominent orchestrator of genetic and epigenetic regulation of solid tumor microenvironment and immunotherapy. Biochim Biophys Acta BBA - Rev Cancer. 2022 Mar;1877(2):188700.

93. Ben-Bassat H, Shlomai Z, Kohn G, Prokocimer M. Establishment of a human T-acute lymphoblastic leukemia cell line with a (16;20) chromosome translocation. Cancer Genet Cytogenet. 1990 Oct;49(2):241–8.

94. Kumar A, Drusbosky LM, Meacham A, Turcotte M, Bhargav P, Vasista S, et al. Computational modeling of early T-cell precursor acute lymphoblastic leukemia (ETP-ALL) to identify personalized therapy using genomics. Leuk Res. 2019 Mar;78:3–11.

95. Henry CJ, Marusyk A, DeGregori J. Aging-Associated Changes in Hematopoiesis and Leukemogenesis: What's the Connection? Aging. 2011 Jul 2;3(6):643–56.

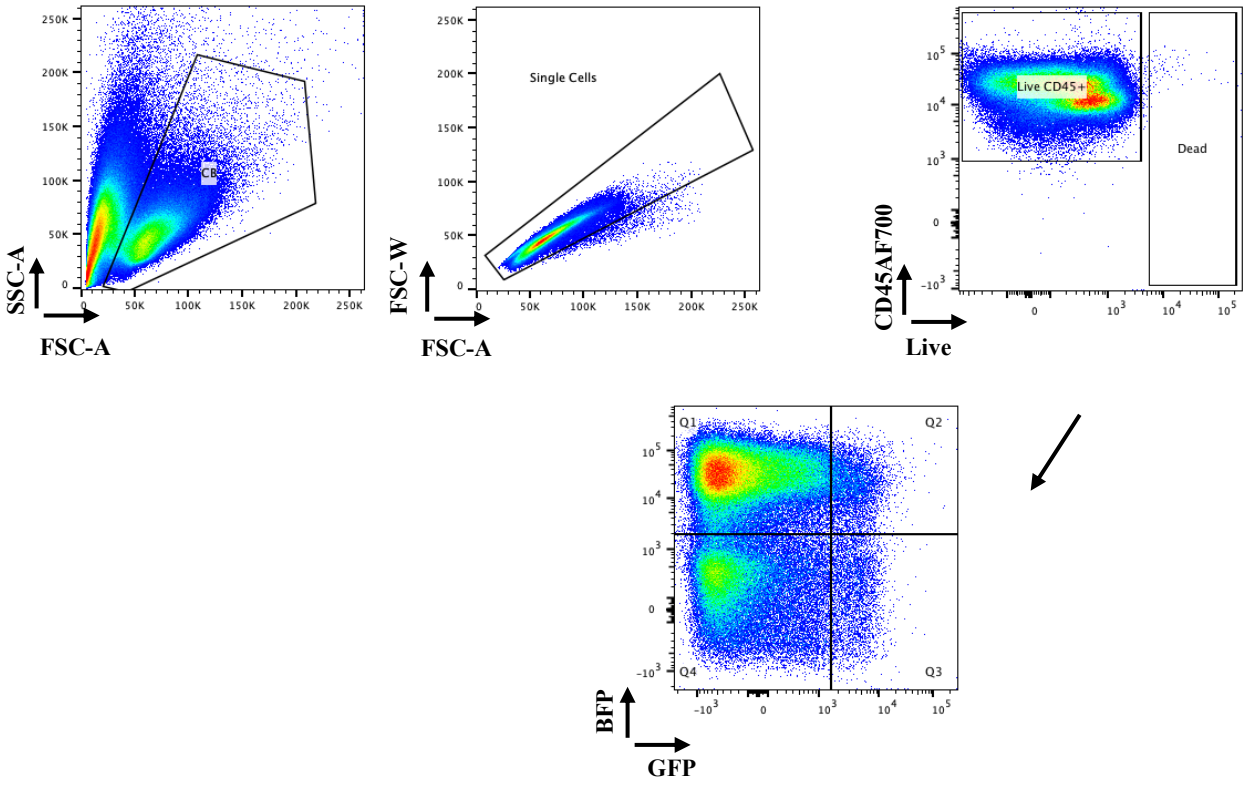
96. Henry CJ, Casás-Selves M, Kim J, Zaberezhnyy V, Aghili L, Daniel AE, et al. Aging-associated inflammation promotes selection for adaptive oncogenic events in B cell progenitors. J Clin Invest. 2015 Nov 9;125(12):4666–80.

97. Irving J, Matheson E, Minto L, Blair H, Case M, Halsey C, et al. Ras pathway mutations are prevalent in relapsed childhood acute lymphoblastic leukemia and confer sensitivity to MEK inhibition. Blood. 2014 Nov 27;124(23):3420–30.

98. Ntziachristos P, Tsirigos A, Vlierberghe PV, Nedjic J, Trimarchi T, Flaherty MS, et al. Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nat Med. 2012 Feb;18(2):298–302.

99. Tarantini F, Cumbo C, Anelli L, Zagaria A, Specchia G, Musto P, et al. Inside the biology of early T-cell precursor acute lymphoblastic leukemia: the perfect trick. Biomark Res. 2021 Dec;9(1):89.

Appendix A An example of cytometer gating strategy



Appendix B Overview of CRISPR/Cas9 strategy

YELLOW = ARM5'
BLU= ARM 3'
BOLD RED =EXON 10 EZH2
crRNAs

EZH2_EXON10

TTTACAAATTCACAATATAGAACTGTCTTGCCTGATTTTTGTACCATTCTAG
GAATATATAGAACCCATATTTCTCTTTCTGAAATTTTCAGACTATCAAAAG
CATATCAAAACAAGTATATTAGACATGCAAATGATTATTTGTGATAAATGG
ATAATGTGATACATTTTTTGACTAACCTGGCTTATATAGTATTTTTTTTTCT
CTTCCATCAAAATGAGTTTTAGAACTTTGCCCTGATGTTGACATTTTTCATT
TCGTAG**GAGGGAGCAAAGGAGTTTGCTGCTGCTCTCACCGCTGAGCGG**
ATAAAGACCCCACCAAAACGTCCAGGAGGCCGCAGAAGAGGACGGCT
TCCCAATAACAGTAGCAGGCCCAGCAAACAGTAGCAGGCCCAGCACC
CCCACCATTAATGTGCTGGAATCAAAGGATACAGACAGTGATAGGGAA
GCAGGGACTGAAACGGGGGGAGAGAACAATGATAAAGAAGAAGAAGA
GAAGAAAGATGAAACTTCGAGCTCCTCTGGTAAGACACGTCTAATAACTG
GGTTTTACTGTTCTGTGAAAGTTCGTGTTGTGAGGATTAAGGATGTATAAT
GCATTGAATAATTTTCTTAGTTGGTTAGTTTCAGTATAAAGACCAGAGTTA
TCTCAAGAATGTGTAGCTGTGTTGTTTCTGTTTCCTGCCTGCACCCGTACCC
TAGCTTGTGGTGTTTGCCCTGCTTGGTTGTTTGAAAGCTGATGATGCACTTT
ATTCTCATTCTTTGTGCCATTTTCTTTTTTCGATCTTTCTCAGGTAGGGAATA
ATGAAATTTGGATCATTTGGTTTAATATTCTTGACGTTCTGTGAGTAGTTTT
GTTGGAAAATGTGAACTACGATGGGTTAGTGTTTTGCCGATTGGATTTGAG
TTGTCTCATCT