Linker histone variant H1-0 dysregulation in *ETV6::RUNX1*$^{+}$ preleukemia and B cell acute lymphoblastic leukemia

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Vera Helena Jepsen
aus Aachen

Düsseldorf, September 2024

aus dem Institut für Kinder-Onkologie, -Hämatologie und Klinische Immunologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

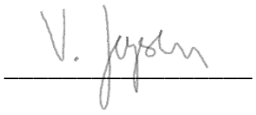
1. Prof. Dr. Arndt Borkhardt
2. Prof. Dr. Sebastian Wesselborg

Tag der mündlichen Prüfung: 24.01.2025

EIDESSTATTLICHE ERKLÄRUNG

Ich versichere an Eides Statt, dass die Dissertation „Linker histone variant H1-0 dysregulation in *ETV6::RUNX1*$^{+}$ preleukemia and B cell acute lymphoblastic leukemia“ von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Die Dissertation wurde in dieser oder ähnlicher Weise noch an keiner anderen Fakultät vorgelegt und es wurde kein vorheriger Promotionsversuch unternommen.



Vera H. Jepsen

Düsseldorf, September 2024

TABLE OF CONTENTS

SUMMARY

The chromosomal translocation t(12;21)(p13;q22) gives rise to *ETV6::RUNX1*, the most common oncogenic fusion gene in pediatric B cell precursor acute lymphoblastic leukemia (BCP-ALL). *ETV6::RUNX1* arises before birth at high frequency and induces a clinically silent state that can persist for over a decade. In <1% of carriers, these preleukemic cells acquire secondary mutations that induce transformation to overt leukemia. The mechanisms contributing to quiescence of *ETV6::RUNX1*$^+$ preleukemic cells still remain elusive. In this thesis, factors involved in *ETV6::RUNX1*$^+$ preleukemia are characterized by generating CRISPR/Cas9-edited human induced pluripotent stem cell (hiPSC) models. These preleukemic hiPSC models express ETV6::RUNX1 at physiological levels via the endogenous *ETV6* promoter. Transcriptional analyses identified upregulation of linker histone H1-0 at the *ETV6::RUNX1*$^+$ preleukemic state, both in hiPSCs and during early B lymphoid differentiation. Moreover, publicly available expression data of 3,026 leukemia patient samples showed significantly elevated H1-0 levels in *ETV6::RUNX1*$^+$ BCP-ALL compared to other leukemia entities. Dual-luciferase promoter assays revealed that *H1-0* promoter activity can be induced by ETV6::RUNX1, but not by RUNX1. While chromatin immunoprecipitation assays did not indicate direct binding of ETV6::RUNX1 to the *H1-0* promoter, H1-0 expression might be indirectly regulated via DNA methylation of a CpG island shore element or histone acetylation. Depletion of H1-0 via RNA interference affected epigenetic processes and inhibited ETV6::RUNX1 signature genes, including *RAG1* and *EPOR*, indicating a key role of H1-0 in regulating the ETV6::RUNX1 transcriptome. Analysis of single-cell sequencing data showed that *H1-0* is highly expressed in quiescent hematopoietic cells. H1-0 expression can be induced in BCP-ALL cell lines by addition of histone deacetylase inhibitors (HDACis) and H1-0 protein levels correspond to susceptibility towards this drug class. Following up on these findings, combinatorial drug treatments using the H1-0-inducing HDACi Quisinostat were performed. These experiments showed promising synergism of Quisinostat with established chemotherapeutic drugs used for B-ALL treatment and the proteasome inhibitor Bortezomib in *ETV6::RUNX1*$^+$ cells. Taken together, these data identify H1-0 as a key regulator of the *ETV6::RUNX1*$^+$ transcriptome and indicate that induction of H1-0 via HDACis might be a potential novel approach to improve *ETV6::RUNX1*$^+$ BCP-ALL treatment.

ZUSAMMENFASSUNG

Die chromosomale Translokation t(12;21)(p13;q22) führt zur Bildung von *ETV6::RUNX1*, dem häufigsten Fusionsgen der akuten lymphatischen Leukämie (ALL) im Kindesalter. ETV6::RUNX1 entsteht bereits vor der Geburt mit hoher Inzidenz und induziert einen präleukämischen Zustand, der über ein Jahrzehnt persistieren kann. Die präleukämischen Zellen erwerben in <1% der *ETV6::RUNX1*-positiven Neugeborenen sekundäre Mutationen, die zum Ausbruch der Erkrankung führen. Die Mechanismen, die zum präleukämischen Zustand beitragen, sind bisher weitgehend unbekannt. In dieser Dissertation wurden humane induzierte pluripotente Stammzellmodelle zur Charakterisierung der *ETV6::RUNX1*-positiven Präleukämie mittels CRISPR/Cas9 entwickelt. Physiologische Expression des Fusionsgens in den gentechnisch veränderten Stammzellen wird über den *ETV6* Promoter gesteuert. Mittels transkriptioneller Analyse wurde eine Hochregulierung des Histons H1-0 in den *ETV6::RUNX1*$^{+}$ Stammzellen sowie frühen B-lymphoiden Vorläuferzellen identifiziert. Außerdem zeigte die Untersuchung publizierter Expressionsdaten von 3.026 ALL Patientenproben ein signifikant erhöhtes H1-0 Level in *ETV6::RUNX1*-positiver ALL im Vergleich zu anderen Leukämiesubtypen. In Luciferase-Reporter-Assays konnte zudem gezeigt werden, dass der *H1-0* Promoter durch ETV6::RUNX1, jedoch nicht durch RUNX1, aktiviert wird. Da publizierte Chromatin-Immunopräzipitationsdaten keine direkte Bindung von ETV6::RUNX1 an den *H1-0* Promoter zeigten, ist eine indirekte Regulation der H1-0 Expression über gezielte DNA Methylierung der *H1-0* CpG *island shore* oder durch Histon-Acetylierung denkbar. Knockdown von H1-0 durch RNA-Interferenz beeinflusste epigenetische Prozesse und hemmte ETV6::RUNX1-spezifische Gene, einschließlich *RAG1* und *EPOR*. Dies lässt auf eine Schlüsselrolle von H1-0 bei der Regulation des ETV6::RUNX1-Transkriptoms schließen. RNA-Sequenzierung von Einzelzellen zeigte, dass H1-0 in ruhenden hämatopoetischen Stammzellen hoch exprimiert wird. Die Expression von H1-0 kann in ALL-Zelllinien durch Zugabe von Histondeacetylase-Inhibitoren (HDACis) induziert werden und ist ein Indikator für die Sensitivität gegenüber dieser Wirkstoffklasse. Aufbauend auf diesen Erkenntnissen wurde die Sensitivität von ALL-Zelllinien gegenüber dem H1-0-induzierenden HDACi Quisinostat in Kombination mit verschiedenen Wirkstoffen getestet. Diese Experimente zeigten vielversprechenden Synergismus von Quisinostat mit etablierten ALL Chemotherapeutika, sowie dem Proteasom-Inhibitor Bortezomib in *ETV6::RUNX1*-positiven Zellen. Zusammengefasst identifizieren diese Daten H1-0 als wichtigen Regulator des ETV6::RUNX1 Transkriptoms. Die gewonnen Erkenntnisse deuten darauf hin, dass eine Induktion von H1-0 über HDACis ein potentieller neuer Ansatz zur besseren Behandlung von *ETV6::RUNX1*-positiven Patienten sein könnte.

1. INTRODUCTION

1.1 Pediatric leukemia

Despite significant treatment progress in recent decades, cancer remains a primary cause of death in children [1]. In Europe and the United States, 16.2 per 100,000 children aged 0-14 years were diagnosed with cancer in 2020 [2] and five-year survival rates were at 85.1% (United States: 2012-2018 [3]). Among all pediatric cancers, leukemia is the most prevalent subtype, followed by central nervous system (CNS) tumors and lymphomas [4] (**Figure 1.1**).

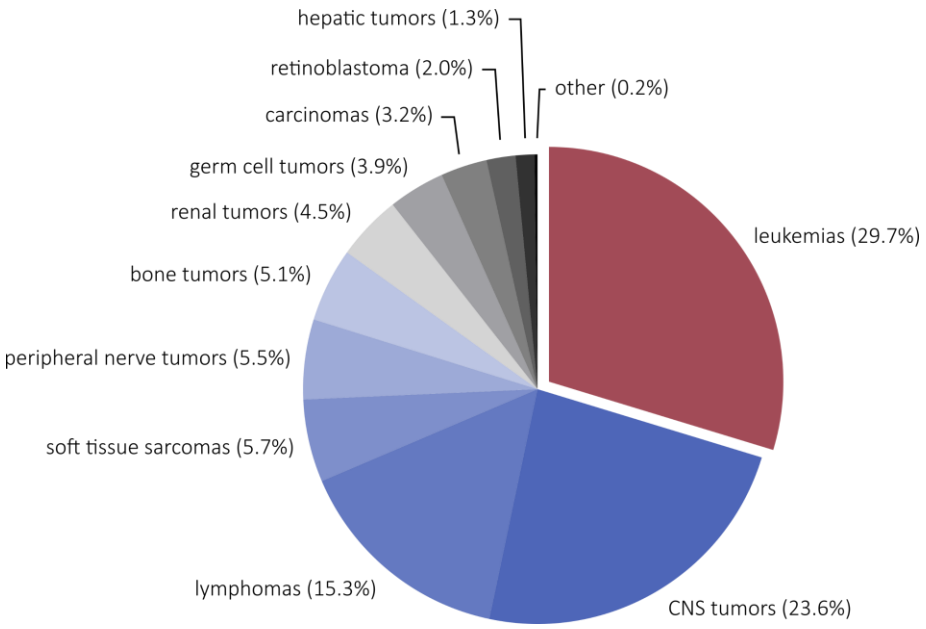


Figure 1.1: Distribution of pediatric cancer cases reported between 2009-2018 in Germany (German Childhood Cancer Registry). Data is based on patients aged <18 years and adapted from [4]. CNS: central nervous system.

Leukemia (greek: *leukós* "white", *haima* "blood") is characterized by the uncontrolled growth of malignant white blood cells and can be stratified into subgroups based on the specific cell lineage (lymphoid or myeloid), differentiation status, speed of progression and genetic alterations. Accordingly, leukemias are classified as acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), chronic lymphoid leukemia (CLL) and chronic myeloid leukemia (CML). While acute leukemias are characterized by rapid disease progression that requires immediate treatment, chronic leukemias have a slower progression. The most frequent type of leukemia in childhood is ALL (75%), whereas chronic leukemias are rare in children [5].

1.2 B cell precursor acute lymphoblastic leukemia (BCP-ALL)

B cell precursor acute lymphoblastic leukemia (BCP-ALL) is the most common subtype of pediatric leukemia and arises from the malignant transformation of a B cell precursor. BCP-ALL has a peak in incidence at the ages of 2-5 years [6]. Immunophenotypically, BCP-ALL presents with strong positivity of at least two of the four antigens CD19, CD10, CD22 and CD79A [7].

B cell development is a tightly regulated process that involves expression of cell state specific transcription factors, cell surface antigens (also referred to as cluster of differentiation (CD) markers) and rearrangement of immunoglobulin (Ig) genes (**Figure 1.2**). All B lymphocytes arise from multipotent hematopoietic stem cells (HSCs). Within the primary B lymphoid organs (in the fetal liver or later in the bone marrow), HSCs differentiate into lymphoid-primed multipotent progenitors (LMPPs) that in turn give rise to common lymphoid progenitors (CLPs). CLPs have the potential to develop into either B cells, T cells, natural killer cells or dendritic cells. Final commitment of CLPs to the B cell lineage requires activation of several transcription factors, including TCF3, FOXO1, EBF1 and PAX5 [8].

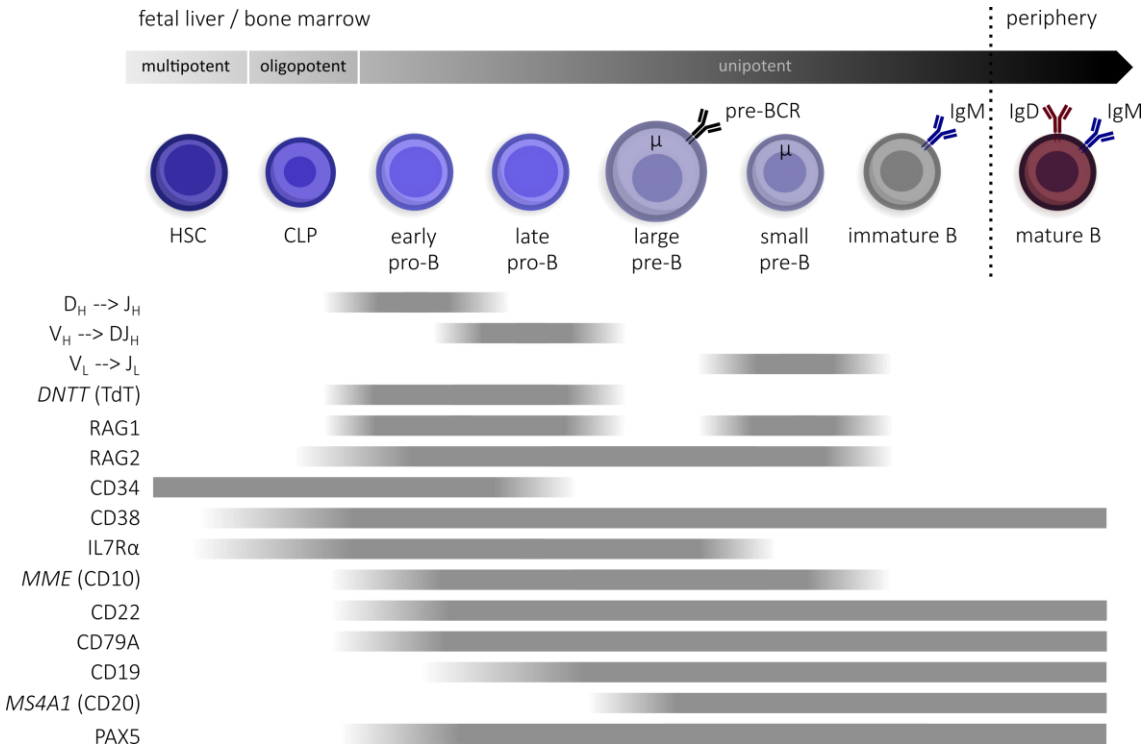


Figure 1.2: Schematic representation of human B cell differentiation. B cell developmental states can be identified by status of immunoglobulin (Ig) gene rearrangements and expression of specific surface molecules or transcription factors. Modified from [9] and [10]. BCR: B cell receptor, HSC: hematopoietic stem cell, CLP: common lymphoid progenitor, μ: μ heavy chain, H: heavy chain, L: light chain, CD: cluster of differentiation.

Ig gene rearrangement within the B cell precursor compartment occurs in a hierarchical manner mediated by the DNA-cleaving recombination-activating genes *RAG1* and *RAG2*. Expression of these genes is restricted to early B lineage cells and crucial for recombination of their variable (V), diverse (D) and joining (J) Ig gene sequences to provide a diverse antibody repertoire and efficient adaptive immune response via mature B cells. RAG1 and RAG2 form a complex that induces DNA double-strand breaks at recombination signal sequences (RSSs). RSSs contain a conserved heptamer site (5'-CACATGT-3'), a spacer of either 12 or 23 nucleotides and a nonamer sequence (5'-ACAAAAACC-3') [11, 12]. The resulting DNA breaks are repaired via the non-homologous end joining (NHEJ) pathway. Additional diversity of immunoglobulin sequences is provided by imprecise repair at the cleaved DNA site, either by nucleotide addition (e.g. via terminal deoxynucleotidyl transferase (TdT) activity) or exonucleolytic cleavage [12]. In the context of specific subtypes of leukemia (e.g. BCP-ALL characterized by the *ETV6::RUNX1* fusion gene), analyses of genomic RSS distribution near breakpoint regions indicated aberrant RAG1 and RAG2 activity as a major cause of secondary oncogenic lesions [13, 14].

In healthy cells, successful in-frame rearrangement of Ig heavy (H) chain genes ($D_H \rightarrow J_H$ and subsequent $V_H \rightarrow DJ_H$ joining) at the pro-B cell state results in the production of μ heavy chains. These assemble with the surrogate light chains VPREB1 and IGLL1, as well as CD79A/B, to form the pre-B cell receptor (pre-BCR). Upon functional assembly of the pre-BCR, Ig light (L) chain gene joining ($V_L \rightarrow J_L$) takes place in the small pre-B cell compartment, resulting in expression of either κ or λ light chains. These participate in the formation of IgM on the cell surface of immature B cells which ultimately differentiate into mature IgM^+IgD^+ B cells that populate the secondary lymphoid organs of the body (e.g. lymph nodes). [15]

1.2.1 Genetic subtypes

BCP-ALL is biologically and clinically heterogeneous. It can be stratified into subgroups based on common chromosomal translocations, aneuploidies or gene expression profiles. These subgroups differ in incidence, age of onset, pathogenesis and outcomes [16] (**Figure 1.3**). Risk stratification based on detection of genetic abnormalities has considerably improved survival of BCP-ALL patients.

The most common BCP-ALL subtypes are high hyperdiploidy and BCP-ALL carrying the t(12;21)(p13;q22) translocation, encoding for the fusion protein ETV6::RUNX1, that will be introduced in more detail in chapter 1.3. High hyperdiploidy was first described by Lampert *et al*. and is characterized by non-random gain of whole chromosomes (52-67 chromosomes)

[17, 18]. Typically, chromosomal gains involve chromosomes X, 4, 6, 8, 10, 14, 17, 18, and 21 [19]. It was suggested that the hyperdiploid genotype pattern is caused by a single abnormal mitosis and simultaneous gain of chromosomes [20]. Overall, high hyperdiploid BCP-ALL has a favorable prognosis [21].

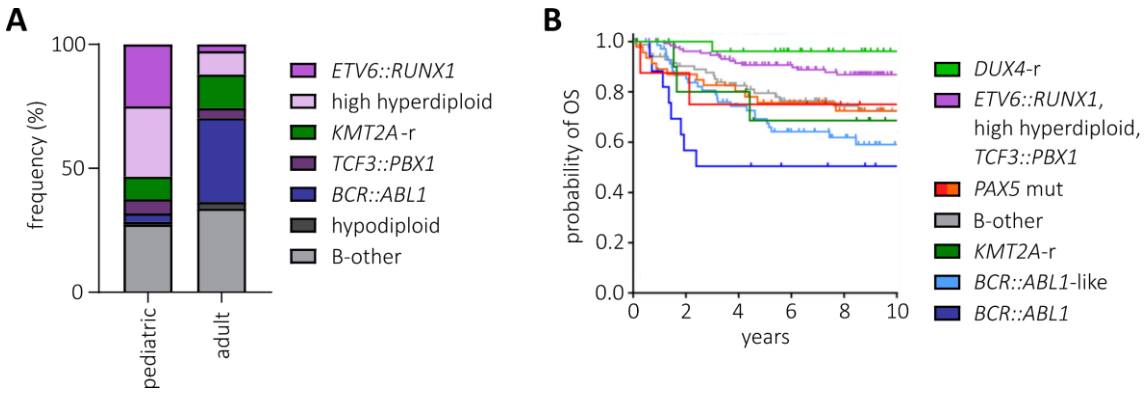


Figure 1.3: Relative frequency and overall survival of BCP-ALL subgroups. Chromosomal abnormalities and mutations present in the subgroups are indicated. Data of **(A)** relative leukemia frequencies is derived from [22] and **(B)** overall survival (OS) data is derived from [23]. -r: rearranged, mut: mutated.

The reciprocal translocation between chromosomes 9q34 and 22q11 results in formation of a shortened chromosome 22, i.e. the Philadelphia chromosome, and expression of the *BCR::ABL1* fusion gene. *BCR::ABL1*$^+$ leukemia is significantly more common in adults (**Figure 1.3A**). Treatment with selective tyrosine kinase inhibitors (TKIs) considerably improved outcome of patients with *BCR::ALB1*$^+$ CML, whereas patients with *BCR::ABL1*$^+$ BCP-ALL still present with poor survival, often accompanied by relapse [24, 25]. Unfavorable prognosis of childhood *BCR::ABL1*$^+$ BCP-ALL treated with the TKI Imatinib was associated with *IKZF1* deletions [26].

KMT2A (formerly known as *MLL*) rearrangements are typically found in infant BCP-ALL (children <1 year) and are associated with poor treatment outcome. Fusion genes involving *KMT2A* are likely sufficient for leukemia development, given the high concordance rate in monozygotic twins and absence of secondary, cooperative mutations in this leukemia subtype [27].

In recent years, gene expression profiling and DNA sequencing studies led to further stratification of BCP-ALL patients. Novel subtypes include the *BCR::ABL1*-like and *ETV6::RUNX1*-like BCP-ALL that lack expression of the typical fusion genes but show similar transcriptome profiles as *BCR::ABL1*$^+$ or *ETV6::RUNX1*$^+$ BCP-ALL [28-30]. Further novel subtypes comprise BCP-ALL with *DUX4* [30, 31], *MEF2D* [32], *ZNF384* [33] and *NUTM1* [34], as well as *IG::MYC* rearrangements [35] and alterations of *PAX5* [22, 36].

1.2.2 Leukemogenesis

Development of BCP-ALL usually requires co-operating mutations, as described by the Knudson or two-hit hypothesis [37]. First, acquisition of an initiating gene fusion or chromosomal aneuploidy gives rise to a clinically covert preleukemic clone. After a variable latency period, secondary genetic lesions ultimately lead to the development of overt leukemia (**Figure 1.4**) [38, 39]. Pediatric BCP-ALL is usually characterized by few genetic hits that drive leukemia development, while adult cancers, such as UV light-induced skin lesions, show higher mutational load on average [40].

As revealed by studies of Ig or T cell receptor (TCR) rearrangements in monozygotic twins with concordant ALL [41], initiating mutations are frequently acquired *in utero* during fetal development, with preleukemic cells spreading from one twin to the other via vascular anastomoses of a shared placenta [42, 43]. Additional evidence for the prenatal origin of ALL was provided by the identification of initiating gene lesions in neonatal blood spots (Guthrie cards) and cord blood samples [44, 45] in backtracking studies. Prenatal origin has been suggested for *KMT2A*-rearranged [46, 47], *ETV6::RUNX1*$^{+}$ [45, 48], *TCF3::PBX1*$^{+}$ [49] and *BCR::ABL1*$^{+}$ ALL [50]. Concordance rates for pediatric ALL vary with age and penetrance of the initiating lesion. While *KMT2A*-rearranged infant BCP-ALL presents with almost 100% concordance in twins that share a monochorionic placenta, concordance rates for older children (>18 months) presenting with other ALL subtypes are much lower (5-15%) [41, 51].

Apart from somatically acquired preleukemia-initiating gene lesions, several germline mutations were found to increase ALL susceptibility. These affect genes that are also targets of somatic alterations in ALL, such as *ETV6* [52-54], *PAX5* [55, 56] and *IKZF1* [57, 58]. Likewise, genetic syndromes, such as Li-Fraumeni, Noonan or Down syndrome confer higher risk to develop leukemia [59-61]. In recent years, genome-wide association studies identified further germline variations that show low penetrance but cumulatively may confer higher risk for BCP-ALL. These include single-nucleotide polymorphisms (SNPs) affecting genes such as *IKZF1*, *CDKN2A*, *GATA3*, *PIP4K2A*, *ARID5B*, *LHPP*, *ELK3* and *CEBPE* [62]. For *ETV6::RUNX1*$^{+}$ BCP-ALL in particular, germline SNPs associated with increased leukemia susceptibility were identified at the *TP63* and *PTPRJ* loci [63].

In addition to the underlying genetic germline predisposition or preleukemia-initiating somatic aberrations, development of overt leukemia commonly requires secondary genetic lesions that arise postnatally. Recurrent secondary mutations found in pediatric BCP-ALL cluster to pathways of B cell development (*IKZF1*, *EBF1*, *PAX5*, *ETV6*), proliferation and cell cycle control (*CDKN2A/B*, *RB1*, *BTG1*) as well as cytokine signaling (*CRLF2*) [64]. Secondary mutations vary between

BCP-ALL subtypes. *BCR::ABL1*$^{+}$ BCP-ALL for instance presents with high frequency of *IKZF1* deletions (in >80% of cases) [65], while high hyperdiploid BCP-ALL is associated with secondary mutations affecting RAS pathway signaling or histone modifiers [19].

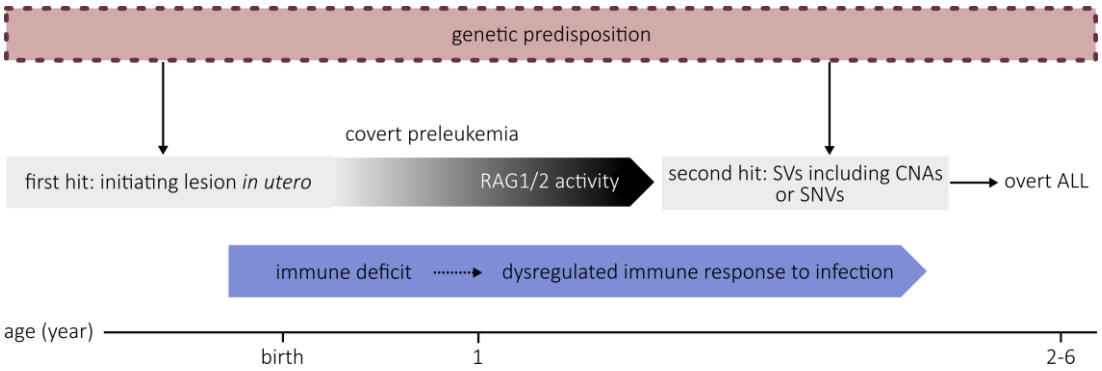


Figure 1.4: Two-hit model of *ETV6::RUNX1*$^{+}$ BCP-ALL development. Leukemia development is a function of genetic background and environmental exposure. Adapted from [6]. SV: structural variant, CNA: copy number alteration, SNV: single-nucleotide variant.

Leukemia develops based on a combination of genetic background and environmental exposures. There is evidence that e.g. ionizing radiation increases the risk to develop BCP-ALL [66]. Additionally, epidemiological studies and leukemia mouse models indicate a role of common infections for leukemogenesis (as reviewed in [6]). These studies suggest that immune training during early childhood protects against BCP-ALL, while an insufficient training may trigger a dysregulated immune response to infection later in life. Accordingly, attendance of daycare [67] was associated with a lower ALL risk and caesarean delivery showed an association with increased risk to develop ALL [68].

1.2.3 Diagnosis and treatment

Patients are diagnosed with BCP-ALL if they present with ≥20% lymphoblast cells in their bone marrow or peripheral blood [69]. Hence, BCP-ALL symptoms are caused by infiltration of the bone marrow and lymphoid system with leukemic cells, and include bone or joint pain, fatigue, bleeding, lymphadenopathy, splenomegaly and/or hepatomegaly [69]. BCP-ALL can be distinguished from lymphoblastic lymphoma, which presents with similar symptoms, primarily by its localization to the bone marrow or blood [70].

Immunophenotyping is a tool routinely used to determine ALL diagnosis. Leukemic subtypes can be distinguished by staining with antibodies directed against specific cell surface or intracellular

antigens, and subsequent flow cytometric analysis. Leukemia entities are defined according to resemblance of their expression profiles to specific cell types (Table 1.1). In addition to immunophenotyping, cytogenetic screening of ALL samples using karyotyping, fluorescence *in situ* hybridization (FISH) or RT-qPCR for fusion detection is performed to enable a more precise prognosis. Modern techniques for stratification of BCP-ALL patients include detection of structural variants via optical genome mapping and transcriptome sequencing to identify -like BCP-ALL subtypes or uncommon fusion genes.

Table 1.1: Surface/intracellular markers for flow cytometric determination of dominant leukemia lineage according to the AIEOP-BFM consensus guidelines (2016). Adapted from [7]. FACS: fluorescence-activated cell sorting.

lineage	FACS markers	criteria
BCP-ALL	$CD19^+$, $CD79A^+$, $CD22^+$, $CD10^+$	≥2 markers positive
T-ALL	$CD3^+$, $CD7^+$, intracellular $MPO^{-/weak}$	all 3 markers present
AML	intracellular MPO^+, $CD13^+$, $CD33^+$, $CD64^+$, $CD65^+$, $CD117^+$	≥2 markers positive

While childhood ALL in the 1960s was generally a deadly disease with 10-year survival rates below 10%, the advent of chemotherapy, first described in 1948 [71], drastically improved patient survival [72]. Over the years, treatment outcomes of pediatric BCP-ALL progressively improved to almost 90% overall survival today [72]. Treatment of pediatric leukemia is therefore considered one of the major success stories of cancer chemotherapy. However, some leukemia subtypes still present with poor prognoses. These include hypodiploidy, *BCR::ABL1*$^+$, *TCF3::HLF*$^+$ and *KMT2A*-rearranged leukemias [73]. Relapse occurs in around 15-20% of lymphoblastic leukemia patients and is accompanied by poor outcome [74]. The most important survival indicator for relapse patients is the time of relapse, with earlier relapses usually presenting with a worse prognosis [75].

ALL treatment is adjusted according to the patient's individual risk features, such as age, gender, white blood cell count at diagnosis, immunophenotype, genetic abnormalities and minimal residual disease values in the bone marrow [76] (Table 1.2). Treatment of ALL consists of different multi-agent chemotherapy courses: remission induction, consolidation and maintenance therapy. To prevent or treat leukemic infiltration of the CNS, CNS-directed therapy is another integral part of ALL therapy. 1-4% of ALL patients present with CNS involvement at diagnosis and around 30% of relapsed leukemias show infiltration of the CNS with blast cells [77, 78].

Table 1.2: Risk factors predicting ALL treatment outcome. Adapted from [76].

factor	favorable	unfavorable
age	1-9 years	<1 or >9 years
sex	female	male
white blood cell count at diagnosis	low (<50 x 10^9/L)	high (>50 x 10^9/L)
immunophenotype	common ALL, pre-B	pro-B, T-ALL
genotype	hyperdiploidy, *ETV6::RUNX1*$^+$	hypodiploidy, *BCR::ABL1*$^+$, *TCF3::HLF*$^+$, *KMT2A*-rearranged

The aim of induction therapy is the reduction of blast cell counts by application of a combination of chemotherapeutic drugs via systemic and intrathecal therapy. Frequently used drugs in this phase are vincristine, glucocorticoids (such as prednisone or dexamethasone), anthracyclines (such as doxorubicin or daunorubicin) and L-asparaginase [76, 79]. Throughout treatment, CNS prophylaxis is applied, usually by administering intrathecal methotrexate [76]. After complete remission (usually defined as <5% lymphoblasts in the bone marrow [80]), therapy intensity is reduced in the consolidation phase and low-intensity therapy is continued for at least 2 years (=maintenance therapy). In few, very high-risk patients, allogeneic hematopoietic stem cell transplantation is applied [76]. Overall, leukemia treatment is connected to increased risk for long-term health effects, e.g. cognitive impairment, cardiac dysfunction, osteonecrosis and secondary cancers [81].

1.3 *ETV6::RUNX1*

t(12;21)(p13;q22) is the most common chromosomal translocation of pediatric B-ALL, accounting for approximately 20% of cases [82], whereas it is much rarer in adult B-ALL (2% of cases) [83]. The translocation leads to the fusion of the two transcription factors ETV6 and RUNX1, which play important roles during hematopoiesis. Within the last decades, characterization of this leukemia entity has deepened our understanding of cancer evolution. *ETV6::RUNX1*$^{+}$ BCP-ALL evolution involves acquisition of a series of genetic lesions, sometimes characterized by long latency between the first genetic hit and development of overt leukemia [39, 84] (**Figure 1.4**). Altogether, *ETV6::RUNX1*$^{+}$ BCP-ALL is associated with high five-year event-free survival rates of around 90% [73]. Nevertheless, approximately 20% of patients suffer from relapse, particularly late relapses ≥3 years after initial diagnosis [85, 86].

1.3.1 Structure

The translocation t(12;21)(p13;q22) results in the fusion of *ETV6* exons 1-5 and *RUNX1* exons 2-8 or 3-8, and subsequent expression of the chimeric transcription factor ETV6::RUNX1 (**Figure 1.5**). Chromosomal breaks occur along the length of *ETV6* intron 5 (14.47 kb) and *RUNX1* intron 1 (155.88 kb) or 2 (5.83 kb). The ETV6::RUNX1 fusion protein therefore contains almost the complete RUNX1 and the first two thirds of the ETV6 protein.

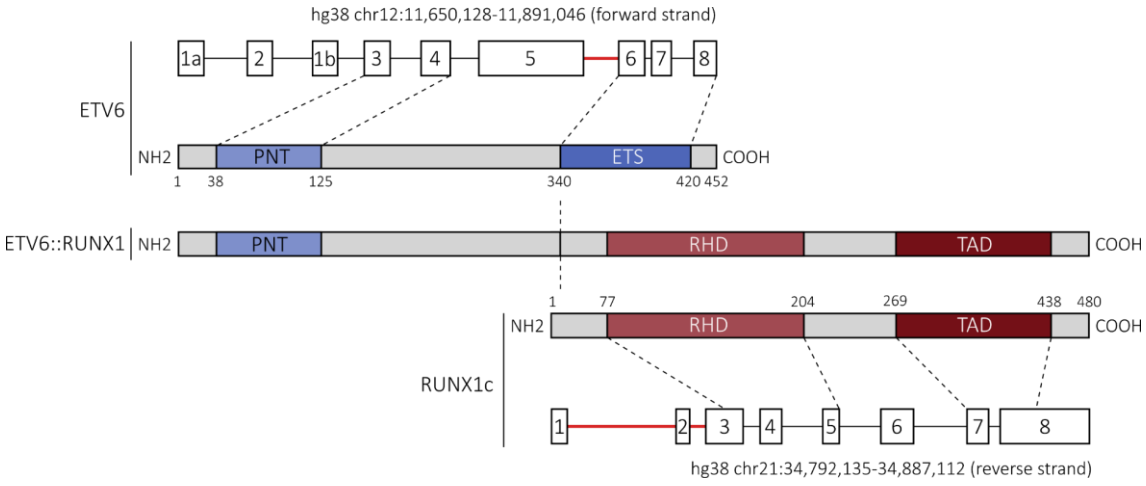


Figure 1.5: Schematic representation of the ETV6::RUNX1 fusion on genome and protein level. Genomic regions and protein domains of *ETV6* and *RUNX1* are depicted. Intronic breakpoint cluster regions are marked as red lines. Exon regions translate to specific protein domains (PNT, ETS, RHD and TAD) indicated by dotted lines. RUNX1 isoform c is shown (transcript: NM_001754.4, protein: NP_001745.2) since isoforms a and b do not contain the breakpoint cluster region. Adapted from [87] and [88]. PNT: pointed domain, ETS: erythroblast transformation specific domain, RHD: Runt homology domain, TAD: transactivation domain.

1.3.2 *ETV6*

The *ETV6* (formerly known as *TEL*) gene encodes for a transcription factor of the ETS (erythroblast transformation specific) family and is ubiquitously expressed [89]. ETV6 consists of a highly conserved ETS DNA-binding domain, a linker domain and an oligomerization domain, i.e. the pointed (PNT) or helix-loop-helix (HLH) domain. ETV6 acts as a transcriptional repressor [90, 91] and can bind to multiple co-repressors, such as L3MBTL1 [92], TRIM28 [93], SIN3A, HDAC3, NCOR1 [94] and KAT5 [95].

Knockout of *ETV6* in mice leads to lethality at embryonic day E10.5 to E11.5 due to defects in blood vessel formation in the yolk sac and apoptosis of mesenchymal and neural cells [89]. While ETV6 is not essential for early hematopoiesis of embryonic stem cells in the yolk sac and fetal liver, it plays a crucial role during hematopoiesis within the bone marrow [96]. In line with this, knockout of ETV6 in HSCs of adult mice leads to a continuous loss of HSCs from the bone marrow [97].

In humans, germline mutations of *ETV6* are associated with ALL predisposition [52]. Studies found that ETV6 haploinsufficiency induces impaired megakaryocyte and platelet formation, leading to thrombocytopenia [53, 54, 98]. Moreover, *ETV6* is involved in a plethora of fusion genes present in leukemia and myelodysplastic syndrome. At least 30 *ETV6* fusion partner genes have been identified so far, including kinases (e.g. *PDGFRB*, *ABL1*) and other transcription factors (e.g. *RUNX1*, *MN1*) [99].

1.3.3 *RUNX1*

RUNX1 (previously known as *AML1* or *CBFA2*) encodes for a transcription factor that is essential for hematopoiesis. RUNX1 contains a DNA-binding Runt homology domain (RHD) and a C-terminal transactivation domain (TAD) which mediates interaction with co-factors. RUNX1 along with the other RUNX proteins (RUNX2 and RUNX3) and the non-DNA-binding CBFβ is part of the core binding factor (CBF) family.

There are three major isoforms of RUNX1 that originate from alternative splicing and activity of either a distal or proximal promoter [100, 101]. Transcription of both RUNX1a (250 amino acids) and RUNX1b (453 amino acids) isoforms is controlled by the distal promoter, while transcription of RUNX1c (483 amino acids) is mediated by the proximal promoter [100]. While RUNX1a is lacking the TAD domain at its C-terminus, RUNX1c differs from RUNX1b only by 32 N-terminal amino acids encoded by an alternative exon. RUNX1a and RUNX1b are expressed throughout hematopoietic development, whereas RUNX1c expression is restricted to definitive HSCs [102].

RUNX1 binds DNA via its RHD at the consensus motif 5'-TGTGGT-3' (or in fewer cases 5'-TGCGGT-3') [103] and forms a transcriptional complex with CBFβ. CBFβ enhances RUNX1 DNA-binding [104] and inhibits ubiquitination as well as subsequent proteasomal degradation of RUNX1 [105]. Depending on its association with other co-factors, RUNX1 may act either as a transcriptional activator or repressor. Interaction partners of the RUNX1-CBFβ complex include ETS1, CEBPA, PU.1, SIN3A and EP300 [106-109].

RUNX1 deficiency in mice leads to embryonic lethality at E11.5 to E12.5 due to necrosis and bleeding in the CNS [110]. Loss of RUNX1 completely abrogates definitive hematopoiesis during mouse embryogenesis [110, 111]. Moreover, RUNX1 haploinsufficiency in mice hinders HSC emergence [112]. In adult hematopoiesis, RUNX1 deficiency leads to impaired B cell, T cell and megakaryocyte differentiation, but does not have an effect on HSC numbers [113]. During early B lymphoid development, RUNX1 is specifically required for the transition to pre-B cells [114]. Altogether, research indicates that RUNX1 regulates the complex process of hematopoietic development in a dose-dependent manner [112, 115].

The essential role of *RUNX1* in hematopoiesis is reflected by its involvement in numerous fusion genes present in AML and ALL. Over 30 chromosomal translocations involving *RUNX1* have been reported so far, most of them including the 5' region of *RUNX1* which contains the RHD (e.g. *RUNX1::RUNX1T1* or *RUNX1::MECOM*) [116, 117]. Germline *RUNX1* mutations lead to familial platelet disorder with associated myeloid malignancy (FPDMM), an autosomal dominant disease that is accompanied by abnormal platelet function and elevated risk to develop hematologic cancers, primarily AML and myelodysplastic syndromes [118, 119]. As is the case with many leukemia predisposing translocations, *RUNX1* mutations alone are not sufficient for leukemia transformation. Additional somatic mutations are required for leukemia onset [120].

1.3.4 Cell of origin

The *ETV6::RUNX1* translocation arises during prenatal development. Depending on the screening method, frequencies of healthy newborns carrying the *ETV6::RUNX1* translocation range from 1% to 5% [45, 84, 121]. Other studies found even lower frequencies (10^{-4}-10^{-5}) [122, 123]. Due to the covert phenotype of *ETV6::RUNX1*$^{+}$ preleukemia, identification of the cell type that initially acquired the fusion gene is challenging. *ETV6::RUNX1*$^{+}$ BCP-ALL blast cells express common markers of B precursor cells (e.g. CD19, CD10, TdT and RAG1/2) and are characterized by ongoing Ig heavy chain rearrangements [14].

RAG1/2 binding sites are absent from the *ETV6::RUNX1* breakpoint cluster region, indicating that *ETV6::RUNX1* arises before the pro-B cell state where RAG1 and RAG2 are expressed [14]. Other studies identified shared partial Ig and TCR rearrangements in *ETV6::RUNX1*$^+$ blast cells of monozygotic twins [43, 124], arguing for a cell of origin with ongoing RAG activity.

Strikingly, while *ETV6::RUNX1*$^+$ ALL is common among pediatric leukemias, only 2% of adult B cell leukemias carry the *ETV6::RUNX1* translocation [83]. This suggests an *ETV6::RUNX1*$^+$ cell of origin that is unique to fetal development or requires a specific microenvironment only present during embryonic development. Using a human induced pluripotent stem cell (hiPSC) model, Böiers *et al*. identified a CD19$^-$IL7R$^+$ cell population present during fetal development that could potentially serve as a candidate for the *ETV6::RUNX1*$^+$ cell of origin [125].

1.3.5 Molecular function

ETV6::RUNX1 acts as an aberrant transcription factor. Since the DNA-binding domain (RHD) of RUNX1 is maintained within the fusion protein, ETV6::RUNX1 competes with RUNX1 for DNA binding sites [126]. It is hypothesized that the N-terminus of ETV6 converts the fusion protein into a repressor. Indeed, several studies could show that ETV6::RUNX1 acts primarily as a transcriptional repressor via recruitment of other co-repressors and therefore antagonizes RUNX1 functions [126-131].

On a molecular level, ETV6::RUNX1 is involved in multiple cellular pathways. Knockdown of ETV6::RUNX1 via small hairpin RNA (shRNA) inhibited PI3K/AKT/mTOR signaling, an essential cellular pathway involved in proliferation and apoptosis [132]. Direct targets of ETV6::RUNX1 identified by chromatin immunoprecipitation (ChIP) assays include the recombination activating gene *RAG1* [133] and the transcription factor *SPIB* that is involved in lymphoid differentiation [134]. Another direct target gene of ETV6::RUNX1 is the erythropoietin receptor *EPOR* that was associated with activation of the JAK/STAT5/BCL-XL signaling cascade, thereby contributing to survival of *ETV6::RUNX1*$^+$ preleukemic cells [135, 136].

While other translocations that include the *RUNX1* gene (such as *RUNX1::RUNX1T1* or *RUNX1::MECOM*) result in myeloid leukemia, *ETV6::RUNX1* is exclusively linked to B lymphoid leukemia. Moreover, expression of ETV6::RUNX1 in myeloid cells led to apoptosis and abrogated myeloid differentiation [137]. This is striking, since aberrant expression of myeloid markers (e.g. CD13, CD33 or CDw65) has been reported for some *ETV6::RUNX1*$^+$ ALL cases [138, 139].

Within the last decades, various *ETV6::RUNX1*$^+$ preleukemia mouse models were generated to analyze ETV6::RUNX1 functions in an *in vivo* setting. These could recapitulate the human disease with varying degrees of success. Schindler *et al*. observed that expression of ETV6::RUNX1 under control of the endogenous *ETV6* promoter leads to malignant transformation in response to chemical mutagenesis in HSCs but not in early lymphoid progenitor cells. When expressed in HSCs, ETV6::RUNX1 induced an increase in quiescent HSC numbers. However, this mouse model exclusively developed leukemias and lymphomas of T cell origin [140]. Another mouse model developed by van der Weyden *et al*. expresses ETV6::RUNX1 driven by the endogenous *ETV6* promoter combined with Sleeping Beauty (SB) transposase expression. This model enabled the identification of cooperating gene mutations introduced by SB, including genes that were identified previously in human *ETV6::RUNX1*$^+$ ALL patients (e.g. *BTG1*, *EBF1*, *EPOR*, *IKZF1*). Moreover, this mouse model was able to develop B cell leukemias, although myeloid and T cell leukemias were still in the majority [141]. An *ETV6::RUNX1*$^+$ mouse model that exclusively developed precursor B-ALL was generated by Rodríguez-Hernández *et al*. by means of expressing ETV6::RUNX1 under the control of the *Sca-1* (*Ly6a*) promoter, thereby restricting expression of the fusion gene to HSCs. Mice showed low leukemia penetrance of about 10% only when exposed to common pathogens, whereas malignant transformation was not observed in a pathogen-free environment [142].

Given that murine models were largely unable to reproduce human B lineage leukemia, Böiers *et al*. used hiPSCs to recapitulate *ETV6::RUNX1*$^+$ preleukemia in a human system. In their model, expression of ETV6::RUNX1 under control of the endogenous *ETV6* promoter led to a partial block of early B lymphopoiesis and co-expression of myeloid genes [125]. A follow-up study using the same hiPSC model could demonstrate an inhibitory effect of ETV6::RUNX1 on cell cycle regulation and suggested targeting CBFβ to improve *ETV6::RUNX1*$^+$ BCP-ALL treatment [126].

Altogether, ETV6::RUNX1 reprograms cells via aberrant transcriptional regulation that induces cell cycle inhibition, altered cell survival and a characteristic developmental block during B cell development. Ultimately, these alterations could create a cellular environment prone to the accumulation of further genetic mutations, e.g. due to aberrant RAG activity.

1.3.6 Secondary mutations

While *ETV6::RUNX1* is a frequent first hit mutation, further mutations are required for leukemia transformation. Hence, only a small fraction (1% or less) of *ETV6::RUNX1* carriers will develop

the actual disease. Due to the weak oncogenic potential of the fusion gene, *ETV6::RUNX1*$^+$ BCP-ALL can arise after long latency periods of up to 14 years [39, 84] and has a low concordance rate of about 10% in monozygotic twins [41].

Secondary mutations of *ETV6::RUNX1*$^+$ BCP-ALL patients comprise mostly copy number alterations (CNAs), specifically deletions of genes important for B lymphocyte differentiation (*PAX5*, *EBF1*, *RAG1*, *RAG2*), immune system regulation (*BTLA*), cell cycle (*CDKN2A*, *CDKN2B*) and hormone response (*TBL1XR1*) [14, 143]. Deletion of the wild-type *ETV6* allele occurs in 59% of *ETV6::RUNX1*$^+$ ALL cases [143], while a gain of the *RUNX1* gene is observed in around 30% of patients [144]. Indeed, loss of the wild-type *RUNX1* gene is rare in *ETV6::RUNX1*$^+$ ALL patients and dependency of *ETV6::RUNX1*$^+$ ALL blast cells on RUNX1 expression was suggested recently [126]. Subclonality of secondary mutations has been demonstrated in twin studies that show identical breakpoint sequences of *ETV6::RUNX1*, but clonally different additional lesions [145, 146]. Altogether, *ETV6::RUNX1*$^+$ ALL patients carry 3.5 secondary mutations on average [143].

RAG-mediated recombination has been proposed as a major mechanism for the accumulation of second hit mutations in *ETV6::RUNX1*$^+$ BCP-ALL. Papaemmanuil *et al*. detected V(D)J RSS motifs specifically at breakpoint regions of recurrent CNAs identified in *ETV6::RUNX1*$^+$ BCP-ALL patients [14]. Moreover, RAG1 expression was shown to be directly upregulated by ETV6::RUNX1 [133].

1.4 Leukemia epigenome

Apart from mutational changes on DNA level, regulation of the epigenome plays an integral role during leukemogenesis and chemoresistance. Epigenetic regulation includes processes that do not alter DNA sequences, such as the spatial organization of nuclear chromatin, DNA methylation and chemical modification of histone tails, e.g. methylation or acetylation (see Figure 1.6).

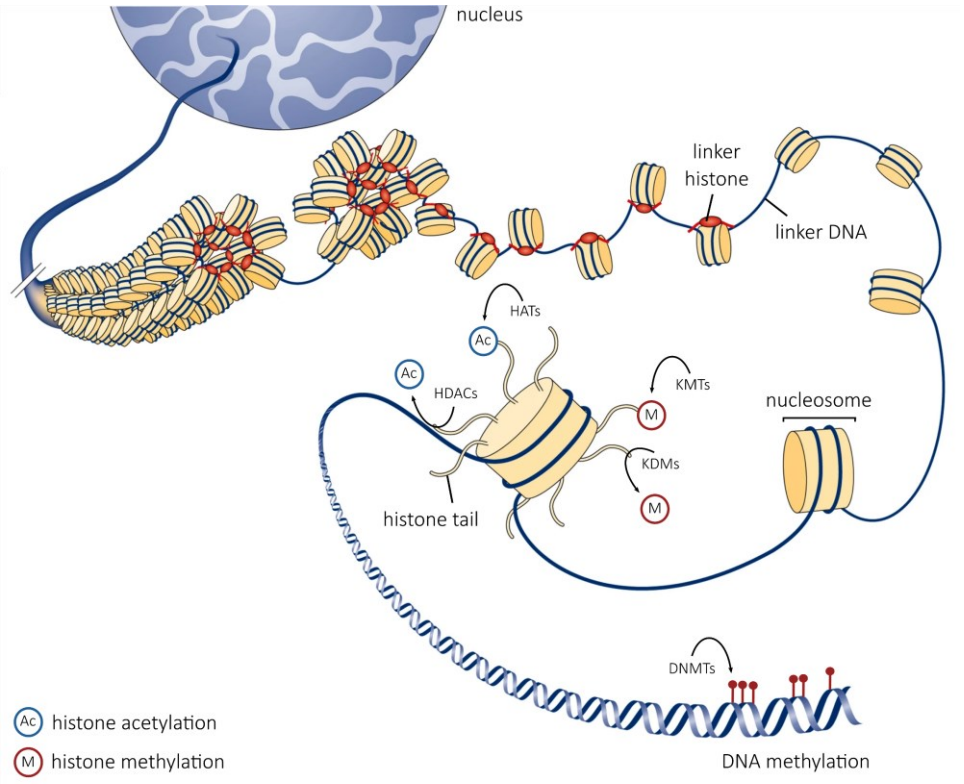


Figure 1.6: Schematic overview of epigenetic mechanisms involved in chromatin compaction. DNA is wrapped around histones, forming nucleosomes. DNA compaction is mediated by different mechanisms, including linker histone binding and chemical modification of DNA or histones. Common modifications include methylation and acetylation. Protein groups involved in addition or removal of chemical groups are indicated in the illustration. Figure was modified from [147]. HDACs: histone deacetylases, HATs: histone acetyltransferases, KMTs: histone lysine methyltransferases, KDMs: histone lysine demethylases, DNMTs: DNA methyltransferases.

Mutations in epigenetic regulators are frequently observed in pediatric cancers, including BCP-ALL [148]. A study by Huether *et al*. identified at least one mutation in an epigenetic gene in 34% of *ETV6::RUNX1*$^{+}$ BCP-ALL samples [148]. Recurrent mutations or translocations at diagnosis in BCP-ALL patients affect histone methyltransferase genes, such as *NSD2* [149, 150] and *KTM2A* [151]. Mutations in epigenetic regulators are also enriched in relapsed B-ALL,

potentially contributing to chemotherapy resistance. Mar *et al*. detected mutations of epigenetic modifiers in 57% of relapsed B-ALL patients (affecting the genes *SETD2*, *CREBBP*, *MSH6*, *KDM6A* and *KMT2D*) [152]. In line with this, Mullighan *et al*. and others could identify mutations of the acetyltransferase gene *CREBBP* in 18.3% of relapsed B-ALL patients [153, 154].

By far the most studied epigenetic modification in B-ALL patients is DNA methylation. Regions of the genome that contain CpG dinucleotide repeats (so called CpG islands (CGIs)) are associated with promoter sequences [155] and methylation at these CGIs is inversely correlated with gene expression [156]. By analyzing global methylation patterns, several studies could show that DNA methylation signatures differ according to B-ALL subtype [156-161] and could potentially be used as biomarkers to complement ALL diagnosis [156, 162].

Apart from DNA methylome alterations, the tails of histone proteins can be post-translationally modified to regulate gene expression. Global changes of histone acetylation were detected in ALL blasts [163, 164]. In line with this, human and murine *ETV6::RUNX1*$^{+}$ B-ALL exhibit recurrent alterations in lysine deacetylase genes, such as *KDM5C*, which are associated with increased H3K4me3 levels [142]. Using an assay for transposase-accessible chromatin combined with high throughput sequencing (ATAC-seq) approach, Barnett *et al*. observed that overall chromatin accessibility differs according to B-ALL subtype, with *ETV6::RUNX1*$^{+}$ BCP-ALL exhibiting lowest number of accessible chromatin sites [165].

1.4.1 Linker histones

Human chromatin is structured into nucleosomes which contain 146 bp of DNA wrapped around a histone octamer made up of the four core histones H2A, H2B, H3 and H4 [166]. Additionally to the core histones, DNA is further compacted by binding of H1 linker histones to the DNA that enters and exits the nucleosome (see **Figure 1.7A**) [167]. In humans, there are 11 linker histone variants with different tissue-restricted expression. These include the somatic variants H1-0 to H1-5 and H1-10, the testis-restricted H1-6, H1-7 and H1-9, as well as the oocyte-specific linker histone H1-8 [147]. Structurally, linker histones contain a conserved, nucleosome-binding central globular domain, while N- and C-terminal domains are less conserved (see **Figure 1.7B**) [168]. Interaction of linker histones with the DNA phosphates is conveyed by positively charged arginine (R) and lysine (K) residues [167].

Linker histones function as epigenetic regulators by interacting with a multitude of epigenetic modifiers, and thereby affect post-translational modification of core histones, DNA replication and DNA damage repair [147]. It was proposed that linker histones sterically hinder access of

histone acetyltransferases, such as KAT2B (also referred to as P300/CBP-associated factor), to nucleosomes [169]. Similarly, linker histones contribute to transcriptional inhibition by mediating RNA polymerase II pausing at the linker DNA region that enters the nucleosome [170]. Moreover, linker histones contribute to silencing of repetitive elements by interacting with histone lysine methyltransferases SUV39H1, SUV39H2 and SETDB1 [171]. In general, linker histones are considered to have a repressive effect on chromatin [171-173].

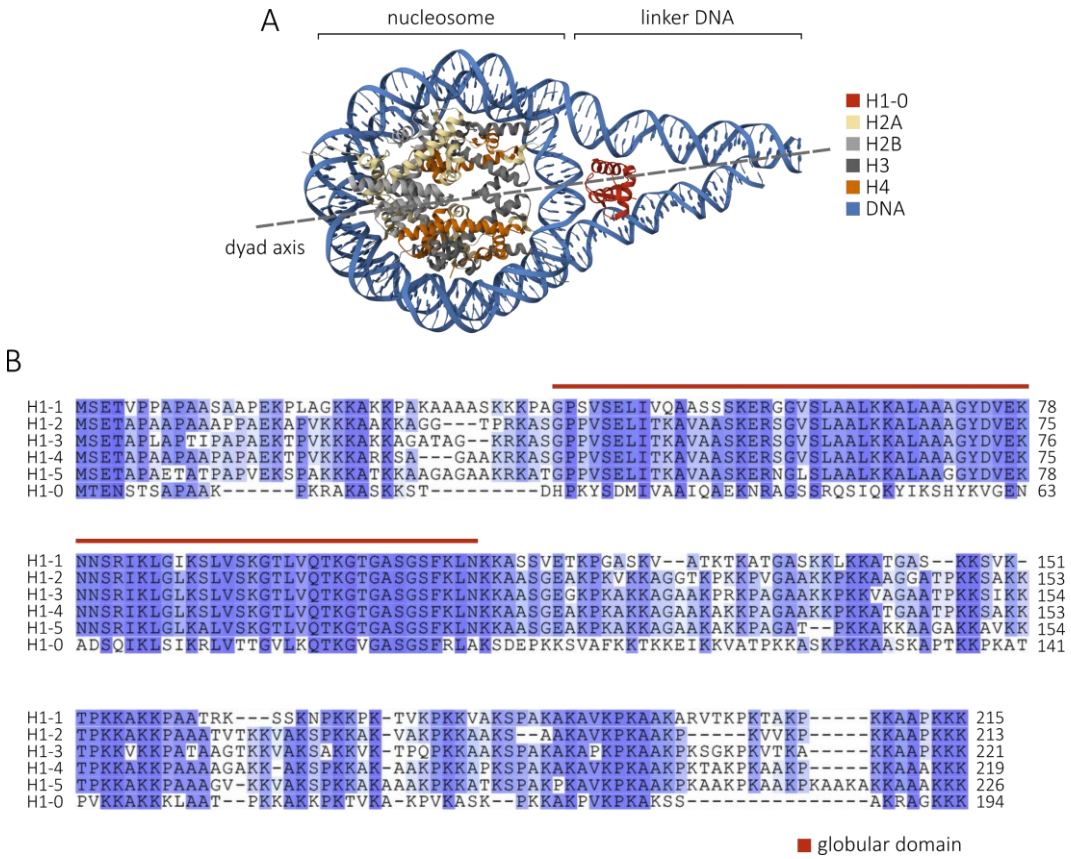


Figure 1.7: Spatial organization of histone proteins within human chromatin and multiple sequence alignment of human somatic linker histones H1-0 to H1-5. (A) Cryogenic electron microscopy structure of a human nucleosome containing linker histone H1-0 (Protein Data Bank identifier: 7k5x [174, 175]), forming the so-called chromatosome. The 3D molecule structure was visualized using the Mol* 3D Viewer web application [176]. **(B)** Alignment of somatic linker histone variants H1-0 to H1-5. The highly conserved globular domain is highlighted in red and the individual amino acids are colored according to degree of conservation (dark blue color marks higher conservation). Adapted from [177].

Although linker histones are highly conserved, the individual linker histone variants show functional differences. Chromatin binding affinities differ between variants: H1-1 and H1-2 have weak chromatin condensation properties, whereas H1-0, H1-4, H1-5 and H1-10 are strong

condensers [178]. Moreover, the different linker histone variants show distinct DNA binding patterns [179] and differences in expression levels during embryogenesis [180].

1.4.2 Linker histone H1-0

While the core histones and somatic linker histone variants *H1-1* to *H1-5* are encoded on chromosome 6, the *H1-0* gene is located on chromosome 22. H1-0 is the shortest linker histone variant, containing only 194 amino acids. While histone synthesis is generally coupled with DNA synthesis, H1-0 expression occurs independently of DNA replication [181, 182] and was found to be enriched at GC-rich genes [183]. In solid tissues, H1-0 accumulates in slowly or non-dividing cells [184], and was therefore detected in many terminally differentiated cell types [185, 186]. Consistently, induction of cell proliferation, for instance during rat liver regeneration, leads to a decrease of H1-0 levels [185]. In human hematopoietic cells, H1-0 is highly expressed in quiescent, immature cell types that have the potential to be reactivated [187]. In line with this, research by Morales Torres *et al*. showed that H1-0-induced expression changes are reversible [183].

Mice lacking H1-0 develop normally but show a partial differentiation block of dendritic cells [188, 189]. The absence of a severe phenotype upon H1-0 knockout in mice suggests that other linker histone variants can compensate for H1-0 loss. This is supported by the observation that only the compound knockout of at least three linker histone variants induces major defects in mice, such as slow growth, a small thymus, low birth rate and neonatal or embryonic lethality [190]. Hence, there appears to be a critical total level of H1 linker histone expression that is necessary for correct embryonic development (see [191] for a comprehensive summary of linker histone single and compound knockout mouse models).

Downregulation of H1-0 correlates with reduced patient survival in various cancer entities [183]. However, the expression profile of linker histone H1-0 depends on the specific cancer subtype. Within solid tumors, H1-0 is heterogeneously expressed, marking cells of different differentiation states and proliferative potential [183]. In leukemic blasts, H1-0 levels were found to be upregulated compared to healthy lymphocytes in a cohort of eight patients [192].

H1-0 expression can be induced by *in vitro* treatment with the histone deacetylase inhibitor (HDACi) sodium butyrate, a compound that induces cell cycle arrest at the G0/G1 cell cycle phase [193]. In line with this, other HDACis, such as Quisinostat, Vorinostat, Abexinostat and Trichostatin A, are able to restore high H1-0 levels due to their cytostatic effect [194].

2. AIMS OF THE THESIS

The *ETV6::RUNX1* fusion gene is a frequent first hit mutation that can be detected in up to 5% of healthy newborns [45]. Of these *ETV6::RUNX1* carriers, only a small fraction (<1%) will go on to develop leukemia. The subtle effects of the *ETV6::RUNX1* fusion gene make modeling *ETV6::RUNX1*$^+$ preleukemia challenging and several preleukemic mouse models failed to reproduce the exclusive association with B lineage leukemia that is observed in human *ETV6::RUNX1*$^+$ BCP-ALL [140, 141]. Recently, promising results have been achieved by expressing the fusion gene at physiological levels from the *ETV6* locus in an hiPSC model [125]. Because of the pluripotent nature of hiPSCs, modeling in these cells allows to recapitulate early B cell development.

While BCP-ALL treatment outcomes have considerably improved over the last decades, chemotherapy remains highly toxic and is associated with severe side effects. Moreover, disease relapse of *ETV6::RUNX1*$^+$ BCP-ALL occurs in 20% of patients [85, 86]. Therefore, further improvement of therapy efficacy and reduction of toxicity are necessary.

The main aim of this work is to identify and further characterize regulators that are involved in *ETV6::RUNX1*-mediated leukemogenesis and might offer novel therapeutic strategies to improve treatment.

Major aims of this work:

1. Generation of preleukemic *ETV6::RUNX1*$^+$ hiPSC lines using a CRISPR/Cas9-mediated knock-in approach.
2. Transcriptome analysis of *ETV6::RUNX1*$^+$ hiPSCs and hematopoietic progenitor cells to detect genes dysregulated by ETV6::RUNX1.
3. Functional analysis of H1-0 in the context of *ETV6::RUNX1*$^+$ leukemia via luciferase promoter assays and siRNA-mediated knockdown in the *ETV6::RUNX1*$^+$ REH cell line.
4. Analysis of vulnerability of *ETV6::RUNX1*$^+$ BCP-ALL towards H1-0-inducing drugs.

3. MATERIALS & METHODS

3.1 Molecular biology

3.1.1 Molecular cloning of a *RUNX1* homology-directed repair template

A *RUNX1* homology-directed repair (HDR) template was constructed by combining exons 2-8 of the *RUNX1* gene with a puromycin resistance gene (puromycin N-acetyltransferase) under control of the human *EF-1α* promoter. Sequences homologous to the insertion site located within *ETV6* intron 5 were PCR-amplified from ChiPSC12 (#Y00280, Takara Bio) genomic DNA. Restriction sites were added to the PCR products using PCR-based cloning (Table 3.1 and Figure 3.1A). PCR fragments were cut with the respective restriction enzymes (see chapter 3.1.4) and fragments were ligated using T4 DNA ligase (#M0202, New England Biolabs, see chapter 3.1.6). Nucleic acid concentrations and purity (A_{260}/A_{280} ratio) was determined by using the Nanodrop ND-1000 (PEQLAB). A schematic representation of the *RUNX1* HDR template is depicted in Figure 3.1B. Primer sequences were purchased from Eurofins Genomics. See Table 3.2 for a list of primer sequences used for molecular cloning of the *RUNX1* HDR template.

Table 3.1: Pipetting scheme and thermocycling conditions used for amplification of HDR template sequences and simultaneous addition of restriction sites. Annealing temperatures of PCRs varied according to primer melting temperature.

	volume	amount
2X Phusion High-Fidelity PCR Master Mix (#F531S, New England Biolabs)	25 µl	1x
fw primer (10 µM)	2.5 µl	0.5 µM
rev primer (10 µM)	2.5 µl	0.5 µM
template DNA	X µl	20-250 ng
nuclease-free H_2O	to 50 µl	

98 °C	———	5 min	
98 °C	———	30 s	35 cycles
65-72 °C	———	30 s	
72 °C	———	30 s/kb	
72 °C	———	4-7 min	
4 °C	———	∞	

The *RUNX1* HDR template was subcloned into the pUC19 plasmid at the EcoRI and KpnI restriction sites (pUC19 was a gift from Joachim Messing, addgene plasmid #50005) and linearized by PCR amplification with primers binding the 5' and 3' end of the insert (see Table 3.2 and 3.3). PCR amplification was performed using the Guide-it long ssDNA production system (#632644, Takara Bio). The linear double-stranded DNA product was column-purified using the QIAquick PCR purification kit (#28104, QIAGEN) and concentrated by isopropanol precipitation. For this, an equal volume of isopropanol as well as 1/10th the volume 3 M sodium acetate (pH 5.2) were added to the eluate. The mixture was thoroughly vortexed and incubated at -20 °C overnight. After centrifugation at 14,000 g for 10 min and 4 °C, the supernatant was

removed carefully and the DNA pellet was air-dried for 10-15 min. The pellet was then dissolved in the appropriate amount of nuclease-free water to obtain a DNA concentration of ≥1 µg/µl. The complete sequence of the *RUNX1* HDR template is depicted in the Appendix (chapter 8.3).

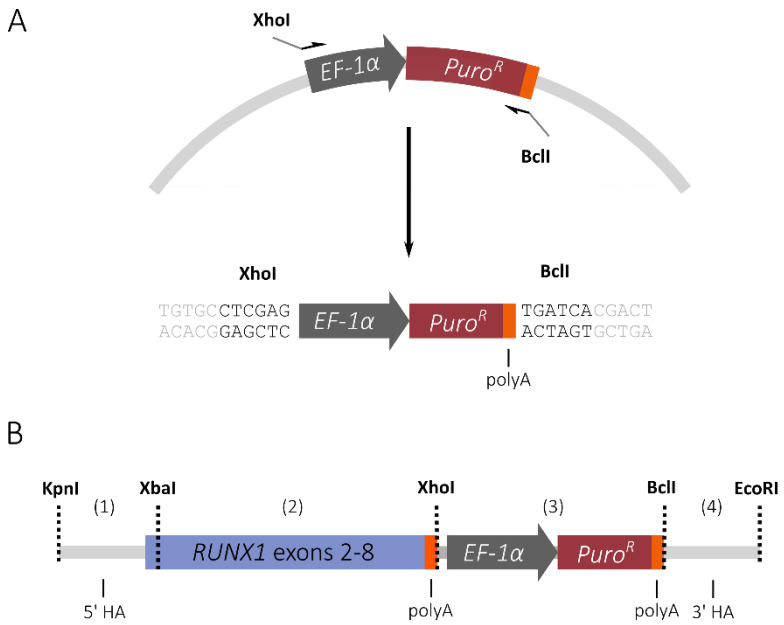


Figure 3.1: Molecular cloning of a *RUNX1* homology-directed repair (HDR) template targeting *ETV6* intron 5. (A) PCR-based cloning was used to add restriction sites, as shown for the *EF-1α* promoter and puromycin resistance gene sequences. **(B)** Schematic representation of the *RUNX1* HDR template. Restriction sites for KpnI, XbaI, XhoI, BclI and EcoRI were added as indicated. HA: homology arm.

Table 3.2: List of primer sequences used for molecular cloning of the *RUNX1* HDR template. Restriction enzyme sites are marked in red. HA: homology arm, fw: forward, rev: reverse, HDRT: homology-directed repair template, lin: linearized.

name	sequence [5' → 3']	amplicon length
KpnI_5'HA_fw XbaI_5'HA_rev	AGTCGGGTACCGCAGCGGCCCCTCCGGTCC GATCGTCTAGAAGGATTCATTCCAAGTATGCATTCTGCTAT TCTCCCAATGGGCATGGC	464 bp
XhoI_RUNX1_rev	TTGAACTCGAGGCACATAAATAGCAATAATAGTGAAAAAGA ATAACATTGACCATTTATTTCAGTAGGGCCTCCACACGGCC TCCTCCAGGCGCGCGG	used with KpnI_5'HA_fw: 1875 bp
XhoI_EF1α_fw BclI_polyA_puro_rev	TGTGCCTCGAGTTCAAAATTTTATCGATACTAG AGTCGTGATCAAAAAAACCTCCCACACCTCCCCCTG	1448 bp
BclI_3'HA_fw EcoRI_3'HA_rev	AGTCGTGATCAGATGGAGTAGTTAATGAGCCTCAGAAATG AGTCGGAATTCTAACCTTGTCACATGGTTAGAGGATTAAGC	522 bp
RUNX1_HDRT_lin_fw RUNX1_HDRT_lin_rev	GCAGCGGCCCCTCCGGTCC TAACCTTGTCACATGGTTAGAGGATTAAGC	3791 bp

Table 3.3: Pipetting scheme and thermocycling conditions for *RUNX1* HDR template linearization.

	volume	amount
2X PrimeSTAR Max Premix (Takara Bio)	50 µl	1x
fw primer (10 µM)	4 µl	0.8 µM
rev primer (10 µM)	4 µl	0.8 µM
template DNA	X µl	20 ng
nuclease-free H_2O	to 100 µl	

98 °C	10 s	40 cycles
55 °C	5 s	
72 °C	15 s (5 s/kb)	
4 °C	∞	

3.1.2 Molecular cloning of sgRNA plasmids

Single guide RNAs (sgRNAs) targeting the 5' region of *ETV6* intron 5 were designed using the online tools CRISPOR (http://crispor.tefor.net/ [195]). Two sgRNAs with high predicted on-target efficiency and low number of off-target sites were chosen for further experiments (Table 3.4).

Table 3.4: sgRNA sequences used in this work.

name	sequence [5' → 3']	position in human genome (hg38)
ETV6_sgRNA_fw	GGATGAGGCTAAATCCCTAA	chr12: 11,870,115-11,870,134, + strand
ETV6_sgRNA_rev	GCCTAATTGGGAATGGTGCG	chr12: 11870054-11870073, - strand

sgRNAs were cloned by hybridization of forward and reverse oligonucleotide sequences creating specific BbsI 5' and 3' overhangs (see Figure 3.2). For this, forward and reverse oligos were dissolved in TE buffer at 100 pmol/µl. 500 pmol of each oligonucleotide and 1.25 µl of a 4 M NaCl stock solution were mixed, and filled to 100 µl with TE buffer. Hybridization was performed overnight by heating the oligonucleotide mixture to 98 °C and slowly cooling to room temperature (RT) in a water bath. The hybridized double-stranded DNA fragments were ligated at the BbsI restriction sites with a pUC19-U6-BbsI-sgRNA vector using T4 DNA ligase (#M0202, New England Biolabs).

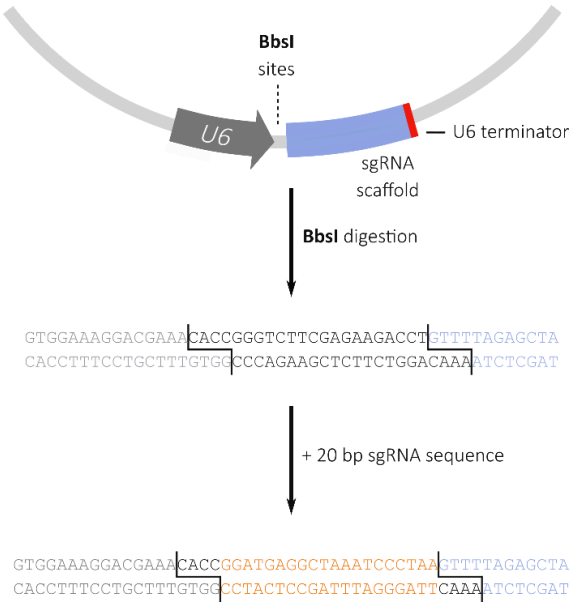


Figure 3.2: Ligation scheme for construction of pUC19-U6-BbsI-sgRNA plasmids. Forward and reverse oligonucleotide sequences were hybridized and ligated with a pUC19-U6-BbsI-sgRNA vector (cut with BbsI). U6 is a polymerase III promoter used for transcription of non-coding RNA molecules requiring only a defined termination site (4-5 thymidines). sgRNA sequences must contain a guanine as the first transcribed nucleotide to ensure efficient transcription.

3.1.3 Molecular cloning of a *H1-0* promoter fragment

Human *H1-0* promoter sequence (nucleotides -351 to +161 from TSS) was PCR-amplified from REH (#ACC 22, DSMZ) genomic DNA (**Tables 3.5** and **3.6**). The amplified product was purified using the QIAquick PCR Purification kit (#28104, QIAGEN) and inserted into Firefly luciferase vector pGL4.22 (#E6771, Promega) at KpnI and HindIII restriction sites.

Table 3.5: Pipetting scheme and thermocycling conditions for PCR-amplification of the *H1-0* promoter.

	volume	amount
5X PrimeSTAR GXL buffer (#R050A, Takara Bio)	10 µl	1x
dNTP mixture (2.5 mM each)	4 µl	200 µM each
H1-0_fw (10 µM)	1 µl	0.2 µM
H1-0_rev (10 µM)	1 µl	0.2 µM
template DNA (genomic DNA)	X µl	100 ng
PrimeSTAR GXL DNA polymerase	1 µl	1.25 U
nuclease-free H_2O	to 50 µl	

98 °C	10 s	40 cycles
68 °C	1 min	

Table 3.6: Primer sequences used for molecular cloning of the *H1-0* promoter region. Restriction enzyme sites are marked in red.

name	sequence [5' → 3']	amplicon length
KpnI_H1-0_fw	AGTCGGGTACCGCCAGGCAGCGGCCCAG	534 bp
HindIII_H1-0_rev	AGTCGAAGCTTGGTGGCCTGTCTGGTCCG	

3.1.4 Restriction enzyme digest

Endonuclease restriction was performed for 1 h according to the manufacturer's instructions (see Table 3.7 for a list of restriction enzymes used in this work). Digestions were performed using the CutSmart buffer (#B6004SVIAL, New England Biolabs). In order to determine optimal reaction parameters, the online tool NEBcloner (https://nebcloner.neb.com, New England Biolabs) was used.

Table 3.7: Restriction enzymes used for molecular cloning.

enzyme	supplier
KpnI-HF	#R3142S, New England Biolabs
XbaI	#R0145S, New England Biolabs
XhoI	#R0146S, New England Biolabs
BclI-HF	#R3160S, New England Biolabs
EcoRI-HF	#R3101S, New England Biolabs
BbsI-HF	#R3539S, New England Biolabs
XmaI	# R0180S, New England Biolabs
AgeI-HF	#R3552S, New England Biolabs
HindIII-HF	#R3104S, New England Biolabs

3.1.5 Agarose gel electrophoresis and gel extraction

DNA fragments were separated according to size by agarose gel electrophoresis. DNA fragments were run on 1% (w/v) agarose gels at 100 V for 30-40 minutes. The intercalating agent Midori Green (#MG04, Nippon Genetics) was used to stain DNA. If required, DNA bands were cut out using a scalpel and DNA was extracted using the QIAquick Gel Extraction kit (#28704, QIAGEN) according to the manufacturer's instructions.

3.1.6 DNA ligation

Ligation of DNA fragments was performed using T4 ligase (#M0202, New England Biolabs). Reaction components were mixed in a microcentrifugation tube on ice (see Table 3.8) and

incubated at 16 °C overnight. The mixture was then heat-inactivated at 65 °C for 10 minutes and used for transformation of chemically competent bacteria. Insert DNA fragments and vector DNA were combined in a molar ratio of 1:1 to 7:1 (insert:vector, 100-150 ng DNA total) according to the NEBioCalculator (https://nebiocalculator.neb.com, New England Biolabs).

Table 3.8: Pipetting scheme for DNA ligation using T4 DNA ligase (#M0202, New England Biolabs).

	volume	
T4 DNA ligase buffer (10x)	2 µl	
vector DNA	X µl	1:1 to 7:1 molar ratio
insert DNA	X µl	(100-150 ng total)
T4 DNA ligase	1 µl	
nuclease-free H_2O	to 20 µl	

For cloning of the *RUNX1* HDR template, fragment (1) was first ligated to fragment (2) and fragment (3) was ligated to fragment (4) (see Figure 3.1B). The resulting ligation products were PCR-amplified, extracted from an agarose gel, cut with restriction enzymes KpnI and EcoRI, and ligated with the pUC19 plasmid at the KpnI and EcoRI restriction sites in a ratio of 1:1:1.

3.1.7 Transformation of bacteria

Chemically competent XL10-Gold Ultracompetent *E. coli* bacteria (50 µl, #200314, Agilent) were shortly thawed on ice and mixed with 10-100 ng of plasmid DNA or 3 µl ligation mixture by gently swirling. The mixture was incubated for 30 minutes on ice, heat-shocked for 30 seconds at 42 °C using a heating block and immediately put back on ice for 2 minutes. 250 µl pre-warmed LB medium (#X964.4, Carl Roth) were added to the bacteria and grown at 37 °C for 1 h in a shaking incubator. 25 µl of the transformation mixture bacteria were then plated onto a 10 cm LB agar (#X965.2, Carl Roth) plate containing 100 µg/ml ampicillin (#A9518-5G, Merck) and incubated at 37 °C overnight.

3.1.8 Preparation of chemically competent bacteria

15 ml LB medium was inoculated with one vial of XL10-Gold Ultracompetent *E. coli* (#200314, Agilent) and incubated in a shaking incubator at 37 °C overnight. The next day, the optical density at 600 nm (OD_{600}) of the bacteria culture was measured with a spectrophotometer (Eppendorf). At an OD_{600} of 0.1, the liquid culture was filled to 200 ml with LB medium and

incubated at 37 °C in a shaking incubator until it reached an OD_{600} of 0.5-0.6. The bacteria were pelleted at 6,600 rpm for 5 min and 4 °C, the supernatant was discarded and the pellet was dissolved in 17 ml B1 buffer (Table 3.9). After an incubation of 5 h on ice, the cell suspension was centrifuged again as described before and the pellet was resuspended in 4 ml B2 buffer (Table 3.10). The cell suspension was incubated on ice for 15-30 minutes, split into 50 µl aliquots, shock-frozen in liquid nitrogen and stored at -80 °C until further use.

Table 3.9: Components of buffer B1 for generation of chemically competent cells. The buffer was adjusted to pH 5.8 with 0.2 M acetic acid and sterile-filtered.

reagent	amount	supplier
rubidium chloride (RbCl)	100 mM	#R2252, Sigma-Aldrich
manganese(II) chloride ($MnCl_2$)	50 mM	#105934, Merck
potassium acetate (Kac)	30 mM	#104830, Merck
calcium chloride ($CaCl_2$)	10 mM	#2382, Merck
glycerol (85%)	15%	#3783, Carl Roth

Table 3.10: Components of buffer B2 for generation of chemically competent cells. The buffer was adjusted to pH 6.8 with 0.5 M sodium hydroxide (NaOH) and sterile-filtered.

reagent	amount	supplier
3-morpholinopropanesulfonic acid (MOPS)	10 mM	#A1077, AppliChem
rubidium chloride (RbCl)	10 mM	#R2252, Sigma-Aldrich
calcium chloride ($CaCl_2$)	75 mM	#2382, Merck
glycerol (85%)	15%	#3783, Carl Roth

3.1.9 DNA preparation

A single XL10-Gold Ultracompetent *E. coli* (#200314, Agilent) colony was used to inoculate either 5 ml (miniprep) or 250-300 ml (maxiprep) of LB medium. Liquid cultures were grown at 37 °C and 150-200 rpm in 15 ml Falcon tubes or Erlenmeyer flasks. After 16-18 h, cells were pelleted and DNA was prepared using either the Monarch Plasmid DNA Miniprep kit (#T1010L, New England Biolabs) or Plasmid Maxi kit (#12163, QIAGEN).

3.1.10 RT-qPCR

RNA was isolated using the RNeasy Mini kit (#74106, QIAGEN) with an on-column DNA removal step via RNase-free DNase (#79254, QIAGEN). For RT-qPCR detection of H1-0, an additional DNase digest was performed with the TURBO DNA-free kit (#AM1907, Thermo Fisher Scientific). Total RNA concentration was determined using the Nanodrop ND-1000 spectrophotometer and

1-2 µg RNA were reverse transcribed to cDNA using M-MLV reverse transcriptase (#M1701, Promega) as shown in Table 3.11. Due to limited amount of RNA following siRNA-mediated H1-0 knockdown in REH cells, 200 ng of total RNA were reverse transcribed.

Table 3.11: Pipetting scheme and protocol for reverse transcription.

	volume	amount
RNA	X µl	0.2-2 µg
Oligo(dT)18 primer (50 µM, Eurofins Genomics)	0.93 µl	250 ng
random primers (#C118A, Promega)	0.5 µl	250 ng
nuclease-free H_2O	to 17.125 µl	

- incubate for 5 minutes at 70 °C (to remove secondary RNA structures)
- immediately cool on ice for 5 minutes
- centrifuge briefly and keep on ice

	volume	amount
M-MLV reaction buffer (5x, #M531A, Promega)	5 µl	1x
dNTPs (10 mM each, #U151A, Promega)	1.25 µl	0.5 µM
RNase inhibitor (#N251A, Promega)	0.625 µl	25 U
M-MLV reverse transcriptase (#M170A, Promega)	1 µl	200 U
total volume	25 µl	

- incubate at 37 °C for 1 h
- dilute with nuclease-free H_2O and store at -20 °C until further use

RT-qPCR was performed in triplicates using SYBR Green PCR Mix (#4309155, Thermo Fisher Scientific) or TaqMan Universal PCR Master Mix (#4304437, Thermo Fisher Scientific) according to the manufacturer's instructions. No-template controls (NTC) and no-reverse-transcriptase controls (NRT) were included on each plate. Fluorescence was measured with the CFX384 Touch Real-Time PCR Detection System (Bio-Rad). Specific RT-qPCR protocols for detection of *ETV6::RUNX1* and *ETV6* are described in chapter 3.3.3. RT-qPCR primer and probe sequences used in this work are listed in Table 3.12. Efficacy and specificity of primers used in this work was tested by cDNA dilution, assessment of melt curves and gel electrophoresis of amplified PCR fragments (Supplementary Figure 8.1).

Table 3.12: Primer and probe sequences used for RT-qPCR detection of transcript levels.

name	sequence [5' → 3']	amplicon length
H1-0_qPCR_fw	CTGGCTGCCACGCCCAAGAA	124 bp
H1-0_qPCR_rev	GGAGGGAGTGTCCGCAAG	
ETV6::RUNX1	ipsogen ETV6::RUNX1 Kit (#675113, Qiagen)	-
ETV6_qPCR_fw	GAAGAGCACGCCATGCCCAT	155 bp
ETV6_qPCR_rev	GCCAGTCCGTTGGGATCCA	
ETV6_ qPCR_probe	[FAM]ACAGCCGGTACGAAAACTTCATCCGATG[Eclipse]	
GAPDH	Hs_GAPDH_1_SG QuantiTect Primer assay (#QT00079247, Qiagen)	-
ATP5PB_ qPCR_fw	CCACACCTTGTCCCTGTACC	190 bp
ATP5PB_ qPCR_rev	CAGTGAAGGTCTCTGCGCTAA	
PGK1	Applied Biosystems Human PGK1 Endogenous Control (FAM/MGB probe, #4333765T, Thermo Fisher Scientific)	160 bp
DNMT3B_ qPCR_fw	GCTCACAGGGCCCGATACTT	93 bp
DNMT3B_ qPCR_rev	GCAGTCCTGCAGCTCGAGTTTA	
GDF3_ qPCR_fw	AGACTTATGCTACGTAAAGGAGCT	149 bp
GDF3_ qPCR_rev	CTTTGATGGCAGACAGGTTAAAGTA	
POU5F1_ qPCR_fw	GAAGGAGAAGCTGGAGCAAA	94 bp
POU5F1_ qPCR_rev	CTTCTGCTTCAGGAGCTTG	
POU5F1_ qPCR_probe	[FAM]AGGACATCAAAGCTCTGCAGAAA[Eclipse]	
NANOG_ qPCR_fw	CCTGTGATTTGTGGGCCTG	124 bp
NANOG_ qPCR_rev	CAGGGCTGTCCTGAATAAGC	
NANOG_ qPCR_probe	[FAM]ATGCCTCACACGGAGACTGTCT[Eclipse]	

3.1.11 Dual-luciferase promoter assay

293T cells at 50-70% confluency were transfected with 755 ng plasmid DNA using Xfect Transfection Reagent (#631317, Clontech Laboratories) in 24-well plates according to the manufacturer's instructions. Each well was transfected with 500 ng pGL4.22 vector (#E6771, Promega,) with or without *H1-0* promoter expression as well as 5 ng Renilla luciferase control plasmid pGL4.73 (#E6911, Promega), and 250 ng of the respective pcDNA3.1 vectors (#V79020, Thermo Fisher Scientific) for expression of *ETV6::RUNX1* or *RUNX1* or empty vector in triplicates. Cells were lysed after 48 h with Passive Lysis buffer and luciferase signal was measured on a Tecan SPARK 10M reader using the Dual-Luciferase Reporter Assay System (#E1910, Promega). Firefly luciferase signal was normalized to Renilla luciferase activity. Adequate protein expression of ETV6::RUNX1 and RUNX1 was determined by Western blot.

3.2 Cell culture

HW8 hiPSCs were a kind gift from Dr. Herui Wang (Zhuang lab, National Cancer Institute, National Institutes of Health (NIH)) and were generated from PBMCs of a sarcoma patient using the CytoTune-iPS 2.0 Sendai Reprogramming kit (#A16517, Thermo Fisher Scientific) following written informed consent. Reprogramming via Sendai virus results in transient expression of POU5F1 (OCT3/4), SOX2, KLF4 and MYC. Study approval was obtained by the internal review board at the National Institutes of Health (protocol number: 16CN 069). Cellartis human iPSC line 12 (#Y00280, ChiPSC12) was purchased from Takara Bio. ChiPSC12 cells were induced from skin fibroblasts using replication-defective polycistronic retrovirus technology, leading to stable integration of *POU5F1*, *SOX2*, *KLF4* and *MYC*.

B-ALL cell lines and 293 cells were purchased from the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ). In-house leukemia patient samples for injection into NSG mice (The Jackson Laboratory) were retrieved from the Biobank of the University Hospital of Düsseldorf following informed consent in accordance with the Declaration of Helsinki. Study approval was obtained by the ethics committee of the Medical Faculty of the Heinrich Heine University (study number: 2019-566). All animal experiments adhered to regulatory guidelines set by the official committee at LANUV (Akt. 81-02.04.2017.A441) and were authorized by the animal research institute (ZETT) at Heinrich Heine University Düsseldorf. Patient blasts were injected intravenously into 6-week-old NSG mice (The Jackson Laboratory) and engraftment was assessed regularly by flow cytometric detection of human CD45+ cell percentage (#304011, BioLegend) in peripheral blood starting four weeks post injection [196]. Mice were sacrificed at predetermined timepoints and human CD45+ cells were isolated from bone marrow and spleen using the mouse cell depletion kit (#130-104-694, Miltenyi Biotec) to achieve >90% purity of human cells.

BCP-ALL cell lines were maintained at 37 °C and 5% CO_2 in RPMI GlutaMAX (#61870036, Thermo Fisher Scientific) with 20% fetal bovine serum (#S181H-500, Biowest). 293 and 293T cells were cultured in DMEM GlutaMAX (#31966, Thermo Fisher Scientific) with 10% fetal bovine serum.

3.2.1 hiPSC maintenance

hiPSCs were maintained in feeder cell-free conditions on Geltrex-coated plates (#A1413302, Thermo Fisher Scientific) in mTeSR Plus medium (#100-0276, STEMCELL Technologies). Geltrex was diluted 1:100 in DMEM/F-12 (#11320033, Thermo Fisher Scientific) and plates were coated for at least 1 hour at 37 °C. Cells were passaged according to culture density every 3-7 days by

incubation with Versene solution (#15040066, Thermo Fisher Scientific) for 7-8 min at RT and dissociated into aggregates of 20-50 cells by pipetting. Medium was exchanged every 1-2 days.

hiPSCs were cryopreserved as aggregates using mFreSR cryopreservation medium (#05855, STEMCELL Technologies). Cells were harvested using Versene solution, centrifuged at 150 g for 3 minutes and gently detached using cold mFreSR (4 °C). Cryovials (#374115, Thermo Fisher Scientific) were frozen using a Mr. Frosty container (#5100-0001, Thermo Fisher Scientific) to ensure slow cooling of -1 °C/minute, followed by long-term storage at -135 °C (liquid nitrogen).

Cryopreserved hiPSCs were thawed in a 37 °C water bath or by hand until a small frozen pellet remained. 500 µl mTeSR Plus medium were added dropwise to the cells and the cell suspension was transferred to a 15 ml Falcon tube containing 5 ml DMEM/F-12. The hiPSCs were then centrifuged at 150 g for 3 minutes and the cell pellet was resuspended in fresh mTeSR Plus with 10 µM ROCK inhibitor (Y-27632, #72304, STEMCELL Technologies). Cells were handled with during thawing to keep cell aggregates intact.

3.2.2 Nucleofection of hiPSCs

hiPSCs were incubated for at least 2 hours in mTeSR Plus with 10 µM Y-27632 before starting the Nucleofection procedure. Optimal Nucleofection conditions were determined using the hiPSC line HW8 by addition of 0.4 µg pmaxGFP control vector in P3 solution (#V4XP-3024 and #V4XP-3032, Lonza) and flow cytometric measurement of cells after 48 hours. Program CD-118 was selected for subsequent Nucleofections of hiPSCs.

First, an hiPSC single-cell suspension was prepared by incubation with Stempro Accutase solution (#A1110501, Thermo Fisher Scientific) for 5 min at 37 °C. About 10 x 10^6 hiPSCs were resuspended in 100 µl P3 solution with added supplement and transfected with 2.5 µg linearized double-stranded *RUNX1* HDR template (see chapter 3.1.1) as well as 4 µg each of pCW-Cas9 plasmid (a gift from Eric Lander and David Sabatini, Addgene plasmid #50661) coding for *Streptococcus pyogenes* Cas9 (SpCas9) plasmid (including 5' and 3' nuclear localization sequences (NLS)), and sgRNA plasmids (see chapter 3.1.2). Cells were resuspended in mTeSR Plus/Y-27632 and transferred onto a 100 mm Geltrex-coated plate. To remove floating dead cells, mTeSR Plus/Y-27632 was exchanged 2 hours after Nucleofection when cells became adherent and changed to mTeSR Plus without Y-27632 after 24 hours with further daily medium changes. Selection with 0.5 µg/ml puromycin (#A1113803, Thermo Fisher Scientific) was commenced 48 hours after Nucleofection and single colonies were picked under microscopic guidance into single wells of a 96-well plate at day 7-10. Clones were expanded for subsequent

confirmation of correct template insertion and expression. See **Figure 3.3** for a schematic representation of the hiPSC Nucleofection approach used in this work.

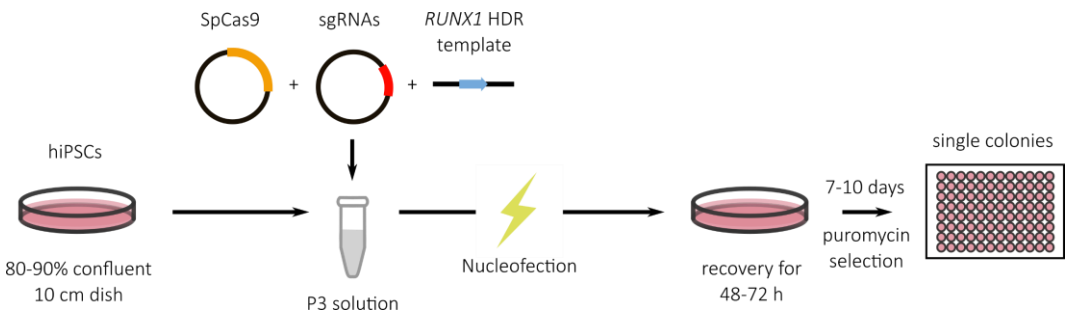


Figure 3.3: CRISPR/Cas9-mediated genome engineering approach in hiPSCs.

3.2.3 Transfection of 293 cells

ETV6::RUNX1 overexpressing 293 cells (#ACC 305, DSMZ) were generated as controls for Western blot analyses to test antibody specificity. 1.2 µg pMC3-HA-ETV6-RUNX1-puro vector were mixed with Opti-MEM reduced serum medium (#31985062, Thermo Fisher Scientific) to a final volume of 100 µl. 4.5 µl Attractene transfection reagent (#301005, QIAGEN) were added and the mixture was incubated for 10-15 minutes at RT to allow formation of transfection complexes. The DNA mixture was added dropwise to a single well of a 6-well plate of HEK293 cells at 50% confluence. After 2 days of incubation at 37 °C and 5% CO_2, the medium was exchanged to selective medium containing 0.75 µg/ml puromycin. After 48 hours, resistant cells were pelleted and either directly used for downstream applications or stored at -20 °C.

3.2.4 Nucleofection of REH cells

Transient knockdown of H1-0 in REH cells was performed by RNA interference (RNAi). For this, siRNAs were designed using the Eurofins siRNA design tool available online (https://eurofinsgenomics.eu/en/ecom/tools/sirna-design/). All siRNAs were purchased from Eurofins Genomics and are listed in **Table 3.13**. siRNA oligonucleotides were diluted to 100 µM with siMAX dilution buffer (30 mM HEPES, 100 mM KCl, 1 mM $MgCl_2$, pH 7.3; provided by Eurofins Genomics). Knockdown was performed with the 4D-Nucleofector system (#V4XC-2024, Lonza, SF solution, program DS-150) and 1 x 10^6 REH cells were transfected with 200 pmol of each siRNA in the 100 µl Nucleocuvette format according to the manufacturer's instructions.

Table 3.13: List of siRNA sequences used for H1-0 silencing experiments in REH cells.

name	sequence [5' → 3']	pool
H1-0_siRNA_1	GUAUAUCAAGAGCCACUAC	siH1-0_1
H1-0_siRNA_2	GAAGUCAGUGGCCUUCAAG	
H1-0_siRNA_3	GUCCAUUCAGAAGUAUAUC	siH1-0_2
H1-0_siRNA_4	GUCAGUGGCCUUCAAGAAG	
negative Ctrl siRNA (31% GC)	UAAUGUAUUGGAACGCAUA	siCtrl
negative Ctrl siRNA (31% GC)	AGGUAGUGUAAUCGCCUUG	

3.2.5 Hematopoietic differentiation of hiPSCs

To generate hematopoietic progenitor cells, *in vitro* differentiation of hiPSCs was performed by using the STEMdiff Hematopoietic kit (#05310, Stemcell Technologies) according to the manufacturer's instructions. On day 12, suspension cells were harvested and used for downstream analyses.

3.3 Screening of *ETV6::RUNX1*⁺ hiPSCs

Genetically engineered hiPSCs were analyzed for correct insertion of the *RUNX1* HDR template on DNA level and expression of the fusion gene both on RNA and protein level. The workflow of *ETV6::RUNX1*⁺ hiPSC selection is depicted in **Figure 3.4**.

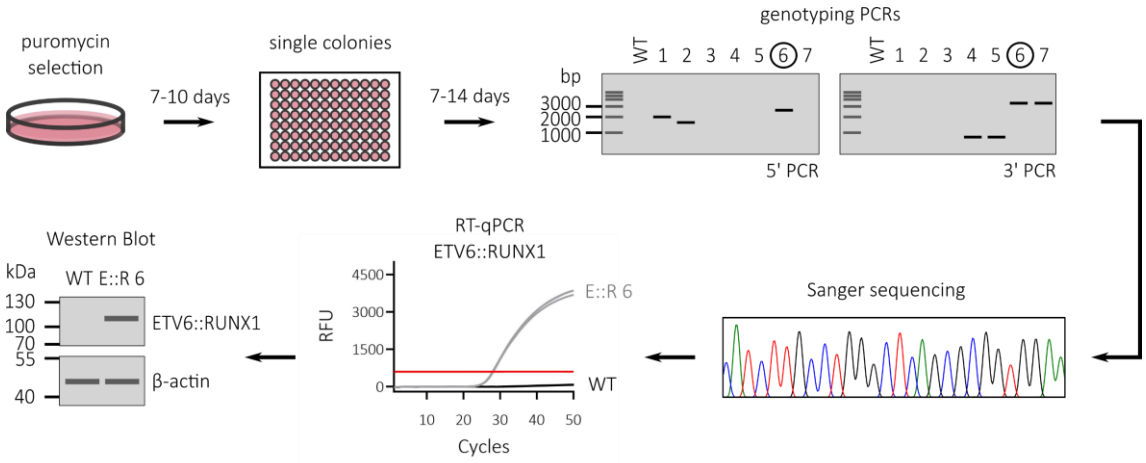


Figure 3.4: Workflow of *ETV6::RUNX1*⁺ hiPSC selection. Following Nucleofection, puromycin-resistant hiPSC clones were transferred into single wells and expanded. Clones correctly expressing ETV6::RUNX1 were selected by genotyping PCRs, Sanger sequencing of the complete insert site, RT-qPCR and Western blot.

3.3.1 Genotyping PCRs

Following puromycin selection, genomic DNA was isolated from each hiPSC clone using the QIAamp DNA Blood Mini kit (#51306, QIAGEN). A series of PCRs (Table 3.14) was performed to ensure correct incorporation of *RUNX1* HDR template at the *ETV6* locus. Two sets of primers spanning each of the homology arms (HA) were used to confirm proper orientation of the template (5'_fw and 5'_rev, 3'_fw and 3'_rev). A third PCR using three primers (3'_fw and 5'_fw2: 0.25 µM each, 3'_rev: 0.5 µM), two annealing outside of the HA region and one primer binding within the *RUNX1* HDR template sequence, was used to determine if the *RUNX1* HDR template was inserted at one or both alleles. Primers used for genotyping are listed in Table 3.15. The resulting PCR products were loaded onto an agarose gel to determine DNA fragment lengths and purified from the gel for analysis via Sanger sequencing.

Table 3.14: Pipetting scheme and thermocycling conditions used for PCR-based verification of correct *RUNX1* HDR template insertion. Extension times of PCRs varied according to amplicon lengths (1.5 minutes for 5'_fw + 5'_rev and 3'_fw + 3'_rev, 50 seconds for 5'_fw + 3'_rev + 5'_fw2).

	volume	amount
2X Phusion High-Fidelity PCR Master Mix	10 µl	1x
fw primer (10 µM)	1 µl	0.5 µM
rev primer (10 µM)	1 µl	0.5 µM
template DNA	X µl	100 ng
nuclease-free H_2O	to 20 µl	

98 °C	30 s	
98 °C	10 s	40 cycles
67-71 °C	30 s	
72 °C	30 s/kb	
72 °C	5 min	
4 °C	∞	

Table 3.15: Genotyping PCR primer sequences for screening of *ETV6::RUNX1*⁺ hiPSCs.

name	sequence [5' → 3']	amplicon length
5'_fw	CATACCTACACGCTCCTCCATTTAC	2639 bp
5'_rev	GCTTGGATCTGTAACGGCGCAG	
3'_fw	CAGAGTCAGATGCAGGATACAAGGC	2936 bp
3'_rev	CCTCCAGTCAAAACACACCTTCC	
5'_fw	CATACCTACACGCTCCTCCATTTAC	no insert: 1631 bp insert: 1631 bp and 1005 bp
3'_rev	CCTCCAGTCAAAACACACCTTCC	
5'_fw2	GCTAGCTTCGATCCAGACATG	

3.3.2 Sanger sequencing

Correct sequence of the *RUNX1* HDR template was verified by Sanger sequencing. Primer sequences used for Sanger sequencing are listed in Table 3.16. All Sanger sequencing reactions were performed at the Genomics and Transcriptomics Laboratory (GTL, Heinrich-Heine

University Düsseldorf, Germany). Sequencing was performed on an ABI 3130XL Genetic Analyzer (Thermo Fisher Scientific) and the data was analyzed and converted using the Sequencing Analysis Software v5.3.1 (Thermo Fisher Scientific). Sample files were visualized with the ApE software v3.0.8 [197].

Table 3.16: Sanger sequencing primer to determine correct insertion of the *RUNX1* HDR template.

name	sequence [5' → 3']
5'_seq_1_rev	GGCCACCACCTTGAAAGCGATG
5'_seq_2_fw	GCAAGCTGAGGAGCGGCG
3'_seq_3_rev	GCTTGGATCTGTAACGGCGCAG
3'_seq_4_fw	TAAAGCTCAGGTCGAGACCG
3'_seq_5_fw	GCTAGCTTCGATCCAGACATG

3.3.3 RT-qPCR: ETV6::RUNX1 and ETV6

Expression of ETV6::RUNX1 in hiPSC clones was confirmed by a performing reverse transcription quantitative PCR (RT-qPCR) using the ipsogen ETV6::RUNX1 kit (#675113, QIAGEN). RNA was isolated from the B cell precursor line REH, wild-type hiPSC lines and *ETV6::RUNX1*$^{+}$ hiPSC lines as described in chapter 3.1.10. 2 µg RNA were reverse transcribed to cDNA and diluted 1:4 (20 ng/µl RNA equivalent) for subsequent RT-qPCR. Synthesized cDNA was either stored at -20 °C or directly used as template for RT-qPCR to quantify the amount of ETV6::RUNX1 transcripts (see Table 3.17). ABL1 expression was measured simultaneously to normalize fusion gene expression. For quantification of control and fusion gene transcripts, plasmid DNA standards are included in the ipsogen ETV6::RUNX1 kit.

Table 3.17: Pipetting scheme and thermocycling conditions used for ETV6::RUNX1 RT-qPCR.

	volume	amount
TaqMan Universal qPCR Master Mix (2x, #4304437, Thermo Fisher Scientific)	12.5 µl	1x
Ipsogen primer and probe (25x)	1 µl	1x
template cDNA	5 µl	100 ng RNA equivalent
nuclease-free H_2O	to 25 µl	

50 °C	2 min	
95 °C	10 min	
95 °C	15 s	50 cycles
60 °C	1 min	
acquisition of FAM		

To determine if the *RUNX1* HDR was inserted into one or both *ETV6* alleles, a RT-qPCR using a primer and probe set specific for *ETV6* exons 5 and 6 was performed (see Table 3.12). The *ETV6*

probe contains both a 5' FAM fluorophore and 3' quencher (Eclipse). Primer and probe sequences were purchased from Eurofins Genomics. Total RNA was isolated as described before and 2 µg RNA were reverse transcribed to cDNA (see Table 3.11). Synthesized cDNA was diluted 1:8 and 2 µl diluted cDNA (20 ng RNA equivalent) were used for RT-qPCR (see Table 3.18).

Table 3.18: Pipetting scheme and thermocycling conditions to determine ETV6 expression.

	volume	amount
TaqMan Universal qPCR Master Mix	10 µl	1x
ETV6_ qPCR_fw (10 µM)	1 µl	0.5 µM
ETV6_ qPCR_rev (10 µM)	1 µl	0.5 µM
ETV6_ qPCR_probe (10 µM)	0.5 µl	0.25 µM
template cDNA	2 µl	20 ng RNA equivalent
nuclease-free H_2O	to 20 µl	

50 °C	2 min	
95 °C	10 min	
95 °C	15 s	40 cycles
60 °C	1 min	
acquisition of FAM		

3.4 *In vitro* inhibitor treatments

BCP-ALL cell lines were treated with 1 µM JNJ-26481585/Quisinostat at a concentration of 1×10^6 cells per ml and RNA was extracted after 24 hours for subsequent analysis of *H1-0* expression by RT-qPCR. DMSO-dissolved compounds were purchased from Selleck Chemicals and MedChem Express. For drug synergy analysis, Quisinostat (concentration range: 0.2 nM - 20 nM), Vincristine (concentration range: 0.1 nM - 5 nM), Daunorubicin (concentration range: 1.5 nM - 50 nM) and Bortezomib (concentration range: 1 nM - 10 nM) were printed in a randomized fashion in increasing concentrations in 8 x 8 matrices using a D300e digital dispenser (Tecan) and normalized with DMSO (Sigma-Aldrich, #2650). Plates were incubated for 72 hours and viability was determined by CellTiter-Glo Luminescent viability assay (Promega) using a Tecan SPARK 10M reader. Synergy scores were determined using the zero interaction potency method (ZIP) using the SynergyFinder web application (version 3.0).

3.5 Protein Biochemistry

3.5.1 Protein extraction and BCA assay

Cell pellets were dissolved in ice-cold RIPA lysis buffer (#89900, Thermo Fisher Scientific) with freshly added protease inhibitors (#11836145001, Roche) and incubated at 4 °C with agitation for 30 minutes. Samples were then centrifuged at 16,000 g for 15 minutes and the supernatant was transferred to a new microreaction tube.

Total protein concentrations of the samples were determined with a bicinchoninic acid assay (BCA, #23225, Thermo Fisher Scientific) according to the manufacturer's instructions. The absorbance of the samples was measured at 562 nm with a Tecan Spark microplate reader (#30086376, Tecan) and protein concentrations were calculated by referring to a BSA standard.

Protein lysates were diluted with RIPA buffer and mixed with 6x Laemmli loading buffer (#J61337, Thermo Fisher Scientific) to achieve the desired protein concentration. To denature the proteins, the diluted samples were boiled at 95 °C for 5 minutes before proceeding to Western blotting.

3.5.2 Western blot

5-20 µg protein lysate were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using 8.5% or 10% gels (see Table 3.19) and transferred to Amersham Protran 0.45 µm nitrocellulose membranes (#GE10600012, Merck) by wet blotting using the Mini-Protean Vertical Electrophoresis system (Bio-Rad). For protein size determination, PageRuler Prestained Protein Ladder (#26616, Thermo Fisher Scientific) was used. Membranes were blocked for 1 h with 5% BSA (#A3294, Merck) in T-BST at RT and incubated with primary antibodies diluted in blocking buffer overnight at 4 °C. The next day, membranes were incubated with secondary antibodies diluted in blocking solution for 1 h at RT. Antibodies used in this study are listed in Table 3.20. Signal development was performed using ECL detection reagent (#GERPN2109, Merck) according to the manufacturer's instructions and images were acquired using the JESS Western system (Proteinsimple, Bio-Techne). For further detections, membranes were incubated in T-BST with added 0.1% NaN_3 (#S2002, Merck) in T-BST for 1 h at RT or Re-Blot Plus Strong Antibody Stripping (#2504, Merck) solution for 15 min at RT. Quantification of protein bands was performed using ImageJ analysis software (release 1.53c).

Table 3.19: Composition of SDS-PAGE gels for Western Blot.

separating gel (for two 1.5 mm gels):	8.5%	10%
H_2O	8.22 ml	7.32 ml
1.5 M Tris (pH 8.8, #T1503, Merck)	4.68 ml	4.68 ml
acrylamide (30%, #3029.1, Carl Roth)	5.1 ml	6 ml
APS (10%, #A3678, Sigma-Aldrich)	180 µl	180 µl
TEMED (#110732, Merck)	40 µl	40 µl
total volume	18.58 ml	18.22 ml

stacking gel (for two 1.5 mm gels):	
H_2O	4.16 ml
0.5 M Tris (pH 6.8)	1.82 ml
acrylamide (30%)	933 µl
APS (10%)	70 µl
TEMED	17 µl
total volume	7 ml

Table 3.20: Antibodies used for Western blot detection.

target	species	dilution	supplier
RUNX1/2/3 (detects ETV6::RUNX1)	rabbit IgG monoclonal	1:2000	#ab92336, Abcam
ETV6	mouse monoclonal IgG	1:2000	#sc-166835, Santa Cruz Biotechnology
H1-0	rabbit monoclonal IgG	1:2000	#MA5-35484 (ARC1059), Thermo Fisher Scientific
FLAG (M2)	mouse monoclonal IgG1	1:2000	#F1804, Sigma-Aldrich
ACTB	mouse monoclonal IgG2a	1:5000	#A5316, Merck
HRP-linked secondary anti-rabbit	goat monoclonal IgG	1:1000	#7074S, Cell Signaling Technology
HRP-linked secondary anti-mouse	horse monoclonal IgG	1:2000	#7076S, Cell Signaling Technology

3.6 Next-generation sequencing

3.6.1 Bulk RNA sequencing

Sample preparation. HW8 and ChiPSC12 hiPSCs of similar passage number (±2 passages) were lysed on ice using RLT buffer (#79216, Qiagen) and detached by scraping. RNA was isolated using the RNeasy Mini kit (#74104, Qiagen) with on-column DNA digestion (#79254, Qiagen). REH cells were transfected with siRNA as described above and RNA was extracted after 48 hours. RNA quality was assessed with the 2100 Bioanalyzer system (#G2939BA, Agilent).

Library preparation and sequencing. Barcoded libraries were prepared from 0.5 µg of total RNA using the TruSeq RNA Sample Preparation v2 kit (low-throughput protocol, Illumina) and quantified with the Bioanalyzer system (Agilent). 7.5 pM denatured libraries were used as input for cBot (Illumina) and subjected to deep sequencing using the NovaSeq 6000 (Illumina) for 101 cycles, with an additional 7 cycles for index reading. Sequencing was performed at the next-generation sequencing core facility of the German Cancer Research Center (DKFZ).

Data analysis. Analysis of fastq files was performed by using the Partek Flow software (Partek Incorporated). After assessing the read quality, a trimming step was performed (both ends: 13 bases at the 5' end and 1 base at the 3' end). One technical replicate of siH1-0_1 was excluded from the analysis since it did not pass quality control. After trimming, reads were aligned to the hg38 genome using the STAR v2.4.1d aligner. Unaligned reads were further processed using Bowtie 2 v2.2.5 aligner. Aligned reads were combined and expression was quantified against the Ensembl database (release 84) by the Partek Expectation-Maximization algorithm [198]. Partek flow default settings were used in all analyses.

Unsupervised hierarchical clustering and heatmap visualization of samples was performed after normalizing mean expression to 0 with a standard deviation (SD) of 1 and using Pearson's dissimilarity algorithm and average linkage in Partek Genomics Suite (Partek Incorporated). Ingenuity pathway analysis (IPA, Qiagen) of REH cells was performed taking into account significantly dysregulated genes (absolute fold change (FC)>1.5 and $p<0.05$). The significance cut-off for IPA to identify upstream regulators was set to $p<0.05$. Venn diagrams of overlapping genes were generated using the web-based tool InteractiVenn [199]. RNA-seq analysis was performed by Daniel Picard.

Gene set enrichment analysis (GSEA). GSEA was performed on processed RNA-seq data of REH cells treated with non-targeting or *H1-0*-targeting siRNA pools using the GSEA v4.2.3 software (http://software.broadinstitute.org/gsea). Genes were ranked by the GSEA software using the signal-to-noise metric. The permutation type was set to gene_set and number of permutations to 5000. Canonical pathways or hallmark gene sets were obtained from the Molecular Signatures Database (MSigDB, https://www.gsea-msigdb.org). Genes significantly upregulated upon *ETV6::RUNX1* knockdown in REH and AT-2 cells (cutoffs: $\log_2$ FC>0.9 and adjusted $p<0.05$, n=103 genes) were derived from a published dataset [130]. Visualization of pathway networks was performed using the Cytoscape EnrichmentMap and AutoAnnotate applications [200, 201].

3.6.2 Single-cell RNA sequencing

Sample preparation. Following *in vitro* differentiation of hiPSCs, hematopoietic progenitor cells of 5 wells were pooled, filtered and resuspended in PBS with 0.04% BSA (Biowest). Cells were resuspended in PBS with 0.04% BSA in DNA LoBind microreaction tubes (#022431021, Eppendorf) and cell viability was determined using the BD Rhapsody Single-Cell Analysis system (>70% viability, BD Biosciences).

Library preparation and sequencing. Single-cell suspensions were used for single-cell droplet library generation on the 10X Chromium Controller system using the Chromium Single Cell 3' NextGEM Reagent kit v3.1 (10X Genomics) according to the manufacturer's instructions. Sequencing was performed on a NextSeq 2000 system (Illumina). All scRNA-seq reactions were performed at the GTL (Heinrich Heine University Düsseldorf).

Data analysis. Raw sequencing data was processed using the 10X Genomics CellRanger software (v6.0.2). Raw BCL-files were demultiplexed and processed to fastq files using the CellRanger mkfastq pipeline. Alignment of reads and UMI counting was performed via the CellRanger count pipeline to generate a gene-barcode matrix (genome version: GRCh38, Ensemble release 98). The CellRanger aggr pipeline was used for aggregation and sequencing depth normalization.

Filtered cells were normalized using the PFlog1pPF method [202]. Thus, differences in sequencing depth were normalized via the normalize_total function in Scanpy (version 1.9.1) followed by variance stabilizing log+1 transformation by applying the log1p function, and a second depth normalization. Feature selection based on binomial deviance [203] was performed on raw counts using the package Scry (version 1.10.0, https://doi.org/10.18129/B9.bioc.scry). Subsequently, the 4000 most deviant genes were used to compute principal components (PCs) via the pca function in Scanpy. The top 50 PCs were used to calculate the neighborhood graph via the neighbors function (n_neighbors=15), and Uniform Manifold Approximation and Projection (UMAP) was computed using the umap function.

Cell cycle phase was inferred by scoring the cell cycle gene set as defined by Tirosh *et al*. [204] (from https://github.com/scverse/scanpy_usage/tree/master/180209_cell_cycle) by applying the score_genes_cell_cycle function. For cell type annotation, gene sets underlying cell types defined by Jardine *et al*. [205] were extracted via the rank_genes_groups_df function (pval_cutoff=0.05, log2fc_min=2) and used for scoring via the score_genes function. For visualization purposes, gene expression values were scaled using scale (max_value=10) function. scRNA-seq analysis was performed by Andrea Hanel.

3.7 Quality control

3.7.1 Short tandem repeat (STR) analysis

For cell line authentication, short tandem repeat (STR) analysis was performed using a set of defined regions of repeated genomic DNA as genetic markers. STR loci TH01, TPOX, vWA, CSF1PO, D16S539, D7S820, D13S317, D5S818, D21S11 and AMEL were amplified by PCR using the GenePrint 10 system (#B9510, Promega). An internal DNA length standard was added before separation and fluorescence detection on an ABI 3130XL Genetic Analyzer (Thermo Fisher Scientific). Data analysis was performed with the Microsatellite Analysis tool (Thermo Fisher Scientific). The STR profiles were then compared to authenticated cell lines listed in the DSMZ database (https://www.dsmz.de/services/human-and-animal-cell-lines/online-str-analysis). STR profiling was performed at the GTL (Heinrich Heine University Düsseldorf).

3.7.2 Mycoplasma test

All cell lines were routinely tested for mycoplasma contamination using the Venor GeM Advance mycoplasma detection kit (#11-7024, Minerva Biolabs). The kit includes both a positive control (265-278 bp) and internal control (191 bp).

3.7.3 hiPSC karyotype analysis

Chromosomal integrity of hiPSCs was confirmed by karyotyping, kindly performed by Judith Bartel at the Institute of Human Genetics (lab of PD Dr. Anke Bergmann, Hannover Medical School (MHH), Germany). For each hiPSC line, 10-15 metaphase spreads of Giemsa-stained chromosomes were analyzed, confirming intact karyotypes of 46, XY.

3.7.4 Testing RT-qPCR primer efficiencies

To ensure that RT-qPCR amplification efficiencies were in a range of 85-110%, standard curves were generated for each RT-qPCR primer pair using 1:2 cDNA dilutions. Reaction efficiencies were calculated according to standard curve slope:

$$\text{efficiency (\%)} = (10^{\frac{-1}{\text{slope}}} - 1) \times 100$$

When using SYBR green chemistry, melt curves were generated after each experiment. Additionally, the correct molecular weight of the amplified fragment was determined on a 2% agarose gel.

3.7.5 hiPSC pluripotency assays

To analyze the pluripotent status of hiPSC lines, flow cytometry (see chapter 3.7.6) and RT-qPCR assays were performed to detect expression of pluripotency gene markers. RT-qPCR primer sequences were kindly provided by the lab of Prof. Dr. Alessandro Prigione (DNMT3B) or designed using Primer-BLAST (GDF3, POU5F1, NANOG; https://www.ncbi.nlm.nih.gov). RT-qPCR primers are listed in Table 3.12. As described before, 2 µg total RNA was reverse transcribed to cDNA and diluted 1:10 with nuclease-free water. 2 µl diluted cDNA (20 ng RNA equivalent) were used for RT-qPCR.

3.7.6 Flow cytometry

For fluorescence-activated cell sorting (FACS), hiPSCs were detached with Stempro Accutase solution for 5 min at 37 °C, centrifuged at 200 g for 3 minutes and incubated with human TruStain FcX blocking reagent (#422301, BioLegend) for 10 minutes on ice. hiPSCs were stained with PE-labelled SSEA-4 antibodies (#330405, BioLegend) for 15 minutes on ice, washed with PBS (+2% FCS) and fluorescence was measured with a CytoFLEX S flow cytometer (Beckmann Coulter). Dead cells were excluded by staining with 0.1 µg/ml DAPI (#D9542, Merck).

Output of $CD34^+$ cells following differentiation of hiPSCs was determined by staining with PE/Dazzle594-labeled CD34 antibodies (#343534, BioLegend). Unspecific binding sites were blocked with human TruStain FcX blocking reagent (#422301, BioLegend) and dead cells were excluded by staining with Viobility 405/520 fixable dye (#130109814, Miltenyi Biotec).

3.8 Data analysis

3.8.1 Data availability and bioinformatic analyses

RNA-seq data of preleukemia models was obtained from the ArrayExpress functional genomics data collection website (https://www.ebi.ac.uk/biostudies/arrayexpress) accession number E-MTAB-6382 [125]. RNA expression data of leukemia subtypes and normal B cell developmental stages was obtained from the St. Jude PeCan Data Portal (https://pecan.stjude.cloud) [206, 207] and from the R2 Genomics Analysis Visualization Platform (http://r2.amc.nl; GSE13159 dataset [208], microarray platform u133p2; GSE87070 dataset [209, 210], microarray platform u133p2; GSE24759 dataset [211], microarray platform u133a). Processed DNA methylation (Infinium HumanMethylation450 BeadChip platform) and matched RNA expression data (microarray platform u133p2) of various leukemia entities was

retrieved from Gene Expression Omnibus (NCBI GEO, https://www.ncbi.nlm.nih.gov/geo) accession number GSE49032 [158].

RNA-seq data of normal B cell developmental stages and *ETV6::RUNX1*$^+$ BCP-ALL samples was retrieved from NCBI GEO (accession number GSE115656 [212]) and processed using the Galaxy platform (https://usegalaxy.eu). Fastq files were trimmed using the Trimmomatic tool and aligned to the hg38 genome using the HISAT2 aligner. Expression was quantified using htseq-count against the UCSC database. To exclude effects of underlying predisposing syndromes, one *ETV6::RUNX1*$^+$ BCP-ALL patient presenting with trisomy 21 was omitted from the analysis.

Analysis of scRNA-seq data derived from normal bone marrow precursor B cells was performed as previously described [213]. Fetal liver scRNA-seq data was derived from the Developmental Cell Atlas accession number E-MTAB-7407 [214] (https://www.humancellatlas.org/, Newcastle University).

ChIP-seq datasets of REH cells for H3K4me1, H3K4me3, H3K27ac and RUNX1 (accession number GSE117684 [215]), as well as ETV6::RUNX1 (accession number GSE176084 [133]) were downloaded from NCBI GEO. Fastq files were processed as described for bulk RNA-seq and BAM files were visualized using IGV version 2.9.1 (https://igv.org) [216]. See Table 3.21 for a summary of datasets analyzed in this study.

Table 3.21: Publicly available datasets analyzed in this study.

accession number	description	reference
E-MTAB-6382	RNA-seq data of *ETV6::RUNX1*$^+$ preleukemia	[125]
GSE13159	expression microarray data of leukemia entities (Microarray Innovations in Leukemia (MILE) study cohort, microarray platform u133p2, MAS5.0 normalization)	[208]
GSE87070	expression microarray data of leukemia entities (microarray platform u133p2, MAS5.0 normalization)	[209, 210]
GSE49032	matched DNA methylation and expression microarray data of leukemia entities (Infinium HumanMethylation450 BeadChip platform and microarray platform u133p2)	[158]
GSE115656	RNA-seq data of normal B cell developmental stages and *ETV6::RUNX1*$^+$ BCP-ALL samples	[212]
GSE24759	expression microarray data of normal B cell developmental stages (microarray platform u133a, MAS5.0 normalization)	[211]
GSE117684 GSE176084	ChIP-seq data of REH cells	[215] [129]
E-MTAB-7407	scRNA-seq data of fetal liver cells	[214]

The datasets produced in this study are available on the Gene Expression Omnibus (GEO) database under the accession numbers GSE270944 (RNA-seq) and GSE270945 (scRNA-seq).

3.8.2 Statistical analyses

Statistical analysis of data was performed using GraphPad Prism version 9.5.1. The number n of replicates and statistical tests are indicated in the Figure descriptions. Statistical significance was considered for p-values *p<0.05, **p<0.01 and ***p<0.001.

4. RESULTS

4.1 Generation of *ETV6::RUNX1*$^{+}$ hiPSC preleukemia models

Faithful recapitulation of *ETV6::RUNX1*$^{+}$ preleukemia has been challenging thus far. This is likely a consequence of the subtle effects of *ETV6::RUNX1* as a first hit mutation and is reflected by the observation that the function of ETV6::RUNX1 is dose-dependent [217]. While *ETV6::RUNX1*$^{+}$ mouse models have shown inconsistent phenotypes and largely failed to reproduce B lineage restriction of overt leukemia seen in patients [140, 141, 218, 219], the use of hiPSCs as a modeling approach holds great promise for investigating human *ETV6::RUNX1*$^{+}$ preleukemia, as demonstrated recently by Böiers *et al*. [125].

Hence, a CRISPR/Cas9 genome editing approach was used to generate an hiPSC model system that expresses *ETV6::RUNX1* from the endogenous *ETV6* locus (**Figure 4.1**). For this, a *RUNX1* homology-directed repair (HDR) template was generated using PCR-based cloning that allowed for targeted insertion of *RUNX1* exons 2-8 at the endogenous *ETV6* locus.

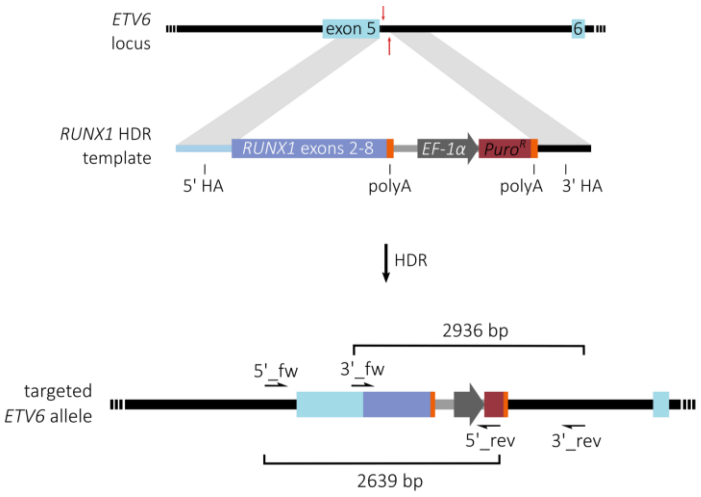


Figure 4.1: Schematic representation of the *RUNX1* HDR template. PCR primers used for detection of correct integration of the *RUNX1* HDR template at the endogenous *ETV6* locus are indicated.

Given that *ETV6* intron 5 is the most common breakpoint cluster region of the *ETV6::RUNX1* translocation in patients [45], sgRNAs were designed to target a region directly downstream of *ETV6* exon 5 (**Figure 4.2**). To enable optimal insertion of the *RUNX1* HDR template, sgRNAs were designed to cut within a maximum distance of 100 bp from the start sites of the repair template's homology arm (HA) sequences (5' HA: 87 bp, 3' HA: 26 bp) in a protospacer adjacent

motif (PAM)-out configuration. Using nucleofection, the *RUNX1* HDR template was inserted into the genome of hiPSCs derived from two donors (HW8 and ChiPSC12) and successfully edited hiPSC colonies were selected via puromycin treatment.

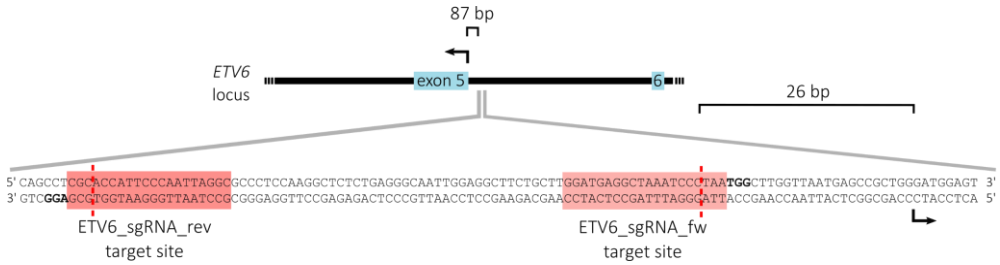


Figure 4.2: Orientation of sgRNA target sites at the *ETV6* locus. Localization of sgRNA target sites at the *ETV6* locus. Start sites of homology arm (HA) regions are indicated by arrows and Cas9 cutting sites are indicated by red dotted lines.

Targeting of the *ETV6* locus was validated on DNA level by specific genotyping PCRs covering the 5' and 3' region of the targeted locus (**Figure 4.1**). Successfully edited puromycin-resistant hiPSCs were identified by PCR product lengths of both 2639 bp (5' PCR) and 2936 bp (3' PCR), as shown representatively for 7 monoclonal hiPSC lines in **Figure 4.3**. Screening of >200 colonies identified 3 *ETV6::RUNX1*$^{+}$ hiPSC lines with correct orientation and sequence of the *RUNX1* HDR template, as confirmed by PCR (**Figure 4.4A**) and Sanger sequencing (**Figure 4.4B**).

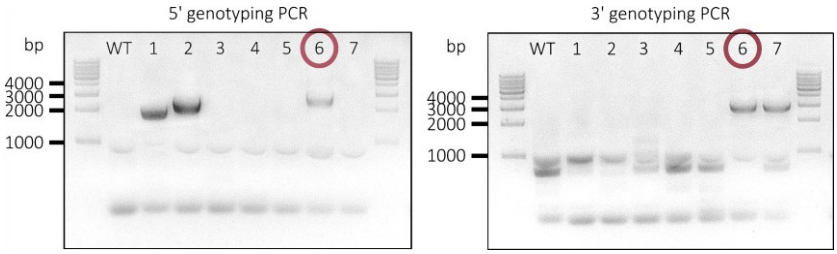


Figure 4.3: Representative genotyping PCR screen of CRISPR/Cas9 genome engineered HW8 hiPSCs. DNA extracted from 7 HW8 hiPSC clones following puromycin selection was subjected to 5' and 3' genotyping PCRs. Clone 6 shows PCR product sizes of correct lengths (5' PCR: 2639 bp, 3' PCR: 2936 bp) and was used for further downstream analyses.

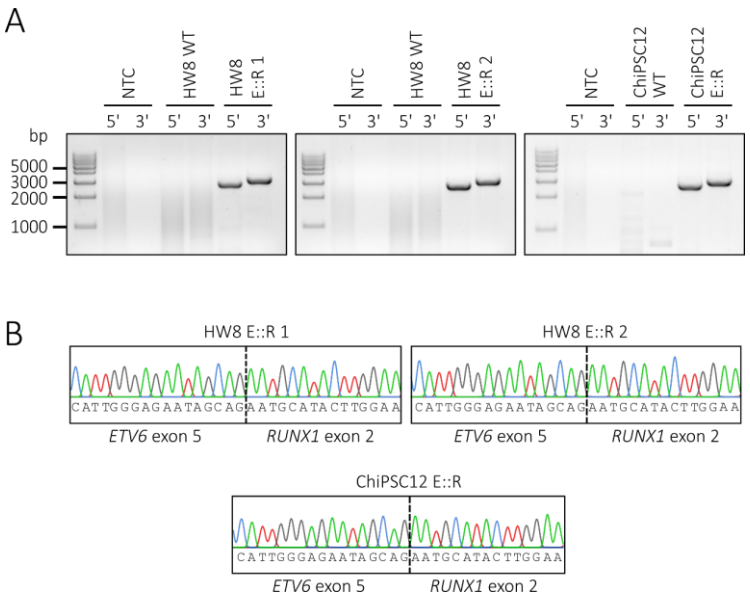


Figure 4.4: CRISPR/Cas-edited hiPSCs show stable integration of the *RUNX1* HDR template into the endogenous *ETV6* locus. (A) Genotyping PCRs confirming correct insertion of the *RUNX1* HDR template into the endogenous *ETV6* locus of HW8 and ChiPSC12 hiPSCs (5' PCR: 2639 bp, 3' PCR: 2936 bp). **(B)** Sanger sequencing traces at the *ETV6::RUNX1* breakpoint region of the 3 successfully targeted hiPSC clones. NTC: no-template control (H_2O).

Correct integration of the *RUNX1* HDR template into the *ETV6* locus translated to expression of the *ETV6::RUNX1* fusion gene on RNA (**Figure 4.5A**) and protein level (**Figure 4.5B**). Of note, RNA and protein levels of ETV6::RUNX1 were lower in the hiPSC lines compared to the *ETV6::RUNX1*$^+$ BCP-ALL cell line REH.

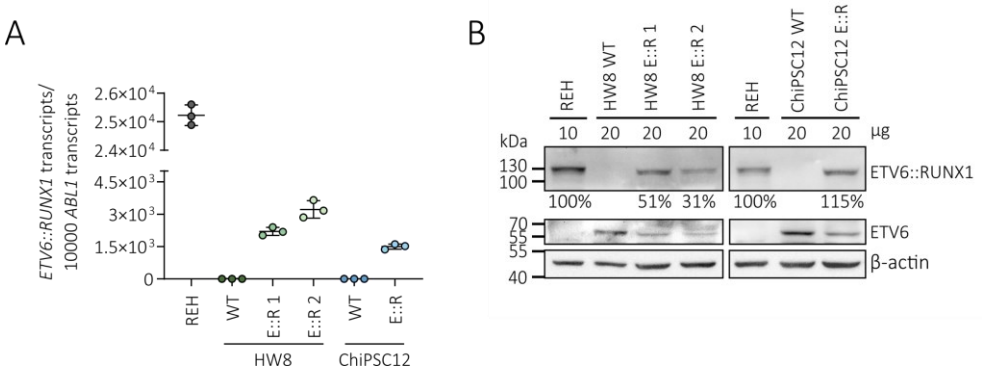


Figure 4.5: CRISPR/Cas-edited hiPSCs express ETV6::RUNX1 on RNA and protein level. (A) Quantification of ETV6::RUNX1 expression by RT-qPCR in REH, HW8 and ChiPSC12 hiPSC lines. Data is presented as the mean ± SD. **(B)** Western blot analysis of REH, HW8 and ChiPSC12 hiPSC lysates for expression of ETV6::RUNX1, ETV6 and β-actin.

Detection of the ETV6::RUNX1 fusion protein was performed based on protein size using a monoclonal antibody raised against a synthetic RUNX1 peptide (#ab92336, Abcam). This antibody detects the three RUNX family proteins RUNX1, RUNX2, RUNX3 (≈45-55 kDa) as well as ETV6::RUNX1 (≈110 kDa) with high specificity and sensitivity, as shown by dilution of protein lysates extracted from REH and 293 cells transiently overexpressing ETV6::RUNX1 (**Figure 4.6**).

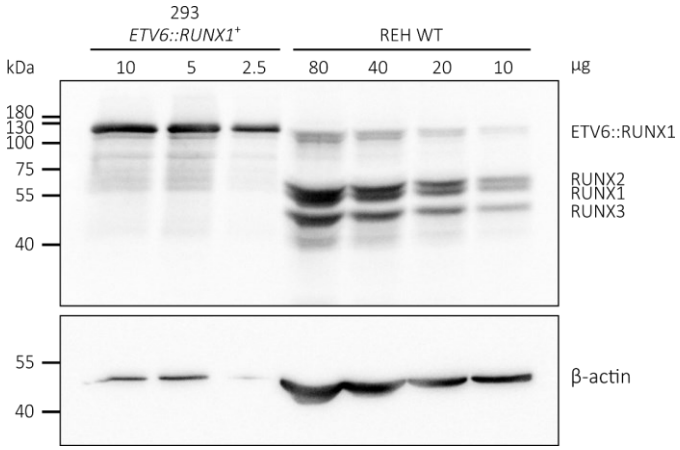


Figure 4.6: Testing of antibody specificity and sensitivity for Western blot detection of ETV6::RUNX1. Western blot analysis of different protein amounts extracted from REH cells and 293 cells overexpressing ETV6::RUNX1.

In BCP-ALL patients, the *ETV6::RUNX1* translocation disrupts one allele of the *ETV6* gene, while loss of the second *ETV6* allele is the most common second hit mutation in these leukemias [143]. Therefore, accurately modeling *ETV6::RUNX1*$^+$ preleukemia requires one *ETV6* allele to remain intact. To identify monoallelic targeting of the *ETV6* gene in the 3 *ETV6::RUNX1*$^+$ hiPSC lines, an additional genotyping PCR was performed that allows to differentiate between targeted (1005 bp) or non-targeted (1631 bp) *ETV6* alleles (**Figure 4.7A**). Using this approach, monoallelic expression of *ETV6::RUNX1* could be confirmed in all 3 hiPSC lines (**Figure 4.7B**). To determine sequence integrity of the remaining wild-type *ETV6* allele, the amplified fragment of the non-targeted *ETV6* allele was Sanger sequenced. Of note, HW8 E::R 1 contained two intronic point mutations at the Cas9 endonuclease target sites (**Figure 4.7C**). However, these indels did not disrupt transcription of wild-type *ETV6*, as shown by Western blot (**Figure 4.5B** and **4.16C**).

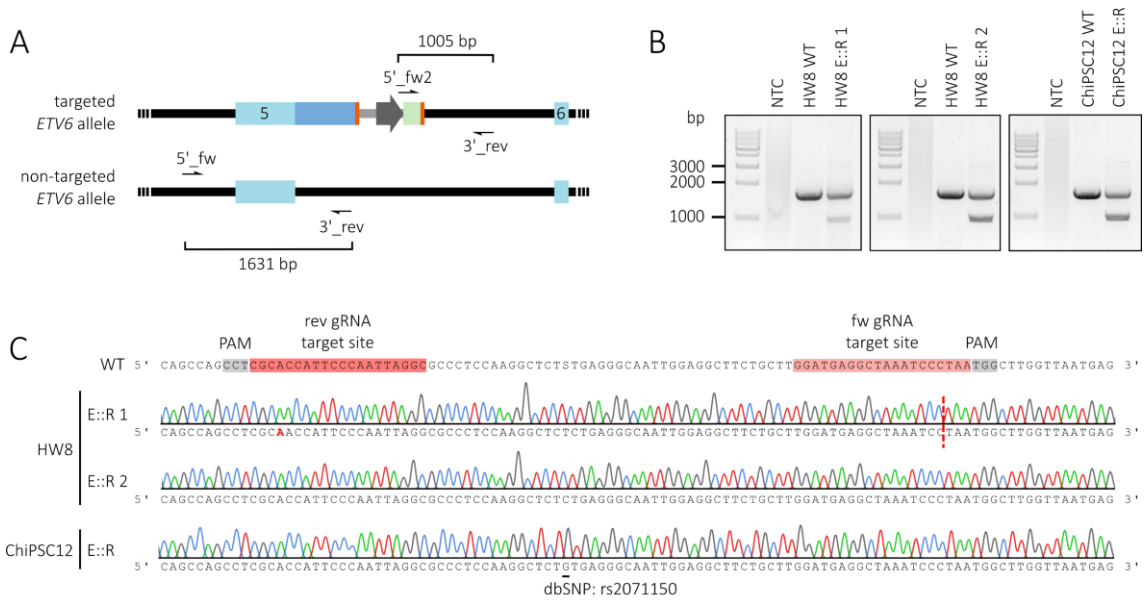


Figure 4.7: Genetically modified hiPSCs display heterozygous expression of *ETV6::RUNX1*. (A) Primer binding sites (arrows) used for PCR detection of *ETV6::RUNX1* homo- or heterozygosity. **(B)** PCRs confirming heterozygous allele status of *ETV6::RUNX1* in HW8 and ChiPSC12 hiPSCs. **(C)** Sanger sequencing of the sgRNA binding region within the non-targeted *ETV6* intron 5 in the *ETV6::RUNX1*$^{+}$ hiPSC clones. sgRNA binding sites are indicated. NTC: no-template control (H_2O), PAM: protospacer adjacent motif, SNP: single-nucleotide polymorphism.

To confirm correct expression of the untargeted *ETV6* allele, an RT-qPCR assay was designed using primers spanning *ETV6* exons 5 and 6 (**Figure 4.8A**). Given that *ETV6* exon 6 is not included in *ETV6::RUNX1*, this assay does not detect expression of the fusion gene, but only of ETV6. Indeed, no signal was detected in REH cells that express *ETV6::RUNX1* and have a deletion of the second *ETV6* allele (**Figure 4.8B**). Reduced *ETV6* levels were observed in all *ETV6::RUNX1*$^{+}$ hiPSCs compared to their respective wild-type counterpart. This was confirmed on protein level by Western blot (**Figure 4.5B** and **4.16C**).

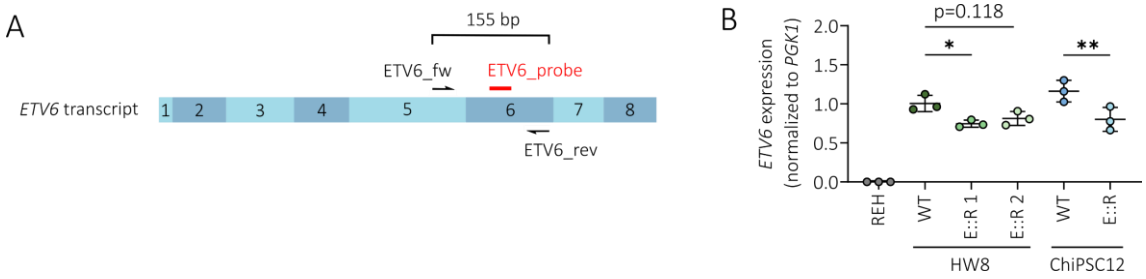


Figure 4.8: *ETV6::RUNX1*$^{+}$ hiPSCs show reduced *ETV6* expression. (A) Schematic representation of the *ETV6* RT-qPCR design. Arrows indicate primer binding sites. The probe sequence is marked in red. **(B)** Relative *ETV6* expression determined by RT-qPCR in REH cells, HW8 and ChiPSC12 hiPSC lines. Data is presented as the mean ± SD and analyzed for statistical significance using an ordinary one-way ANOVA. * = p<0.05, ** = p<0.01

All hiPSC lines were tested for chromosomal integrity by karyotype analysis, which indicated that no additional structural abnormalities had been caused by CRISPR/Cas9-mediated genome editing (**Figure 4.9**). Appropriate morphology and pluripotency marker expression was confirmed via microscopic assessment, FACS and RT-qPCR (**Figure 4.10**). To determine off-target activity for both *ETV6*-targeting sgRNAs, Sanger sequencing of *in silico*-predicted off-target sites with ≤2 mismatches from the target sequences was performed. Analysis of predicted off-target sites in the 3 *ETV6::RUNX1*$^{+}$ hiPSC lines did not show any genetic alterations in the tested regions (**Table 4.1**).

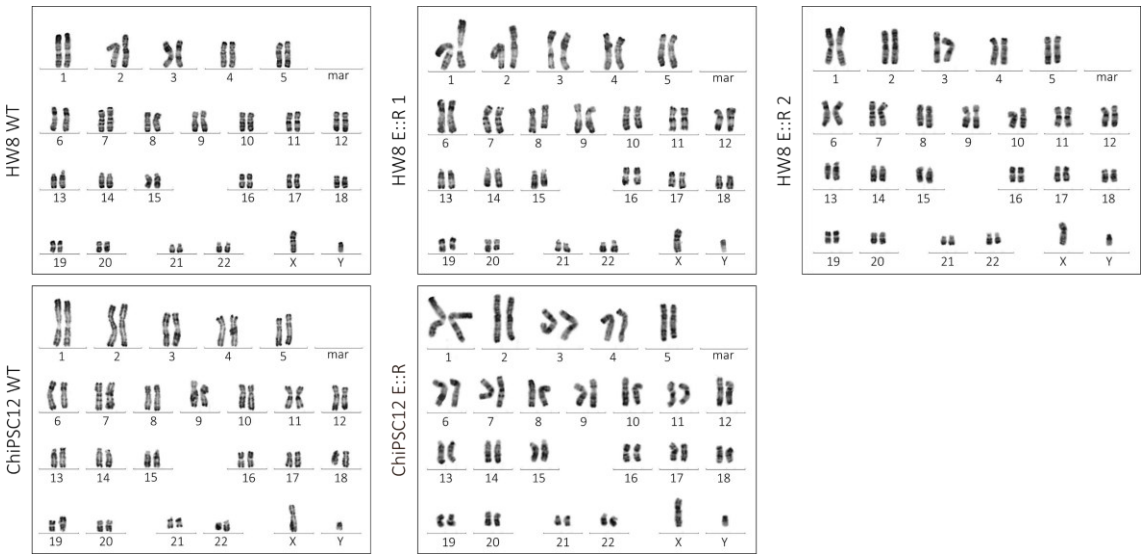


Figure 4.9: Chromosomal integrity is maintained in *ETV6::RUNX1*$^{+}$ hiPSCs following CRISPR/Cas editing. Representative karyotype data for *ETV6::RUNX1*$^{+}$ and wild-type hiPSCs. Data was generously provided by Judith Bartel at the Institute of Human Genetics (lab of Dr. Anke Bergmann, Hannover Medical School).

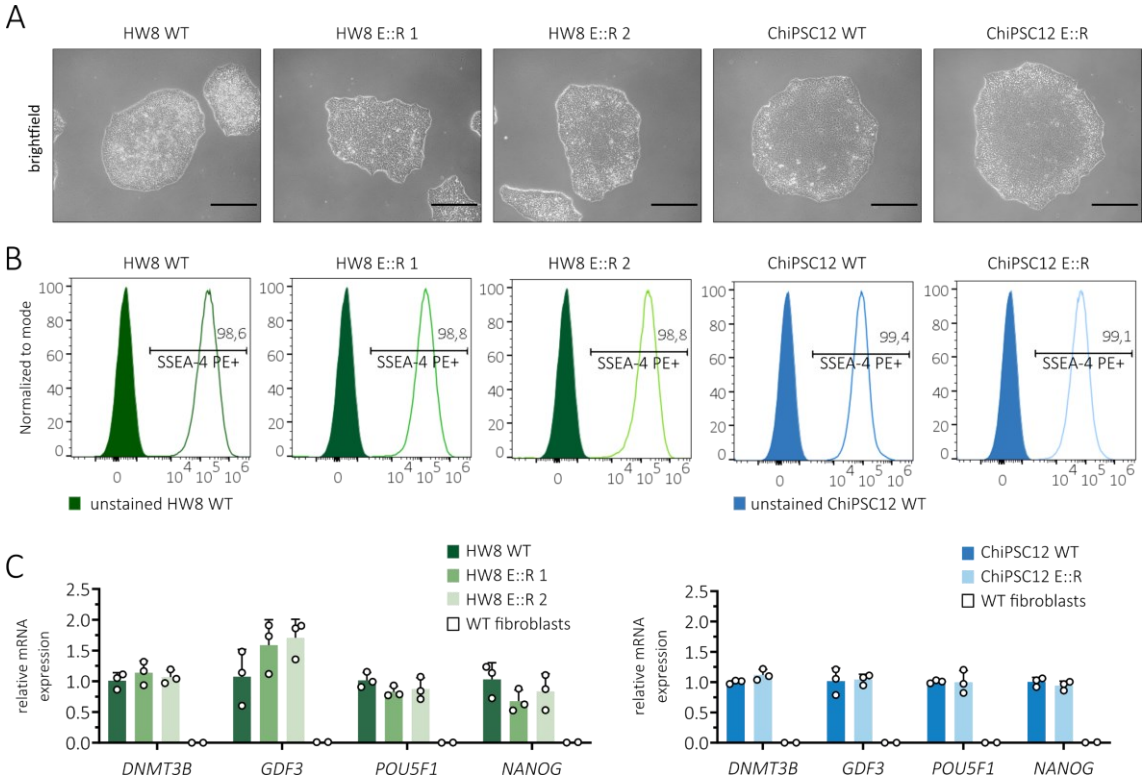


Figure 4.10: Assessment of hiPSC quality. (A) Representative brightfield images confirming normal hiPSC colony morphology (scalebar=300 µm). **(B)** Flow cytometric analysis of stage-specific embryonic antigen 4 (SSEA-4) on HW8 wild-type and ChiPSC12 wild-type cells, as well as CRISPR/Cas9-edited *ETV6::RUNX1*⁺ hiPSC lines HW8 E::R 1, HW8 E::R 2 and ChiPSC12 E::R. **(C)** Representative RT-qPCR analyses of pluripotency marker genes *DNMT3B*, *GDF3*, *POU5F1* and *NANOG* in HW8 wild-type, ChiPSC12 wild-type, as well as the respective *ETV6::RUNX1*⁺ hiPSC clones. Fibroblasts from 2 healthy donors were used as negative controls. Expression of *DNMT3B* and *GDF3* is normalized to *ATP5PB*, expression of *POU5F1* and *NANOG* is normalized to *PGK1*. Data is presented as the mean + SD, and expression of *ETV6::RUNX1*⁺ hiPSC lines is presented relative to the respective wild-type.

Table 4.1: Analysis of CRISPR/Cas9 off-target sites in *ETV6::RUNX1*⁺ hiPSCs. *In silico* prediction of off-target sites was performed using the Cas-OFFinder tool (http://www.rgenome.net/cas-offinder [220]). Off-target sites with ≤2 mismatches from the target sequences and bulge size of ≤2 were considered.

gene	location	predicted off-target sequence	mismatch detected?
intergenic region	chr1: 205039607	GG--GAGGCTAAATCCCaAAAGG	no
intergenic region	chr5: 12289845	GGATGAG--TAAATCCCTtAGGG	no
lncRNA ENSG00000250697	chr5: 32906797	GCCTAATgGGGAATGGTG-GTGG	no
UTRN (intron)	chr6: 144392358	GG--GAGGCTAgATCCCTAAAGG	no
NALF1 (intron)	chr13: 107849120	GGAaGAaGCTAAATCCCTAATGG	no

4.2 Transcriptome analysis of *ETV6::RUNX1*⁺ preleukemia

Following successful genome editing, bulk RNA sequencing (RNA-seq) of *ETV6::RUNX1*⁺ and wild-type hiPSCs was performed. Unsupervised principal component analysis (PCA) taking into account all genes identified by RNA-seq (n=16,328) clearly separated *ETV6::RUNX1*⁺ and wild-type hiPSCs according to genotype (**Figure 4.11A**). Of note, both *ETV6::RUNX1*⁺ HW8 hiPSC lines formed a single cluster, indicating high similarity of expression profiles. Altogether, expression changes were subtle with only 20 genes showing differential expression in all 3 *ETV6::RUNX1*⁺ hiPSC lines compared with their respective wild-type counterpart (cut-offs: absolute FC>2 and p<0.05, **Figure 4.11B-C**). These mild changes in expression underline the subtle oncogenic impact of *ETV6::RUNX1* and requirement of additional mutations for manifestation of overt leukemia.

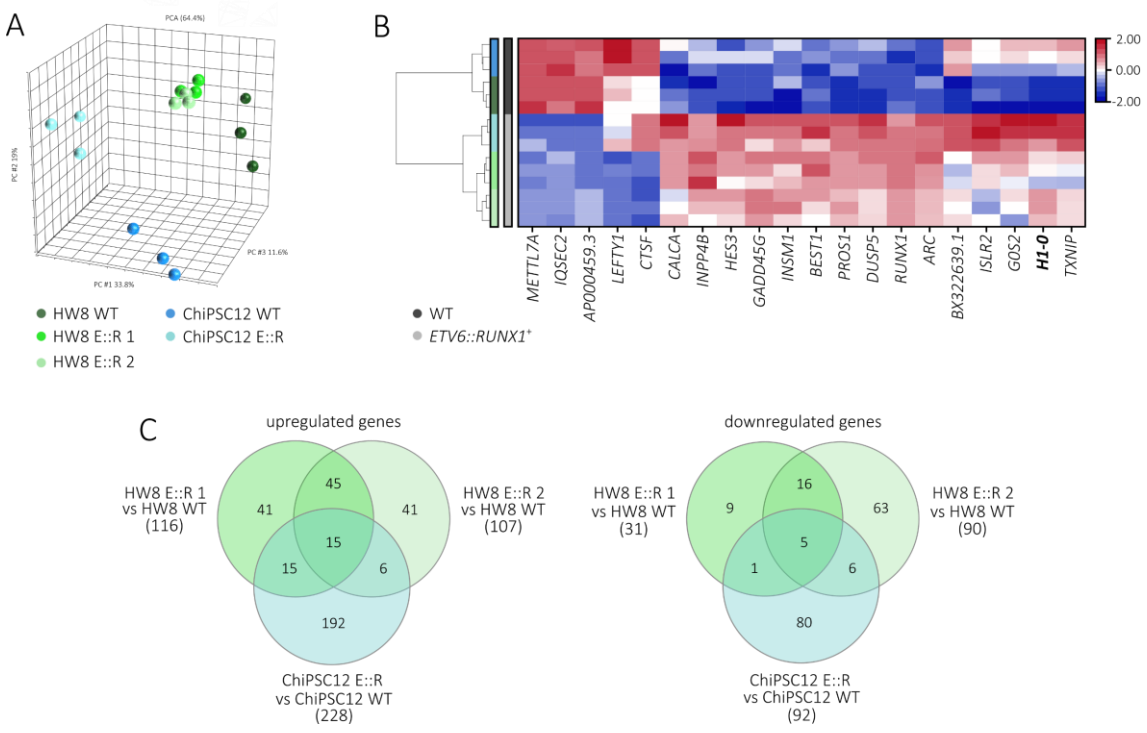


Figure 4.11: Transcriptome analysis reveals subtle expression changes induced by *ETV6::RUNX1* in hiPSCs. **(A)** Unsupervised principal component analysis (PCA) plot of *ETV6::RUNX1*⁺ and wild-type hiPSC transcriptome profiles determined by RNA-seq taking into account all genes. **(B)** Hierarchical clustering analysis of the 20 differentially expressed genes with an absolute FC>2 and p<0.05 detected in *ETV6::RUNX1*⁺ hiPSCs by RNA-seq compared to their respective wild-type counterpart. **(C)** Venn diagrams of upregulated and downregulated differentially expressed genes using the cut-offs mentioned in (B).

While expression changes were mild, *ETV6::RUNX1*⁺ hiPSCs displayed dysregulation of multiple previously identified ETV6::RUNX1 target genes (**Figure 4.12**). These include the direct

ETV6::RUNX1 target genes *RAG1* [133], *SPIB* [134] and *ICAM1* [221], as well as other genes previously found to be dysregulated upon *ETV6::RUNX1* knockdown, such as *S100A4* and *LYN* [130].

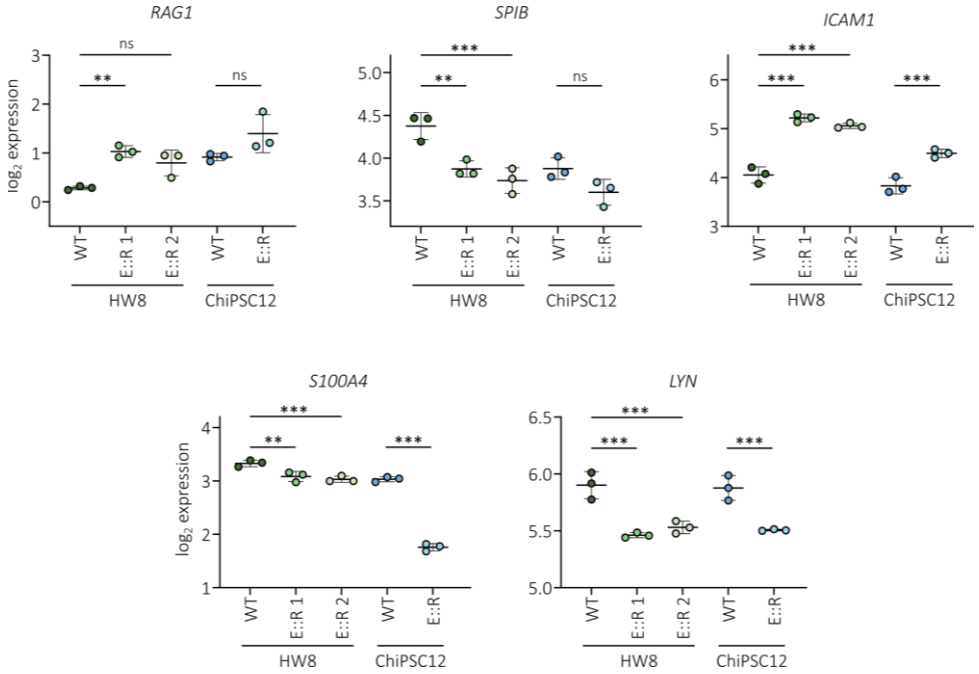


Figure 4.12: *ETV6::RUNX1*⁺ hiPSCs recapitulate dysregulation of common ETV6::RUNX1 target genes. RNA-seq expression of ETV6::RUNX1 target genes depicted in HW8 and ChiPSC12 hiPSCs. Data is shown as mean ± SD and analyzed for statistical significance using an ordinary one-way ANOVA. ** = p<0.01, *** = p<0.001

Given that the *ETV6::RUNX1* fusion gene likely arises in a hematopoietic precursor cell (HPC) [125], hiPSCs were differentiated towards HPCs using a serum- and feeder-free differentiation protocol (STEMdiff Hematopoietic kit, Stemcell Technologies). Differentiation was performed in two steps: hiPSCs were first induced towards mesodermal cells for three days and then differentiated for another nine days towards the hematopoietic lineage (**Figure 4.13**).

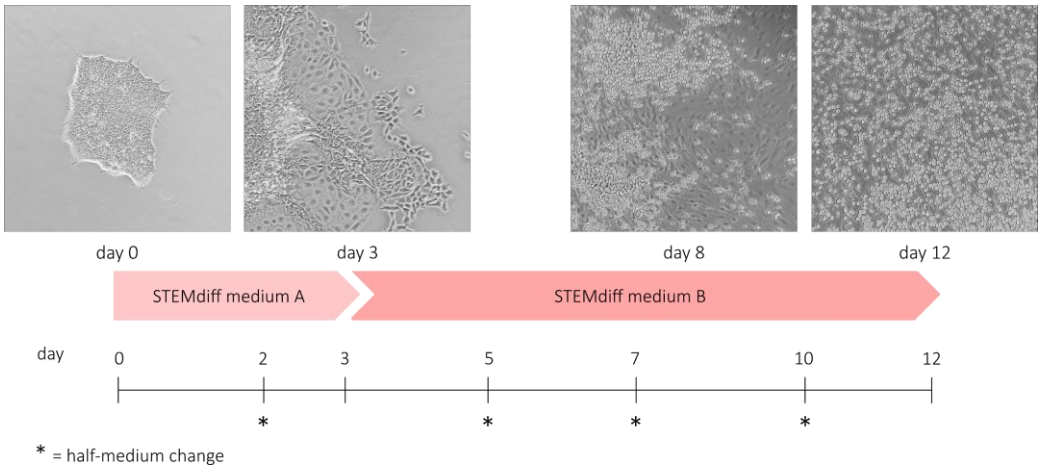


Figure 4.13: Overview of hiPSC differentiation towards HPCs. Representative brightfield images of HW8 wild-type hiPSCs during differentiation using the STEMdiff Hematopoietic kit (Stemcell Technologies) are shown.

ETV6::RUNX1 expression in the emerging HPCs was determined via RT-qPCR, indicating that fusion gene levels increased during hematopoietic differentiation to levels comparable to REH cells (**Figure 4.14A**). Additionally, single-cell RNA-seq (scRNA-seq) analyses of HPCs were performed to characterize cellular heterogeneity of *ETV6::RUNX1*$^+$ and wild-type HPCs. Flow cytometric analysis revealed significantly increased output of CD34$^+$ cells derived from *ETV6::RUNX1*$^+$ hiPSCs compared to wild-type hiPSCs (**Figure 4.14B-C**). In line with this, scRNA-seq identified upregulation of hematopoietic stem cell (HSC) marker genes *CD34*, *MECOM*, *NIRP1*, *NKAIN2* and *PROM1* in *ETV6::RUNX1*$^+$ hiPSC-derived HPCs (**Figure 4.14D**). Of note, while flow cytometry detected the CD34 surface protein on 78.8-97.4% of HPCs, the fraction of CD34$^+$ HPCs identified by scRNA-seq was much lower (11-22%). The discrepancy between RNA and surface protein levels is a common observation in scRNA-seq and has been previously decribed for CD34 [222].

Additionally, annotation of cell stages was performed by using gene set signatures derived from fetal bone marrow scRNA-seq data [205]. As expected, gene signatures of HSCs and multipotent progenitors (MPPs) were enriched in *ETV6::RUNX1*$^+$ HPCs compared to wild-type HPCs (**Figure 4.14E**).

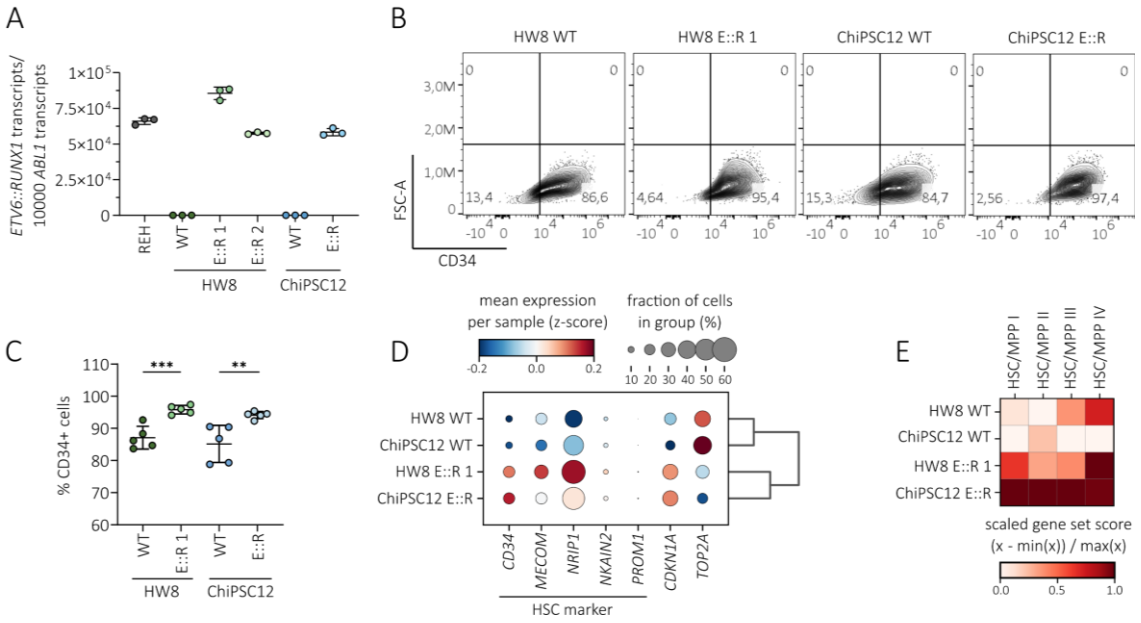


Figure 4.14: *ETV6::RUNX1* expression enhances HSC output during hiPSC differentiation. **(A)** Quantification of *ETV6::RUNX1* expression by RT-qPCR in REH cells and hiPSC-derived HPCs. Mean expression ± SD is indicated. **(B,C)** Frequencies of CD34$^+$ hiPSC-derived HPCs determined by flow cytometry. (B) Representative flow cytometry plots of wild-type and *ETV6::RUNX1*$^+$ hiPSCs. (C) Each dot represents a technical replicate (n=5) and replicates were pooled for scRNA-seq. Mean expression ± SD is indicated and data was analyzed for statistical significance using an unpaired t-test. ** = p<0.01, *** = p<0.001 **(D)** Dot plot showing mean expression per sample (z-score) of HSC marker genes *CD34*, *MECOM*, *NRIP1*, *NKAIN2* and *PROM1*, cell-cycle inhibitor *CDKN1A* (p21) and G2/M marker DNA topoisomerase *TOP2A* determined by scRNA-seq. Gene expression frequency (fraction of cells per sample) is indicated by spot size and expression level is indicated by color intensity. **(E)** Heat map visualizing the HSC/MPP I-IV scores of scRNA-seq data derived from *ETV6::RUNX1*$^+$ and wild-type HPCs using gene set signatures from Jardine *et al*. [205]. Analysis of scRNA-seq data was performed by Andrea Hanel.

ETV6::RUNX1 has been implicated in cell cycle repression [126]. Hence, annotation of cell cycle states was performed based on cell cycle marker gene scoring. Cell cycle distribution was skewed towards the G0/G1 state in *ETV6::RUNX1*$^+$ HPCs (**Figure 4.15A**), possibly due to increased numbers of quiescent CD34$^+$ cells. In line with this, cell cycle kinase inhibitor *CDKN1A* expression was elevated in *ETV6::RUNX1*$^+$ HPCs, while expression of the G2/M phase-specific DNA topoisomerase *TOP2A* was markedly enriched in wild-type HPCs (**Figure 4.14D**).

To measure overall transcriptional activity, the number of detected genes per cell (n_genes) and number of unique transcripts per cell (n_UMIs, UMI: unique molecular identifier) was determined. Both gene counts and unique transcripts per cell were significantly decreased in *ETV6::RUNX1*$^+$ HPCs compared to wild-type HPCs (**Figure 4.15B-C**).

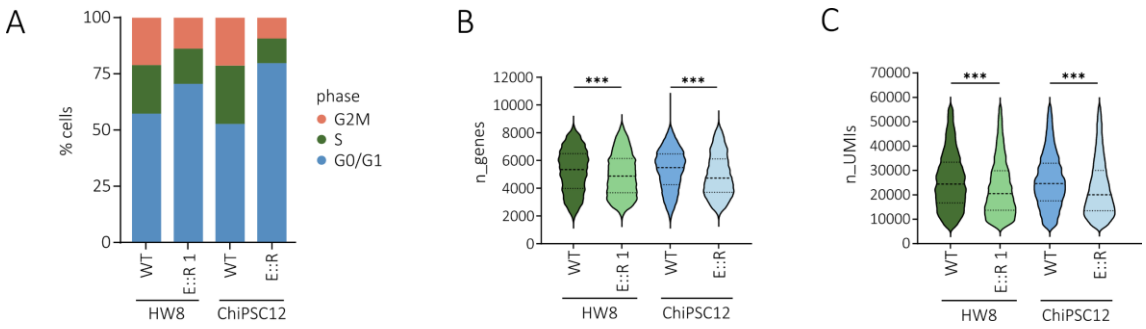


Figure 4.15: *ETV6::RUNX1* expression is associated with slower cell cycle and decreased transcriptional activity in hiPSC-derived HPCs. (A) Stacked bar plot showing distribution of cell cycle stages of samples analyzed by scRNA-seq. Analysis of scRNA-seq data was performed by Andrea Hanel. **(B, C)** Violin plots depicting number of detected genes (n_genes) and unique molecular identifier counts (n_UMIs) per cell. Median expression and quartiles are indicated. Data was analyzed for statistical significance using an unpaired t-test. *** = p<0.001

4.3 Linker histone *H1-0* upregulation in *ETV6::RUNX1*⁺ preleukemia and BCP-ALL

Transcriptome analysis of *ETV6::RUNX1*⁺ hiPSCs revealed consistent upregulation of linker histone *H1-0* (**Figure 4.11B**), an epigentic regulator that mediates chromatin remodeling [183]. Elevated expression levels of *H1-0* identified by RNA-seq (**Figure 4.16A**) were confirmed both by RT-qPCR (**Figure 4.16B**) and immunoblot (**Figure 4.16C**). Again, ETV6 protein levels were markedly reduced in *ETV6::RUNX1*⁺ hiPSCs as a result of the *RUNX1* HDR template disrupting one *ETV6* allele.

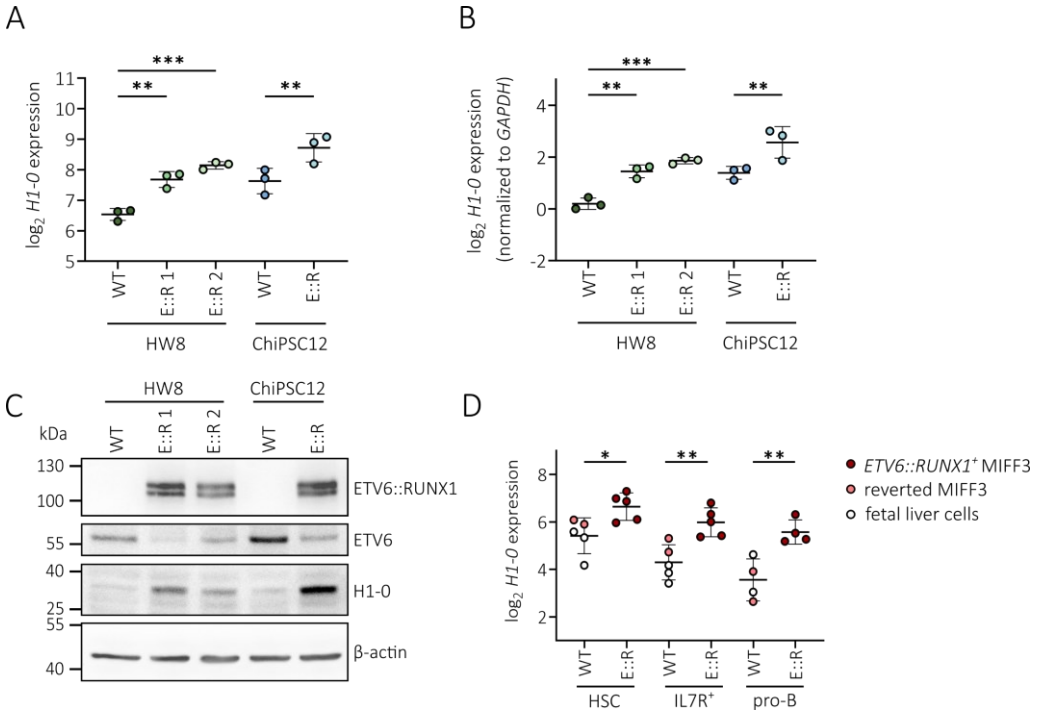


Figure 4.16: Human *ETV6::RUNX1*⁺ preleukemia models exhibit upregulation of linker histone *H1-0*. **(A)** *H1-0* expression levels determined by RNA-seq in *ETV6::RUNX1*⁺ and wild-type hiPSCs. **(B)** *H1-0* expression levels determined by RT-qPCR in *ETV6::RUNX1*⁺ and wild-type hiPSCs. Values are normalized to HW8 wild-type expression levels and to *GAPDH* expression. Samples shown in (B) were subjected to RNA-seq. **(C)** Western blot analysis of ETV6::RUNX1, H1-0, ETV6 and β-actin levels in *ETV6::RUNX1*⁺ and wild-type hiPSCs. **(D)** H1-0 levels in HSCs ($CD19^-CD34^+CD45RA^-$), $IL7R^+$ ($CD19^-CD34^+CD45RA^+$ $IL7R^+$) and pro-B ($CD19^+CD34^+$) cells differentiated from *ETV6::RUNX1*⁺ or reverted MIFF3 hiPSCs, and fetal liver cells. Data is derived from a published RNA-seq dataset (accession number: E-MTAB-6382 [125]). **(A, B, D)** Mean expression ± SD is indicated and data was analyzed for statistical significance using an ordinary one-way ANOVA. * = p<0.05, ** = p<0.01, *** = p<0.001

Furthermore, analysis of a published RNA-seq dataset (accession number: E-MTAB-6382 [125]) revealed that upregulation of *H1-0* in *ETV6::RUNX1*⁺ preleukemic cells is preserved during differentiation of hiPSCs along the B lymphoid lineage and *H1-0* levels were restored to a

degree comparable to sorted fetal liver cells in reverted hiPSC-derived cells that lack *ETV6::RUNX1* expression (**Figure 4.16D**).

To assess *H1-0* levels in overt leukemia, RNA-seq and expression microarray datasets derived from three patient cohorts, encompassing a total of 3,026 samples (PeCan St. Jude database, GSE87070 [199] and GSE13159 [198]) were analyzed. Across all cohorts, *ETV6::RUNX1*$^{+}$ BCP-ALL showed highest *H1-0* levels (**Figure 4.17**).

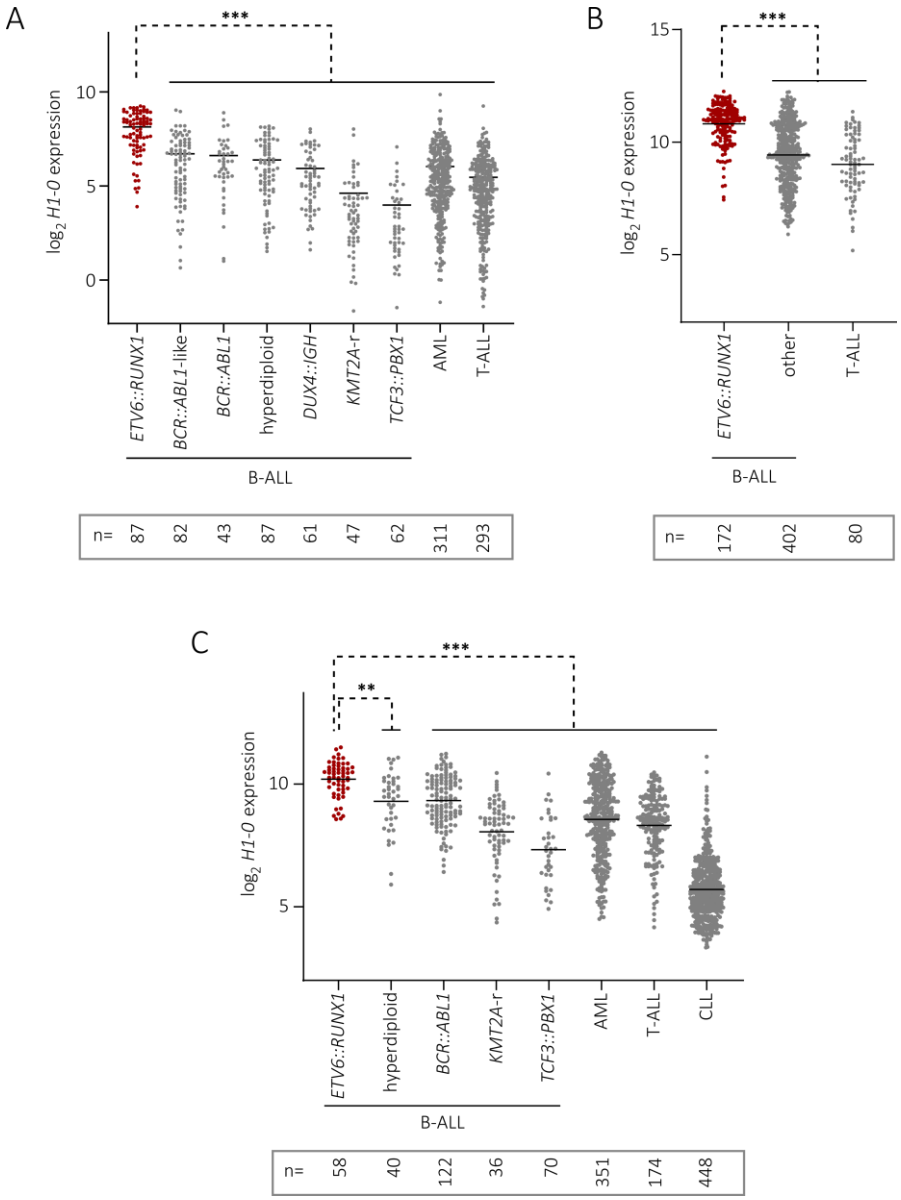


Figure 4.17: *H1-0* levels are consistently upregulated in *ETV6::RUNX1*$^{+}$ BCP-ALL patients. (A-C) *H1-0* levels across various ALL patient cohorts derived from the (A) PeCan St. Jude database and two published expression microarray datasets (accession numbers: GSE87070 [210] (B) and GSE13159 [208] (C)). Number of patients per leukemia entity is indicated. Mean expression is indicated and data was analyzed for statistical significance using an ordinary one-way ANOVA. ** = p<0.01, *** = p<0.001

Additionally, 9 patient-derived xenograft (PDX) samples were analyzed for *H1-0* expression via RT-qPCR. In line with my previous observations, *ETV6::RUNX1*$^+$ PDX samples displayed highest *H1-0* expression compared to high hyperdiploid, *BCR::ABL1*$^+$ or *TCF3::PBX1*$^+$ BCP-ALL (**Figure 4.18A**). Moreover, immunoblot analysis of B-ALL cell lines was performed to determine H1-0 expression levels (**Figure 4.18B**). The *ETV6::RUNX1*$^+$ cell line REH and high hyperdiploid cell line MHH-CALL-2 (derived from a near-haploid clone) showed particluarly high H1-0 protein expression, while cell lines bearing other chromosomal translocations exhibited low or undetectable levels of H1-0. Moreover, H1-0 protein levels were found to reflect doubling times in the B-ALL cell lines.

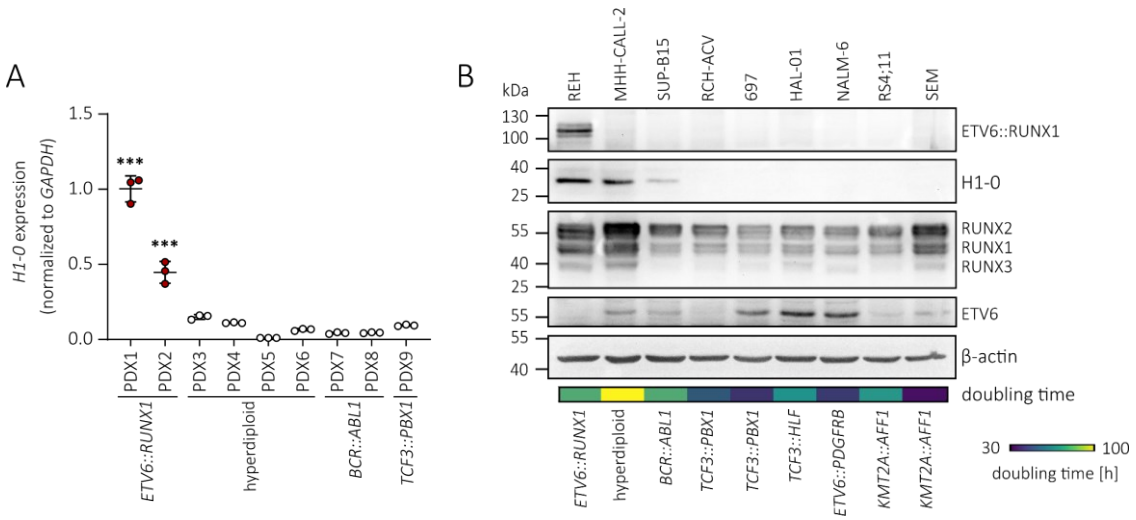


Figure 4.18: *H1-0* levels are consistently upregulated in *ETV6::RUNX1*$^+$ PDX samples and cell lines. **(A)** *H1-0* expression was quantified by RT-qPCR. Technical replicates (n=3) are shown with mean expression ± SD. Statistical significance of H1-0 upregulation in *ETV6::RUNX1*$^+$ PDX was determined using an ordinary one-way ANOVA. *** = p<0.001 **(B)** Protein levels of ETV6::RUNX1, H1-0, RUNX variants, ETV6 and β-actin in BCP-ALL cell lines quantified by Western blot. Chromosomal aberrations present in the respective cell lines are indicated. Doubling times are derived from the DSMZ database (https://www.dsmz.de).

4.4 Regulation of *H1-0* expression

Given the association between *ETV6::RUNX1* and *H1-0* expression, direct binding of the ETV6::RUNX1 transcription factor to the *H1-0* promoter region might be conceivable. However, reanalysis of the *H1-0* promoter using published chromatin immunoprecipitation sequencing (ChIP-seq) data of *ETV6::RUNX1*$^{+}$ REH cells (accession numbers: GSE176084 [215] and GSE117684 [133]) showed no direct binding of either the fusion protein or RUNX1 to the *H1-0* promoter region (**Figure 4.19**).

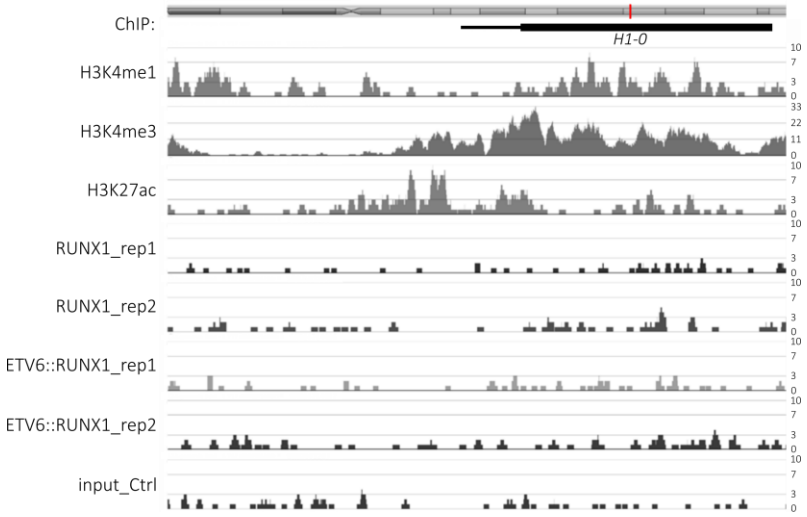


Figure 4.19: Chromatin immunoprecipitation does not show direct binding of ETV6::RUNX1 or RUNX1 to the *H1-0* promoter region. Chromatin immunoprecipitation (ChIP) peak visualization within the human H1-0 gene region (hg38) from ChIP-seq data of REH cells for H3K4me1, H3K4me3, H3K27ac and RUNX1 (accession number: GSE117684 [215]) as well as ETV6::RUNX1 (accession number: GSE176084 [129]).

To determine the potential of ETV6::RUNX1 to transactivate the *H1-0* promoter, dual-luciferase promoter assays were performed. For this, the *H1-0* promoter region (-351 to +161 from TSS) was cloned into a luciferase reporter plasmid (**Figure 4.20A**), which was transfected into 293T cells along with either an empty vector or vectors containing the FLAG-tagged *ETV6::RUNX1* or *RUNX1* sequence. Luciferase activity measurements confirmed that expression of *ETV6::RUNX1* is sufficient to activate the *H1-0* promoter (2.2-fold), while *RUNX1* expression induces reduction of luciferase activity (3.1-fold; **Figure 4.20B**).

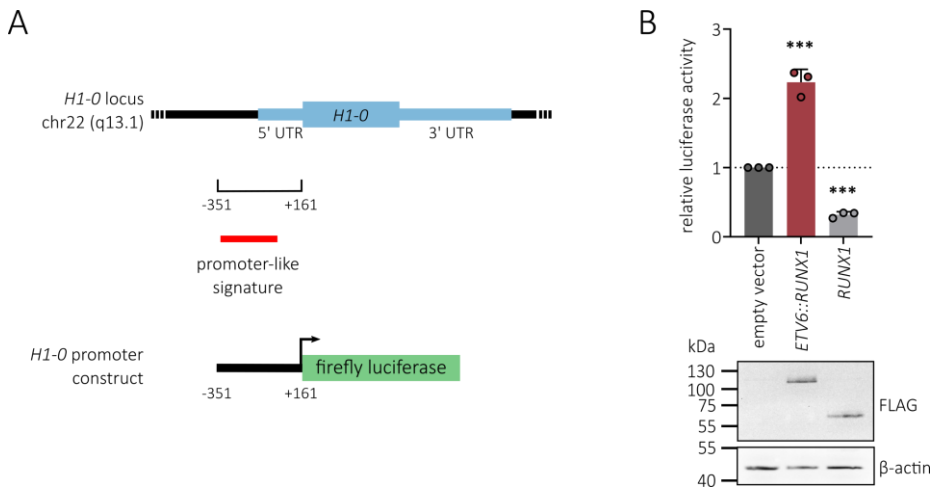


Figure 4.20: ETV6::RUNX1 activates the *H1-0* promoter. (A) Schematic representation of the *H1-0* locus, including the 512-bp region (nucleotides -351 to +161 from TSS) encompassing promoter-like signature EH38E2163184 (ENCODE). The *H1-0* CpG island (CGI) shore and 450K Infinium array probes are indicated. **(B)** 293T cells were transfected with a vector encoding the *H1-0* promoter-like signature indicated in (A), together with the empty pcDNA3 vector or pcDNA3 plasmids expressing either *ETV6::RUNX1* or *RUNX1*, and a vector expressing Renilla luciferase. Luciferase activities were normalized to Renilla luciferase activity and the empty vector control. Data represent mean values of three independent replicates + SD. Significance was calculated using an ordinary one-way ANOVA (*** = p<0.001). Representative protein levels of ETV6::RUNX1, RUNX1 and β-actin determined by Western blot are shown.

Apart from direct binding, *H1-0* expression might be indirectly affected by ETV6::RUNX1. Indirect regulation of *H1-0* expression is conceivable via epigenetic regulation, such as methylation or acetylation. Recently, differential DNA methylation of the *H1-0* CpG island (CGI) shore has been implicated in regulating *H1-0* expression [183]. A schematic representation of the *H1-0* CGI shore region is depicted in **Figure 4.21A**. The 450K Infinium probes covering the indicated genomic locus are indicated.

To analyze differential methylation of the *H1-0* CGI shore region, publicly available 450K Infinium microarray DNA methylation data comprising patient samples of T-ALL (n=101) and six B-ALL subtypes (n=445; accession number: GSE49032 [158]) were analyzed. These analyses revealed that mean *H1-0* CGI shore methylation comprising probes cg07141002 and cg01883777 was lowest in *ETV6::RUNX1*$^{+}$ ALL (**Figure 4.21B**) and inversely correlated with *H1-0* RNA expression (Pearson r=0.645, p<0.0001; **Figure 4.21C**).

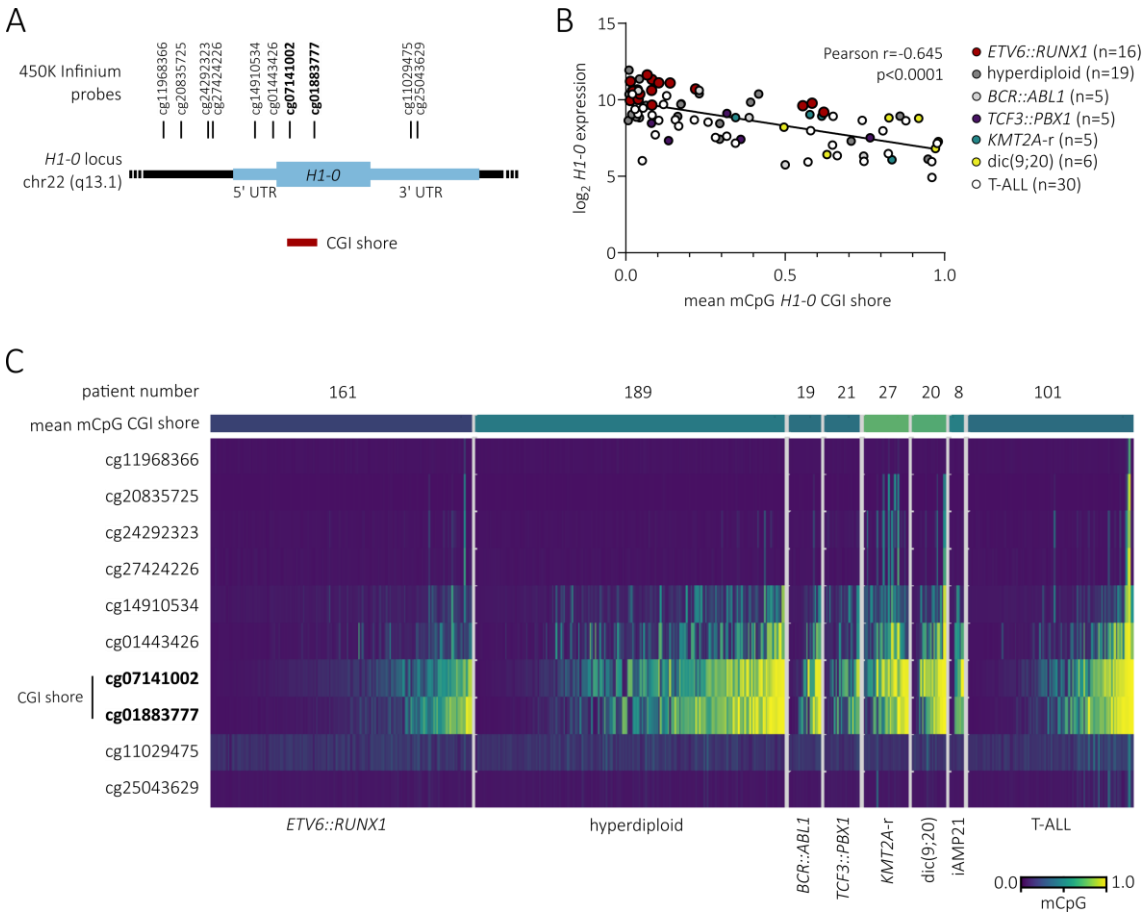


Figure 4.21: *H1-0* RNA expression is inversely correlated with *H1-0* CGI shore methylation in ALL patient samples. (A) Schematic representation of the *H1-0* CGI shore region and position of 450K Infinium microarray DNA probes. Adapted from [183]. **(B)** Pearson correlation of *H1-0* RNA expression and mean DNA methylation of the *H1-0* CGI shore probes cg07141002 and cg01883777 in leukemia patients. Data was retrieved from NCBI GEO (accession number: GSE49032 [158]). Expression is shown for microarray probe 208886_at. Each dot represents a single patient. **(C)** *H1-0* DNA methylation in different leukemia entities is visualized as a heatmap with each column corresponding to a single patient. Within each entity, patients are sorted according to mean DNA methylation of CGI shore probes cg07141002 and cg01883777. The total number of patients per entity is indicated. Data was retrieved from NCBI GEO (accession number: GSE49032 [158]) and processed by Dr. Rabea Wagener.

H1-0 expression can be induced by various extrinsic and intrinsic signals. Early accounts of *H1-0* accumulation via the HDACis sodium butyrate [223] and Trichostatin A (TSA) [224] treatment suggested a potential role of histone acetylation in modulating *H1-0* expression. Induction of *H1-0* expression by HDACis was confirmed by Morales Torres *et al*. in a screen of >4000 compounds in HCC1569 breast cancer cells and tumor differentiation factor (TDF)-transformed dermal fibroblasts [194]. This screen identified the three HDACis JNJ-26481585/Quisinostat, PCI-24781/Abexinostat and suberanilohydroxamic acid (SAHA)/Vorinostat as potent inducers of *H1-0* expression.

Hence, induction of *H1-0* after 24-hour treatment with these three HDACis was tested in hiPSCs. As can be observed in **Figure 4.22A**, all compounds led to significant increase of *H1-0* RNA, albeit with different potencies. Importantly, drug doses of 1 µM Abexinostat and ≥100 nM Quisinostat led to lethality. Among the three drugs, Quisinostat showed highest potency, leading to significantly elevated *H1-0* RNA levels at low nanomolar concentrations of 5-15 nM after 24 hours compared to DMSO-treated hiPSCs (**Figure 4.22B**).

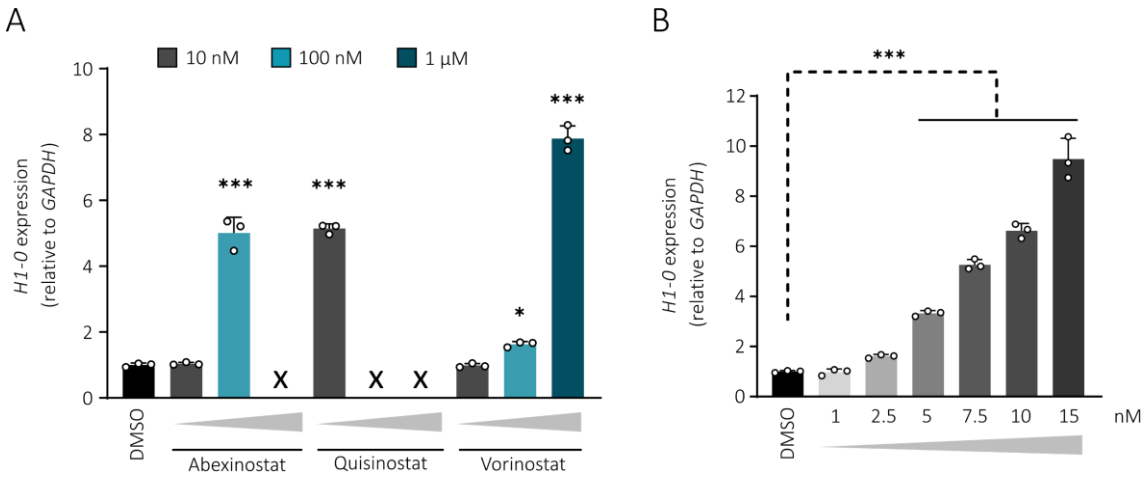


Figure 4.22: Quisinostat induces *H1-0* expression at low nanomolar concentrations in hiPSCs. HW8 hiPSCs were treated with the indicated concentrations of **(A)** Abexinostat, Quisinostat or Vorinostat, and **(B)** Quisinostat for 24 hours. *H1-0* levels were quantified by RT-qPCR and normalized to *GAPDH* levels. Values represent technical replicates (n=3) and mean expression + SD is indicated. One-way ANOVA was used to compare DMSO-treated and HDACi-treated hiPSCs. * = p<0.05, *** = p<0.001

Next, the effect of Quisinostat treatment on transcriptional activation via the *H1-0* promoter was assessed by transfection of 293T cells with the exogenous *H1-0* promoter sequence cloned upstream of a luciferase reporter gene (as previously shown in **Figure 4.20A**). Quisinostat induces dose-dependent activation of transcription via the *H1-0* promoter (**Figure 4.23A**) and induces expression of endogenous *H1-0* RNA levels in 293T cells, as measured by RT-qPCR (**Figure 4.23B**).

As shown in **Figure 4.18**, Western blot analysis of H1-0 protein levels in BCP-ALL cell lines indicated high expression in the *ETV6::RUNX1*$^{+}$ cell line REH as well as the hyperdiploid cell line MHH-CALL-2, while cell lines bearing other chromosomal translocations displayed low (SUP-B15) or undetectable levels. To test whether treatment with Quisinostat induces *H1-0* expression in leukemic cells of the B lineage, BCP-ALL cell lines with high (REH and MHH-CALL-2), medium (SUP-B15) and low *H1-0* expression (RS4;11) were treated with 1 µM

Quisinostat for 24 hours. All cell lines showed strong upregulation of *H1-0* upon Quisinostat treatment, dependent on their basal levels (**Figure 4.24**). The increase of *H1-0* expression observed upon Quisinostat treatment was highest in RS4;11 (343.7-fold), indicating that Quisinostat can efficiently induce *H1-0* expression in B-ALL cells with low basal H1-0 expression levels. Of note, expression of H1-0 could not be detected in wild-type RS4;11 on Western blot level (**Figure 4.18**), while the *H1-0* RT-qPCR approach was able to pick up low *H1-0* expression with high sensitivity.

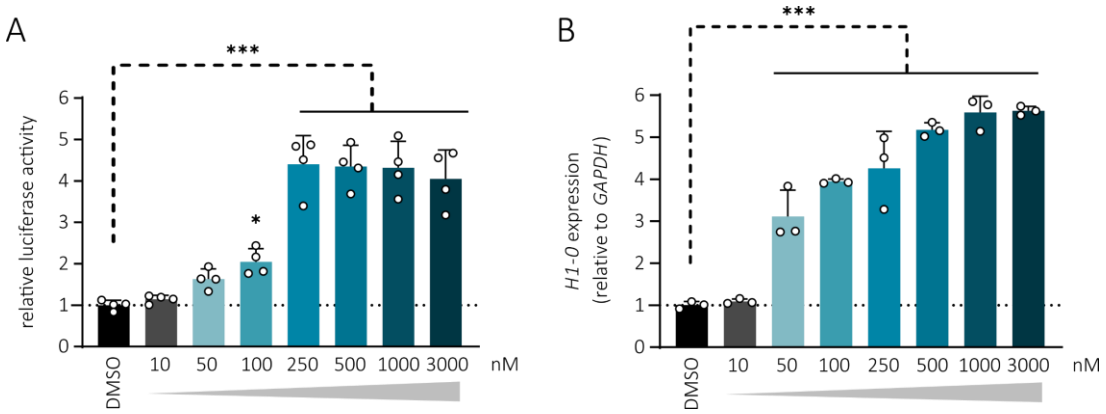


Figure 4.23: Quisinostat activates the *H1-0* promoter at nanomolar concentrations. 293T cells were transfected with vectors encoding for the *H1-0* promoter-like signature upstream of Firefly luciferase and for Renilla luciferase. **(A)** Luciferase activities were normalized to Renilla luciferase activity and the empty vector control. **(B)** *H1-0* expression was quantified by RT-qPCR and normalized to *GAPDH* levels. Data represent mean values of three independent replicates + SD. Significance was calculated using an ordinary one-way ANOVA. * = p<0.05, *** = p<0.001

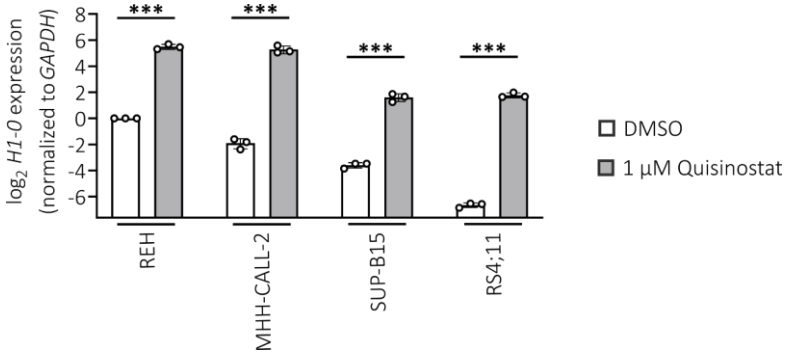


Figure 4.24: Quisinostat induces strong *H1-0* upregulation in BCP-ALL cell lines. RT-qPCR quantifying *H1-0* levels 24 hours after treatment with DMSO or 1 µM Quisinostat in BCP-ALL cell lines. Values represent mean ± SD from three independent replicates and data was analyzed for statistical significance using an ordinary one-way ANOVA. *** = p<0.001.

4.5 Functional characterization of linker histone *H1-0*

4.5.1 *H1-0* knockdown in REH cells

Given that *H1-0* is specifically upregulated in preleukemia and overt BCP-ALL carrying the *ETV6::RUNX1* translocation, I aimed to characterize the contribution of *H1-0* to *ETV6::RUNX1*$^+$ BCP-ALL pathology. To this end, a transient knockdown of *H1-0* was performed in the *ETV6::RUNX1*$^+$ B-ALL cell line REH using siRNA pools delivered by nucleofection. siRNA-mediated knockdown resulted in reduction of *H1-0* RNA levels by ≈50-60% compared to non-targeting control siRNA (siCtrl) treatment after 48 h (**Figure 4.25A-B**). Knockdown was confirmed on protein level by performing Western blot analysis (**Figure 4.25C**).

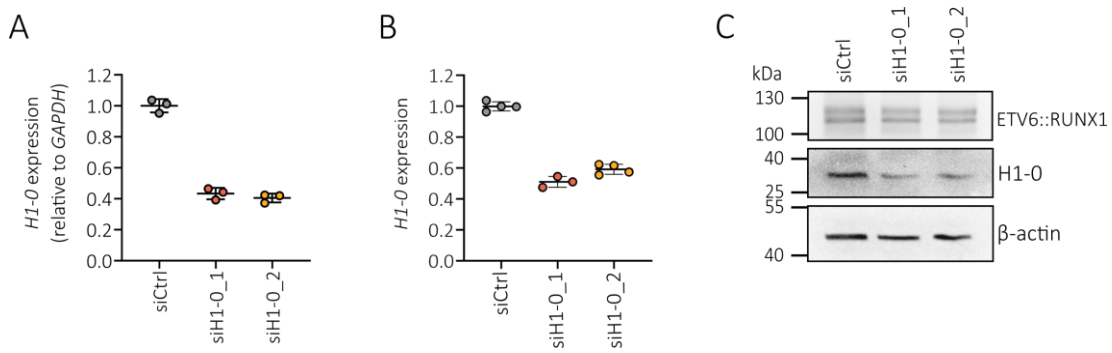


Figure 4.25: siRNA-mediated knockdown of *H1-0* in REH cells. (A) Relative *H1-0* expression determined by RNA-seq in REH cells treated for 48 hours with an siRNA control pool (siCtrl) or *H1-0*-targeting siRNA pools siH1-0_1 or siH1-0_2. Values are normalized to *H1-0* expression levels in siCtrl-treated cells. Data is presented as the mean ± SD. **(B)** *H1-0* expression levels determined by RT-qPCR in REH cells treated for 48 hours with siCtrl or *H1-0*-targeting siRNA pools. Values are normalized to *GAPDH* expression and *H1-0* expression levels in siCtrl-treated cells. Samples measured in (B) are derived from an experiment independent of samples shown in (A). Data is presented as the mean ± SD. **(C)** Representative Western blot analysis of REH lysates following siRNA-mediated knockdown of *H1-0* for 48 hours, detecting ETV6::RUNX1, H1-0 and β-actin.

Transcriptome analysis by RNA-seq of REH cells treated with *H1-0*-targeting siRNA compared to siCtrl treatment revealed moderate gene dysregulation, with 38 genes showing recurrent >1.5 fold changes ($p<0.05$, 23 upregulated and 15 downregulated, **Figure 4.26A**). Gene set enrichment analysis (GSEA) of dysregulated genes identified by RNA-seq upon *H1-0* knockdown revealed significant enrichment (cut-offs: $p<0.005$, false discovery rate (FDR) q-value<0.1) of gene signatures associated with DNA replication, histone modification, DNA repair and protein ubiquitination in siCtrl-treated REH cells (**Figure 4.26B**, **Supplementary Table 8.1**), while no significantly enriched pathways were identified in REH cells treated with *H1-0*-targeting siRNA using the same cut-offs (**Supplementary Table 8.2**),. Notably, GSEA uncovered dysregulation of

genes linked to histone acetylation and methylation (**Supplementary Table 8.1**), consistent with previous reports highlighting strong correlation between *H1-0* gene expression and chromatin methylation or acetylation [194, 224].

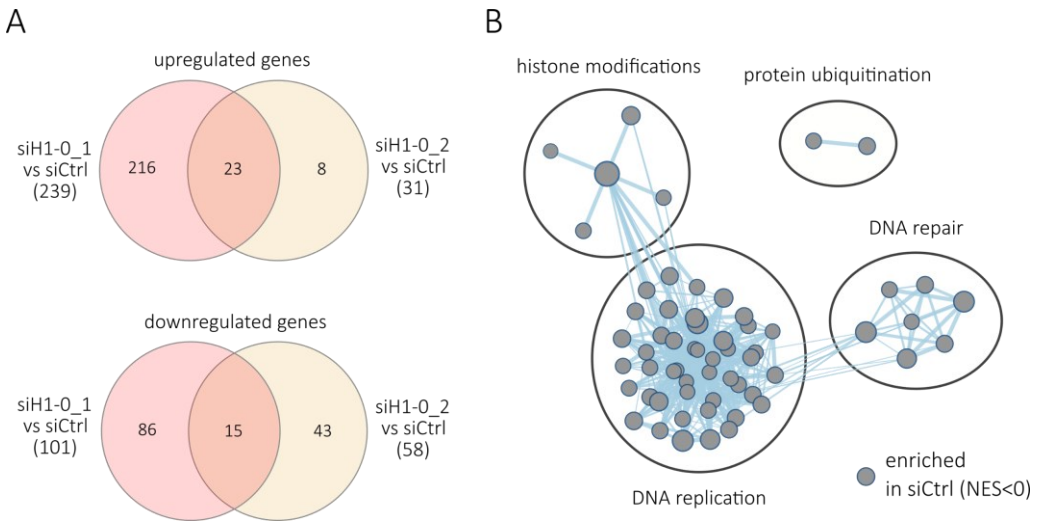


Figure 4.26: Canonical signaling pathways enriched in REH cells upon siRNA-mediated *H1-0* knockdown. **(A)** Significantly dyregulated genes in REH cells treated with *H1-0*-targeting siRNA for 48 hours compared to siCtrl treatment (cut-offs: absolute FC>1.5 and p<0.05) determined by RNA-seq. **(B)** Enrichment map of gene sets enriched in siCtrl-treated REH cells compared to siRNA-mediated knockdown of *H1-0* (cut-offs: p<0.005, FDR q-value<0.1) using the canonical pathways gene set collection (Human MSigDB Collections). No significantly enriched gene sets were found in siH1-0-treated REH cells using the indicated cut offs. Groups of similar pathways are indicated.

To ascertain common molecular drivers of the observed gene expression changes upon *H1-0* knockdown, upstream regulator analysis (Ingenuity Pathway Analysis, QIAGEN) was performed (**Supplementary Tables 8.3** and **8.4**). Notably, the most significant potential drivers of expression changes were *ETV6::RUNX1* and *TP53* (**Figure 4.27A**). Negative activation z-scores indicate predicted inhibition of the ETV6::RUNX1 transcription factor upon *H1-0* knockdown. Given that ETV6::RUNX1 primarily functions as a repressor of RUNX1 target genes [128], a set of genes upregulated upon ETV6::RUNX1 knockdown (cut-offs: $\log_2$ FC>0.9 and p<0.05) [130] was used to validate the previous findings. Indeed, GSEA revealed significant upregulation of *ETV6::RUNX1* signature genes upon *H1-0* knockdown (normalized enrichment score (NES)=1.63, FDR q-value=0.001; **Figure 4.27B**).

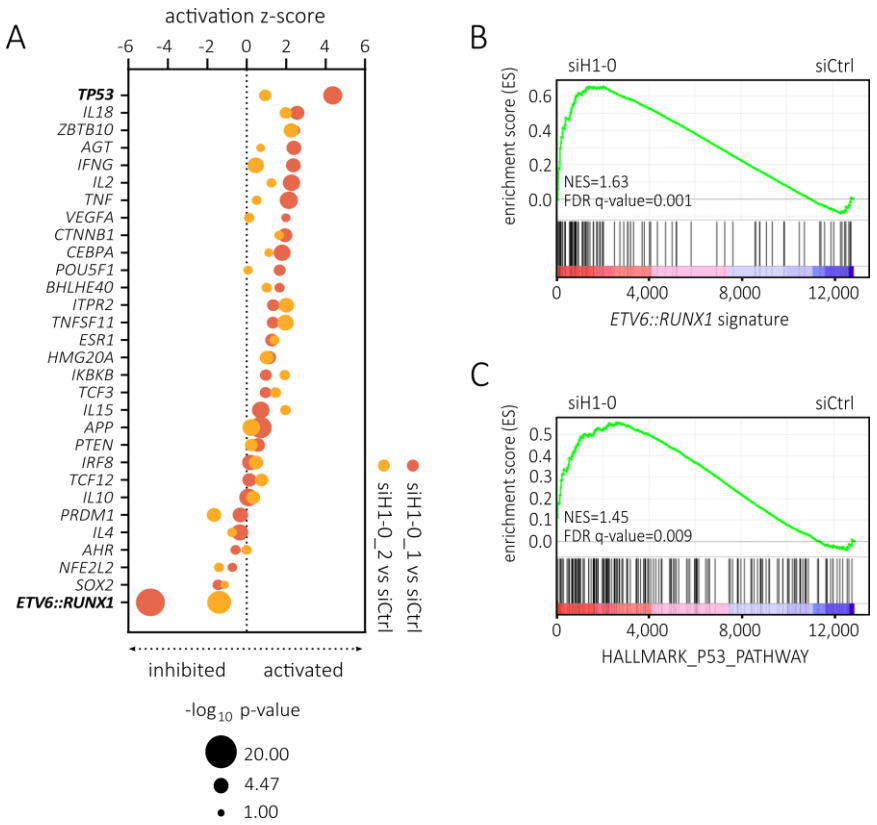


Figure 4.27: *ETV6::RUNX1*-specfic gene signature is reverted in REH cells upon siRNA-mediated knockdown of *H1-0*. Ingenuity Pathway analysis (IPA) of upstream regulators significantly enriched in both siH1-0_1 versus siCtrl and siH1-0_2 versus siCtrl ($p<0.05$). The *ETV6::RUNX1* signature is marked in red. IPA was performed by Daniel Picard. (B, C) Gene set enrichment analysis (GSEA) results of (B) siH1-0_1 versus siCtrl and (C) siH1-0_2 versus siCtrl using a published gene set of significantly upregulated genes in REH and AT-2 cells upon *ETV6::RUNX1* knockdown (cut-offs: $\log_2$ FC>0.9 and $p<0.05$, [130]). Normalized enrichment score (NES) and false discovery rate (FDR) are indicated.

Significant activation of *TP53* (encoding for p53) following *H1-0* knockdown was confirmed by GSEA (NES=1.45, FDR q-value=0.009; **Figure 4.27C**). Indeed, previous studies have demonstrated that ETV6::RUNX1 suppresses p53 activity by upregulating *MDM2* [225], which I found to be downregulated in REH cells upon *H1-0* knockdown (**Figure 4.28A**). Moreover, both *EPOR* and *RAG1*, two genes upregulated by ETV6::RUNX1 and imperative to *ETV6::RUNX1*$^+$ BCP-ALL pathophysiology [14, 133, 135, 226], exhibited reduced levels upon *H1-0* knockdown (**Figure 4.28A**) as well as significant correlation with *H1-0* RNA expression in *ETV6::RUNX1*$^+$ BCP-ALL patient samples derived from the PeCan St. Jude cohort (n=87; https://pecan.stjude.cloud [206, 207]; **Figure 4.28B**).

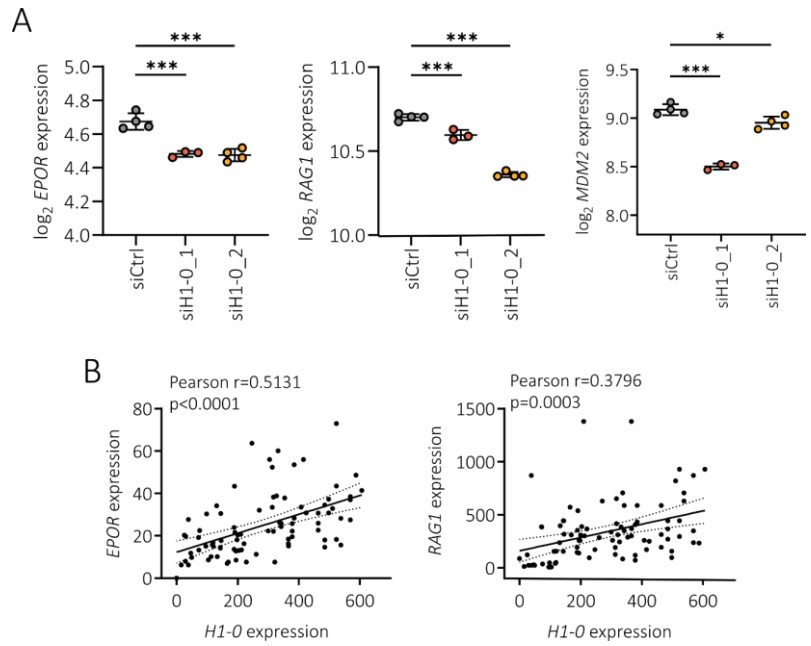


Figure 4.28: Dysregulation of ETV6::RUNX1-regulated signature genes upon *H1-0* knockdown. (A) RNA expression levels of *EPOR*, *RAG1* and *MGM2* determined by RNA-seq in siCtrl and siH1-0 REH. **(B)** Pearson correlation of *H1-0* expression with *EPOR* or *RAG1* expression in *ETV6::RUNX1*$^{+}$ BCP-ALL patient samples derived from the PeCan St. Jude Cloud (n=87, RNA-seq data, https://pecan.stjude.cloud [206, 207]).

4.5.2 Correlation analyses in healthy and leukemic cells

To gain an understanding of *H1-0* function, tissue expression was analyzed using a publicly-available RNA-seq tissue expression dataset (Human Protein Atlas portal, HPA tissue dataset, https://www.proteinatlas.org). Notably, bone marrow exhibited highest *H1-0* expression levels compared to 32 other tissues (sorted according to *H1-0* expression level, Figure 4.29).

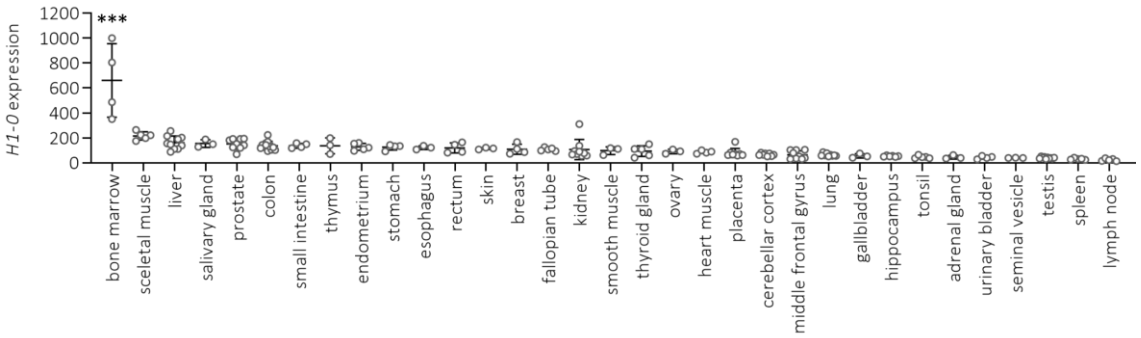


Figure 4.29: *H1-0* is highly expressed in human bone marrow. Tissue-based *H1-0* RNA-seq data (nTPM) obtained from the Human Protein Atlas portal depicting the 33 tissues with highest mean expression of *H1-0*. Tissue types are sorted according to mean expression. Each dot represents a single patient. Data is shown as mean ± SD. *** = p<0.001

Given that *H1-0* has previously been shown to be restricted to undifferentiated, quiescent progenitor cells in hematopoiesis [187], *H1-0* expression was analyzed in hematopoietic precursor cells along the B cell lineage. Analysis of published RNA-seq and expression microarray datasets revealed a stepwise decrease of mean *H1-0* expression during early B cell development in bone marrow (Figure 4.30A), umbilical cord blood and peripheral blood (Figure 4.30B), as well as fetal liver-derived cells (Figure 4.30C). Moreover, *H1-0* expression levels of *ETV6::RUNX1*$^{+}$ BCP-ALL patient samples were compared to healthy HSC and B cell progenitors stages. *ETV6::RUNX1*$^{+}$ BCP-ALL samples showed significantly higher expression of *H1-0* compared to early hematopoietic cell stages (Figure 4.30A).

H1-0 has previously been identified as a heterogenously expressed gene in human solid tumors [194]. To address cell-to-cell variation, scRNA-seq data of B cell precursor cells derived from healthy bone marrow of 8 donors was analyzed. Again, *H1-0*$^{+}$ cell numbers decreased along the B lineage trajectory, clustering predominantly to more immature and G0/G1 cell states (Figure 4.31).

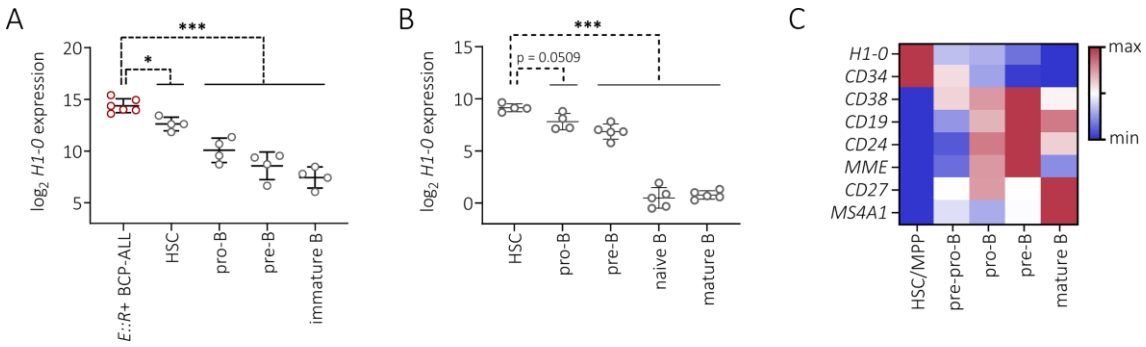


Figure 4.30: Expression of *H1-0* decreases during hematopoietic differentiation in healthy individuals. **(A)** *H1-0* expression in *ETV6::RUNX1*$^+$ BCP-ALL (n=6) and healthy B cell precursor stages derived from a published RNA-seq dataset (accession number: GSE115656). B cell precursor fractions are HSCs ($CD34^+CD19^-IgM^-$), pro-B cells ($CD34^+CD19^+IgM^-$), pre-B cells ($CD34^-CD19^+IgM^-$) and immature B cells ($CD34^-CD19^+IgM^+$). **(B)** *H1-0* expression in healthy B cell precursor stages derived from a published expression microarray dataset (accession number: GSE24759). B cell precursor fractions are HSCs ($CD34^+CD38^-$), pro-B cells ($CD34^+CD10^+CD19^+$), pre-B cells ($CD34^-CD10^+CD19^+$), naïve B cells ($CD19^+IgD^+CD27^-$) and mature B cells ($CD19^+IgD^+CD27^+$). **(C)** Min-max-normalized RNA expression of *H1-0* is depicted in healthy fetal liver B cell developmental stages obtained from a published scRNA-seq dataset (accession number: E-MTAB-7407, Developmental Cell Atlas Newcastle University, https://www.humancellatlas.org/). (A, B) Mean expression ± SD is indicated and data was analyzed for statistical significance using an ordinary one-way ANOVA (* = p<0.05, *** = p<0.001).

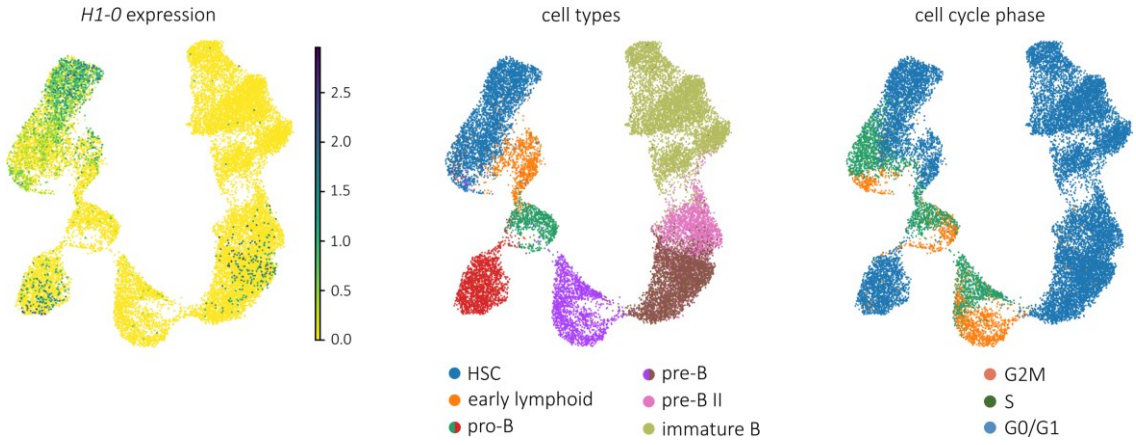


Figure 4.31: *H1-0* expression is primarily restricted to G0/G1 phase cells during normal early B-lymphoid development. *H1-0* expression levels across normal B-lymphoid differentiation depicted in a scRNA-seq UMAP visualization of B cell precursor cells from bone marrow of eight healthy donors. Data analysis was performed by Dr. Juha Mehtonen. Cell stage and cell cycle annotation has been published previously [195].

4.6 Targeting BCP-ALL via Quisinostat-induced H1-0 upregulation

Due to their cytostatic activity, HDACis are potent inducers of *H1-0* expression [194, 224]. Therefore, I assessed whether addition of HDACis can potentiate cell death induced by chemotherapeutic drugs commonly used in B-ALL treatment protocols.

To test correlation of H1-0 protein levels and drug sensitivity towards HDACis, I took advantage of the Functional Omics Resource of Acute Lymphoblastic Leukemia (FORALL) platform [227, 228] (https://proteomics.se/forall/). Indeed, H1-0 protein levels anti-correlated with sensitivity towards HDACis in a panel of 25 B-ALL cell lines (**Figure 4.32**). In particular, highly significant anti-correlation of H1-0 protein levels and sensitivity towards AR-42 and Vorinostat ($p<0.001$) was found, as well as towards 11 other HDACis, including Abexinostat and Quisinostat ($p<0.01$).

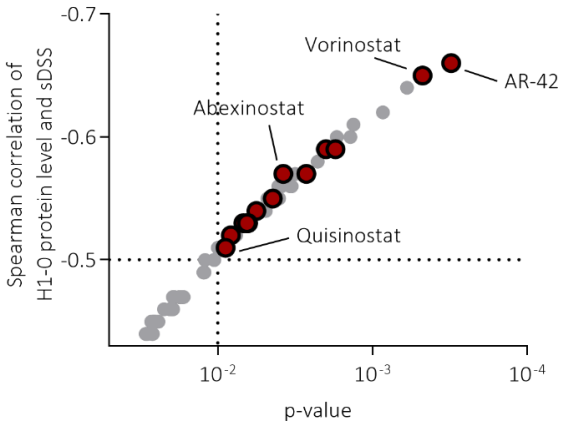


Figure 4.32: Susceptibility towards the HDACis correlates with H1-0 protein expression. Spearman correlation and Spearman p-values of H1-0 protein levels and selective drug sensitivity scores (sDSS) in 25 BCP-ALL cell lines derived from the FORALL platform [227, 228] (https://proteomics.se/forall/; cut-offs: $p<0.05$, FDR<0.25). HDACis are marked in red.

Next, to determine whether addition of Quisinostat to existing chemotherapeutic B-ALL regiments might improve treatment, synergy drug screenings were performed. Combination of Quisinostat with Vincristine and Daunorubicin, two chemotherapeutic drugs that are commonly used in B-ALL treatment protocols [229], exhibited strong synergism in REH cells (**Figure 4.33A-C**), whereas either no additive effect or a less pronounced additive effect was observed in MHH-CALL-2, SUP-B15 and RS4;11 cells. Moreover, strong synergism of Quisinostat with the proteasome inhibitor Bortezomib in REH with ZIP synergy scores >30 (**Figure 4.33D**), MHH-CALL-2 and SUP-B15 cell lines was detected, while synergism was less pronounced in RS4;11 cells (**Figure 4.33A**).

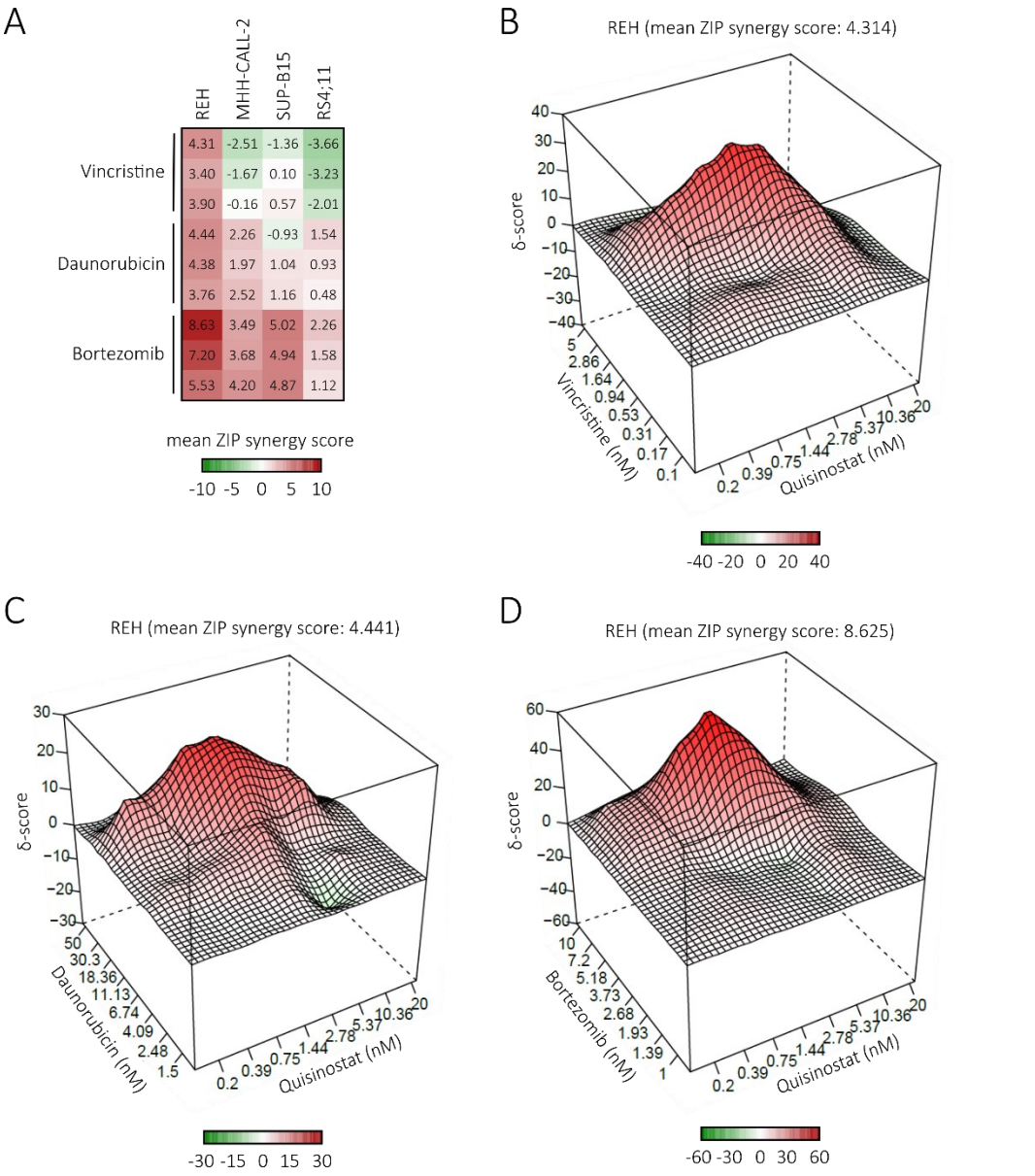


Figure 4.33: Quisinostat synergizes with B-ALL frontline drugs and the proteasome inhibitor Bortezomib in *ETV6::RUNX1*⁺ BCP-ALL. (A) Heatmap indicating mean ZIP synergy scores of Vincristine (0.1-5 nM), Daunorubicin (1.5-50 nM) or Bortezomib (1-10 nM) in combination with Quisinostat (0.2-20 nM) in four BCP-ALL cell lines treated for 72 hours. Data of three independent experiments is shown. Colors indicate synergy (red), lack of synergy (white) or antagonism (green). **(B-D)** Representative synergy plots visualizing drug combinations with high mean ZIP synergy scores in REH cells. Drug screening experiments and analyses were supported by Julian Schliehe-Diecks and Rebecca Hasselmann.

5. DISCUSSION

5.1 Generation of preleukemic *ETV6::RUNX1*$^{+}$ hiPSCs

The CRISPR/Cas9 technology has greatly facilitated the targeted introduction of point mutations into a variety of cell types. However, the efficacy of homology-directed repair (HDR) still remains lower, in particular in hard-to-transfect cell lines, such as hiPSCs. Low single-cell survival rates, time-consuming clonal expansion and maintenance to preserve pluripotency further complicate genetic editing of hiPSCs.

In this work, I constructed an HDR template of 3791 bp that encodes for *RUNX1* exons 2-8 as well as for a puromycin resistance gene under control of the human *EF1α* promoter. Puromycin resistance allows for pre-selecting hiPSCs that integrated the complete or partial *RUNX1* HDR template into their DNA. Indeed, partial integration of the HDR template was observed in multiple hiPSC clones where genotyping via PCR led to amplification of a product covering the 3' region of the *RUNX1* HDR template that contains the puromycin resistance gene, but gave no signal for the 5' region.

Colony screening via genotyping PCR and Sanger sequencing was enabled by using relatively short homology arm (HA) sequences of ≤500 bp, while previous studies using larger constructs relied on labor-intensive Southern blotting [125]. Altogether, screening ≈200 colonies of two hiPSC lines (HW8 and ChiPSC12) identified 3 clones with successful integration of the *RUNX1* HDR template.

Importantly, since the *RUNX1* HDR template was inserted at one *ETV6* allele, the other *ETV6* allele in the 3 genetically modified hiPSC clones stayed intact. This is in line with haploinsufficiency of *ETV6* but does not replicate the disruption of one *RUNX1* allele observed in patients. The ratio of ETV6::RUNX1 and RUNX1 is crucial because both transcription factors compete for the same DNA binding motifs through their RHD [126]. To improve our preleukemic hiPSC model, it would therefore be interesting to delete one of the remaining *RUNX1* alleles.

By inserting the *RUNX1* HDR template downstream of *ETV6* exon 5, expression of *ETV6::RUNX1* is driven by the endogenous *ETV6* promoter. This approach has been previously demonstrated to produce physiological *ETV6::RUNX1* expression levels in human and mouse models [125, 140], and I observed an increase of *ETV6::RUNX1* RNA expression during hiPSC differentiation towards HPCs, reaching levels comparable to those in the *ETV6::RUNX1*$^{+}$ BCP-ALL cell line REH.

ETV6::RUNX1 expression did not impact hiPSC colony morphology or pluripotency marker levels (SSEA-4, *DNMT3B*, *GDF3*, *POU5F1* and *NANOG*), and the genetically modified hiPSC clones maintained their normal karyotype. This might be due to low *ETV6::RUNX1* levels expressed at the hiPSC state. In line with this, a previous study reported that *ETV6* haploinsufficiency did not impact pluripotency marker expression in hiPSCs [98].

5.2 Transcriptome analysis of *ETV6::RUNX1*$^+$ preleukemia

Accurately modeling *ETV6::RUNX1*$^+$ preleukemia and overt leukemia in mice has largely failed to reproduce restriction to B lineage leukemia seen in humans [140, 141, 218]. This has been attributed to level-dependent effects of ETV6::RUNX1, especially in models overexpressing the fusion gene using viral transduction [217]. Discrepancies between *ETV6::RUNX1*$^+$ mouse and human models were also linked to poor inter-species conservation of GGAA repeat enhancers identified as key regulators of the *ETV6::RUNX1*$^+$ BCP-ALL gene signature [230]. Hence, accurately recapitulating the intricate effects of ETV6::RUNX1 may necessitate modeling its function in a human background with physiological expression levels.

Using whole-transcriptome sequencing, mild expression changes in *ETV6::RUNX1*$^+$ hiPSCs with few differentially expressed genes overlapping between the 3 hiPSC clones were detected. These subtle expression changes are in line with low expression of the fusion gene in hiPSCs as well as the weak oncogenic potential of ETV6::RUNX1. Indeed, expression of ETV6::RUNX1 alone is not sufficient to induce leukemia but requires additional mutations. While I did not observe major expression changes, hiPSC lines clustered according to genotype using principal component analysis of RNA-seq data. Moreover, common ETV6::RUNX1 target genes, including *RAG1* [133], *SPIB* [134] and *ICAM1* [221], were dysregulated. This indicates that *ETV6::RUNX1* expression was sufficient to induce a distinct transcriptional landscape even on hiPSC level.

ETV6::RUNX1$^+$ preleukemia in human cells has been previously studied by Böiers *et al*. in a developmental context using hiPSCs [125]. This study observed a partial block of B cell development at the IL7R$^+$ progenitor to pro-B cell transition. However, given the variability of *in vitro* B lineage differentiation protocols (e.g. due to differentiation potential of hiPSC lines or quality of feeder cell layers), high cost of reagents and low yield of hematopoietic precursor cells, downstream analysis of B lineage cells is challenging.

Here, I focused on the HSC compartment during early hematopoietic *in vitro* differentiation of hiPSCs. For reproducibility and to eliminate variability introduced by feeder cells, a commercially available feeder-free differentiation kit (STEMdiff Hematopoietic kit, Stemcell Technologies) was

used. This differentiation system has been shown to produce multilineage HPCs [231]. Characterization of HPCs by flow cytometry and scRNA-seq showed significantly increased numbers of HSCs derived from *ETV6::RUNX1*$^{+}$ hiPSCs compared to wild-type hiPSCs. Correlation of *ETV6::RUNX1* expression and HSC accumulation has been identified previously in mouse models [140, 232, 233]. Concomitant with HSC accumulation, I observed that transcriptional diversity (i.e. number of expressed genes per cell) was decreased in *ETV6::RUNX1*$^{+}$ HPCs compared to wild-type HPCs. Indeed, HSCs generally have reduced RNA contents and low rate of protein synthesis [234], and transcriptional diversity has previously been reported to be associated with cellular differentiation state [235].

Through differentiation, hiPSCs provide the opportunity to generate model systems in disease-relevant cell types. However, hiPSC modeling used in this work does not recapitulate the effects of cell-cell interactions observed in the bone marrow microenvironment. In future experiments, this might be addressed by using bone marrow organoid differentiation protocols that have recently been published [236, 237].

Engraftment of hiPSC-derived HSCs is highly inefficient and engraftment of either wild-type or *ETV6::RUNX1*$^{+}$ HSCs intravenously injected into mice failed (not shown). This is in line with previous studies that were unable to generate engraftable HSCs differentiated from hiPSCs without teratoma formation [238] or ectopic expression of transcription factors [239].

5.3 H1-0 upregulation in *ETV6::RUNX1*$^{+}$ preleukemia and BCP-ALL

Transcriptional analysis detected upregulation of linker histone *H1-0* across multiple data sets, both in preleukemic cells and BCP-ALL expressing the *ETV6::RUNX1* fusion gene. As a regulator of chromatin compaction, H1-0 is involved in epigenetic regulation and affects cellular differentiation [183]. Of note, *H1-0* upregulation in preleukemic cells was preserved during differentiation. Moreover, increased *H1-0* levels were also observed in leukemic BCP-ALL patient samples, suggesting retention of chromatin compaction throughout *ETV6::RUNX1*$^{+}$ BCP-ALL development. This is consistent with a recent study that showed lowest number of accessible chromatin sites, i.e. more compact chromatin, in *ETV6::RUNX1*$^{+}$ BCP-ALL compared to other ALL subtypes [165]. Similar loss of chromatin accessibility and cell cycle arrest has been detected in myeloid progenitors harboring the *RUNX1::ETO* translocation that retains the DNA-binding RHD, allowing it to bind to RUNX1 target sites [240].

Furthermore, detection of H1-0 protein levels in B-ALL cell lines indicated higher expression in slow-cycling cells, such as *ETV6::RUNX1*$^{+}$ REH and high hyperdiploid MHH-CALL-2 cells. This

supports previous findings of H1-0 accumulation in slowly or non-dividing cells [184]. Additionally, a study by Ebinger *et al*. characterizing dormant ALL cells identified H1-0 as the most significantly upregulated gene in leukemia stem cell (LSC)-like cells [241]. As expected, this gene signature of LSC-like cells was more highly expressed in *ETV6::RUNX1*$^{+}$ compared to wild-type cells [126].

Future analysis regarding the contribution of H1-0 to hematopoietic differentiation could reveal whether accumulation of *ETV6::RUNX1*$^{+}$ HSCs can be reverted by knockdown of H1-0. Given that epigenetic regulation plays a crucial role during differentiation of HSCs, it would be interesting to determine whether chromatin compaction via H1-0 contributes to B lymphoid lineage restriction seen in *ETV6::RUNX1*$^{+}$ leukemias.

5.4 Regulation of *H1-0* expression

Dual-luciferase promoter assays showed activation of the *H1-0* promoter region upon ETV6::RUNX1 expression in 293T cells. RUNX1 expression on the other hand led to reduction of promoter activation when compared to cells transfected with an empty control vector. Activation of the *H1-0* promoter upon *ETV6::RUNX1* expression is likely an indirect mechanism, since ETV6::RUNX1 does not directly bind to the *H1-0* promoter region in murine [221, 242] or human cells. Indirect activation of the *H1-0* promoter is further supported by the assumption that ETV6::RUNX1 primarily acts as a transcriptional repressor [126-131].

Previous studies have shown that *H1-0* expression is regulated via epigenetic mechanisms, including DNA methylation and acetylation. DNA methylation of a region located within the *H1-0* gene, the so called *H1-0* CGI shore, has been implicated in transcriptional repression in solid cancers [183]. I could show here that DNA methylation of the *H1-0* CGI shore significantly correlates to *H1-0* expression in leukemia and that *H1-0* CGI shore methylation is lowest in *ETV6::RUNX1*$^{+}$ BCP-ALL compared to other ALL entities. Apart from DNA methylation, *H1-0* expression can be induced by HDACis, such as sodium butyrate, Trichostatin A, Vorinostat, Abexinostat and Quisinostat [193, 194, 224]. In dual-luciferase promoter assays, the HDACi Quisinostat activated the *H1-0* promoter region in a dose-dependent manner.

Altogether, H1-0 expression responds to various intracellular and extracellular signals. Given that HDACis have a strong cytostatic effect and H1-0 levels are higher in slow-cycling cells, it is conceivable that H1-0 accumulates due to continued protein synthesis [243]. This hypothesis is supported by the observation that H1-0 synthesis is independent of DNA replication [181, 182].

5.5 Functional characterization of linker histone *H1-0*

To determine how H1-0 contributes to *ETV6::RUNX1*$^{+}$ BCP-ALL, siRNA-mediated knockdowns of H1-0 were performed in the REH cell line. Upstream regulator analysis of RNA-seq data showed a significant inhibition of the ETV6::RUNX1 signature upon H1-0 knockdown, indicating a substantial role of H1-0 in transcriptional control of ETV6::RUNX1 target genes. Genes dysregulated by H1-0 included common ETV6::RUNX1 target genes, such as *RAG1* and *EPOR* [14, 133, 135, 226]. On the other hand, activation of the p53 pathway upon *H1-0* depletion, including downregulation of the negative p53 regulator MDM2, was detected. Repression of the MDM2/p53 pathway via ETV6::RUNX1 has been reported before [225].

Overall, GSEA of RNA-seq data identified enriched gene signatures associated with DNA replication, histone modification, DNA repair and protein ubiquitination in REH cells treated with non-targeting compared to *H1-0*-targeting siRNA. Enrichment of epigenetic pathways, such as histone acetylation, is consistent with correlation of H1-0 expression and chromatin acetylation described previously [194, 224]. Moreover, GSEA detected enrichment of genes involved in chromatin condensation (gene set name: REACTOME_CONDENSATION_OF_ PROPHASE_CHROMOSOMES) which is in line with the structural role of H1-0 as a chromatin condenser [178] and indicates loosening of chromatin upon *H1-0* knockdown.

Protein-protein interaction studies have indicated involvement of H1-0 in RNA splicing and ribosomal RNA (rRNA) biogenesis [244]. Using GSEA, I found enrichment of multiple gene sets associated with rRNA expression. Moreover, evaluation of protein interactions from the STRING database (https://string-db.org/ [245]) indicated association of H1-0 with various histone proteins, cyclin-dependent kinases and E3 ubiquitin-protein ligase RNF168 [246] (**Figure 5.1**). In line with this observation, protein ubiquitination gene sets were enriched in REH cells treated with non-targeting siRNA (gene set names: REACTOME_E3_UBIQUITIN_LIGASES_ UBIQUITINATE_TARGET_PROTEINS and REACTOME_PROTEIN_UBIQUITINATION).

Interestingly, relaxation of chromatin has been suggested to facilitate DNA damage response (DDR) and triple knockout of H1-2, H1-3 and H1-4 improved DDR in murine embryonic stem cells [247]. Given that GSEA identified multiple enriched gene sets associated with DNA repair, it would be interesting to ascertain whether *H1-0*-depleted cells are more resistant to DNA damage, e.g. in response to ionizing radiation. Regarding *ETV6::RUNX1*$^{+}$ preleukemia, impaired DDR due to H1-0 upregulation might make preleukemic cells more prone to acquisition of secondary mutations.

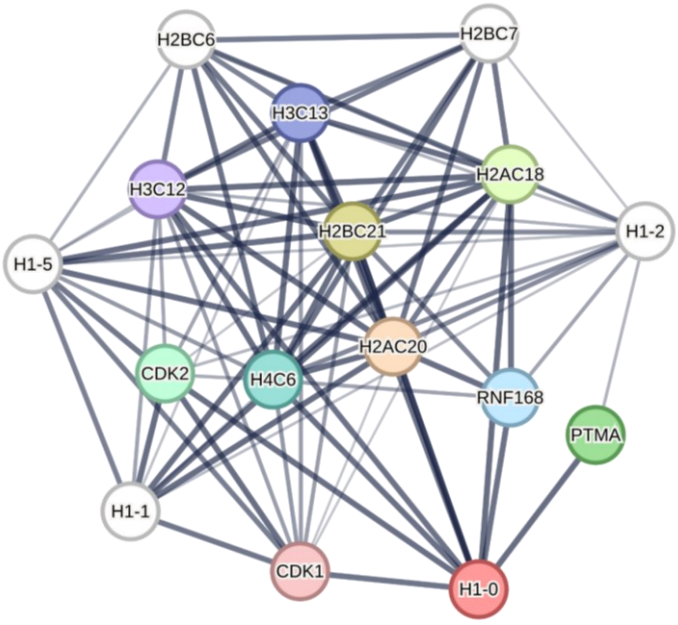


Figure 5.1: STRING analysis of the H1-0 interactome. Shown are interacting proteins determined by textmining and experimental evidence. The confidence level was set to 0.4 (medium) and thickness of network edges indicates the confidence of the respective interactions.

Leukemia cell lines are notoriously difficult to transfect. While nucleofection was able to achieve transient *H1-0* knockdown efficiencies of ≈50% on RNA level in REH cells, this approach was also accompanied by considerable cell death. Low total numbers of viable cells and the transient nature of siRNA treatment hampered functional downstream assays. Therefore, to enable stable downregulation of H1-0, viral transduction using an inducible knockdown system could be beneficial. This approach would allow for sorting of transduced cells and studying the long-term effects of H1-0 ablation.

Binding of H1-0 to chromatin has been found to depend on GC content [183]. Analysis of chromatin accessibility, e.g. via Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq), would allow to determine genomic loci that are affected by H1-0 knockdown or overexpression. These experiments are currently ongoing.

H1-0 accumulation has been associated with terminally differentiated cell types [185, 186]. The data presented here show that H1-0 expression rather indicates cellular quiescence and is highly expressed in human bone marrow and HSCs. High H1-0 levels observed in HSCs are in line with the largely quiescent nature of these cells [187]. Single-cell analysis indicated that the number of *H1-0*^{high} cells gradually decreases during B lymphoid development. This supports the hypothesis that H1-0 accumulates in quiescent cells that have high proliferative capacity [183, 241] during hematopoiesis. Increased quiescence of *ETV6::RUNX1*$^{+}$ preleukemic cells is in

keeping with the presence of these cells in the cord blood of approximately 5% of healthy newborns that was detected previously [45] and might offer an explanation for prolonged latency periods of *ETV6::RUNX1*$^{+}$ leukemia of up to 14 years [39, 84].

5.6 Targeting BCP-ALL via Quisinostat-induced H1-0 upregulation

While treatment outcomes of BCP-ALL patients have considerably improved over the last decades, therapy is still accompanied by severe side effects. Therefore, enhancing the sensitivity of chemotherapeutic regimens, for instance by combination with small molecules, such as HDACis, has been subject of intensive research. To date, there are multiple clinical trials testing the efficacy of HDACis in ALL, in particular for T-ALL and relapsed or refractory ALL [248] and four pan-HDACis (Vorinostat, Romidepsin, Belinostat and Panobinostat) have been approved by the Food and Drug Administration (FDA) for the treatment of T cell lymphoma or multiple myeloma [249]. Given the common involvement of epigenetic dysregulation in B-ALL, targeting epigenetic modifications holds substantial potential to enhance patient outcomes, for instance by reversing the epigenetic silencing of blast cells. However, further clinical studies are required to determine whether HDACis can be successfully integrated into treatment of other hematologic malignancies.

The HDACi Quisinostat has been identified as a potent H1-0 inducer in solid cancers [194] and I could show similar induction of H1-0 in BCP-ALL cell lines. In previous studies, Quisinostat has demonstrated high potency and bioavailability at low nanomolar concentrations [250, 251], while preserving normal stem cell function [194, 252, 253]. However, the predominantly cytostatic activity of HDACis *in vivo* suggests that single-drug treatment is not sufficient to induce cancer remission.

Using leukemic cell line models, I could show here that combinatorial treatment with the pan-HDACi Quisinostat is a promising approach to enhance treatment of *ETV6::RUNX1*$^{+}$ BCP-ALL when administered alongside Vincristine or Daunorubicin. Additionally, combination of the proteasome inhibitor Bortezomib with Quisinostat demonstrated remarkable synergy in targeting *ETV6::RUNX1*$^{+}$ leukemic cells at nanomolar concentrations. Indeed, combination of Bortezomib with Quisinostat showed promising treatment outcomes in a multiple myeloma mouse model [254], while a previous study [255] also reported efficacy of other pan-HDACis used in combination with Bortezomib in preclinical B-ALL models. As shown for solid cancer [194], H1-0 levels might indicate susceptibility towards HDACis in BCP-ALL. Therefore, combination treatments using Quisinostat might be a promising treatment approach for BCP-

ALL, particularly for subtypes with high basal levels of the H1-0. To verify the efficacy of combinatorial treatment using Quisinostat, future studies should include patient samples and *in vivo* treatment in mice.

6. OUTLOOK

Usually, leukemia develops by sequential acquisition of mutations. The subtle predisposing effect of the *ETV6::RUNX1* translocation seen by us and others make it challenging to determine the exact mechanism of how the fusion gene contributes to BCP-ALL development. Rather than inducing a fitness advantage, ETV6::RUNX1 slows proliferation and leads to a partial differentiation block [125, 126]. It is conceivable that ETV6::RUNX1 induces selective pressure for acquisition of certain mutations to overcome the observed fitness disadvantage of preleukemic cells. Indeed, recurrent second hits observed in *ETV6::RUNX1*$^+$ BCP-ALL impact pathways that are disrupted by ETV6::RUNX1, such as B lymphocyte differentiation (e.g. *ETV6*, *PAX5*, *IKZF1*, *EBF1*, *RAG1*, *RAG2*) or cell cycle (e.g. *CDKN2A*, *CDKN2B*). Using the preleukemic hiPSC models generated in this work, it would be interesting to introduce common secondary mutations and characterize their transformative potential on *ETV6::RUNX1*$^+$ preleukemic cells, especially regarding hematopoiesis or proliferation.

The question why certain CNAs are recurrently observed in *ETV6::RUNX1*$^+$ BCP-ALL patients remains incompletely understood. Using CRISPR screens, gene knockouts that result in selective advantage of *ETV6::RUNX1*$^+$ cells could be identified on a genome-wide scale (**Figure 6.1**). This would not only allow to confirm accumulation of lesions commonly seen in *ETV6::RUNX1*$^+$ BCP-ALL patients *in vitro*, but to mimic the sequential acquisition of mutations described by the two-hit hypothesis [6, 37].

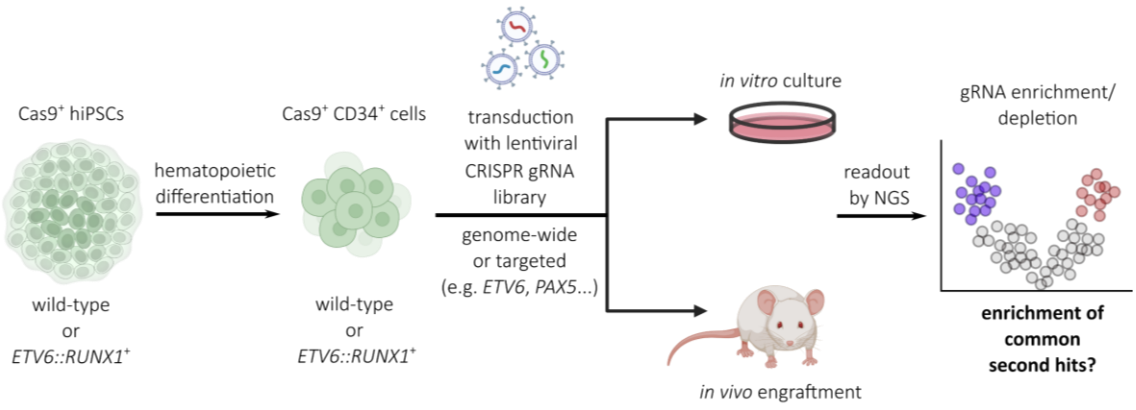


Figure 6.1: Schematic setup to characterize second hits in *ETV6::RUNX1*$^+$ BCP-ALL. Genetic perturbations that affect *in vitro* or *in vivo* proliferation and survival of hiPSC-derived CD34$^+$ cells with or without *ETV6::RUNX1* translocation can be identified by counting of gRNAs sequences delivered via lentivirus. NGS: next-generation sequencing. Figure created with BioRender.

Furthermore, transformed *ETV6::RUNX1*$^+$ HSCs should gain the ability to engraft in mice, which would enable *in vivo* expansion and in-depth characterization of the arising leukemic cells. Combined with CRISPR screens, scRNA-seq offers the possibility to determine the specific expression signatures of gene-edited cells that gained advantageous secondary mutations and led to transformation of preleukemia to leukemia. The feasibility of sequential introduction of mutations into hiPSCs and analysis of leukemia evolution *in vitro* and *in vivo* has been demonstrated for AML in a recent study [256].

The dogma of sequential acquisition of initiating and co-operating secondary mutations during leukemogenesis has been challenged recently for acute megakaryocytic leukemia (AMKL), where the combination of *GATA1* mutations and amplification of chromosome 21 led to disease development, irrespective of the order of lesions [257]. Whether acquisition of *ETV6::RUNX1* as a secondary lesion is able to induce leukemia is unknown. Interestingly, germline mutations predisposing to ALL affect genes that are also common secondary events (such as *ETV6* [52-54], *PAX5* [55, 56] and *IKZF1* [57, 58]), and *ETV6::RUNX1* has been detected in cells with underlying *ETV6* germline mutation [258]. This indicates that it is the combination of mutations affecting genes restricted to a few specific pathways that ultimately leads to leukemia development. Lesions restricted to these pathways may alter gene dosages, resulting in similar expression profiles seen in *ETV6::RUNX1*$^+$ BCP-ALL cases [22]. Again, this hypothesis could be tested via sequential introduction of mutations into hiPSCs and subsequent RNA sequencing.

Of note, the differentiation state of the respective cell of origin may alter the transforming potential of secondary hits. *ETV6::RUNX1*$^+$ BCP-ALL is predominantly a childhood disease and may differ from adult ALL by its origin in cells of different developmental state. Given that *ETV6::RUNX1*$^+$ BCP-ALL is predominantly a disease of childhood, it is conceivable that the *ETV6::RUNX1* fusion gene arises during fetal development in a restricted progenitor cell population or microenvironment. Therefore, different cell types, e.g. HSCs, IL7R$^+$ or pro-B cells, would need to be manipulated in order to study the transforming effect of secondary lesions. To date, differentiation of hiPSC-derived HSCs towards definitive B cells remains difficult, as current protocols are time-consuming and result in a low cell yields [259]. These challenges in hiPSC differentiation limit the ability to produce adequate numbers of B cell progenitors for comprehensive functional studies.

Despite these challenges, the preleukemic hiPSC model developed in this thesis opens up many possibilities to study leukemogenesis via introduction of secondary mutations, e.g. via CRISPR screens, and thereby provides a platform for dissecting the molecular mechanisms underlying *ETV6::RUNX1*$^+$ BCP-ALL.

7. REFERENCES

1. Goldstick, J.E., R.M. Cunningham, and P.M. Carter, *Current causes of death in children and adolescents in the United States.* The New England Journal of Medicine, 2022. **386**(20): p. 1955-1956.
2. Ferlay, J., et al. *Global cancer observatory: Cancer Today. Lyon, France: International agency for research on cancer*. 2020 [cited 2023 april 14]; Available from: https://gco.iarc.fr/today.
3. *ProgramSEER*Explorer: survival (all cancer sites combined)*. [cited 2023 april 17]; Available from: https://seer.cancer.gov/statistics-network/explorer/.
4. Erdmann, F., et al., *German childhood cancer registry - annual report 2019 (1980-2018)*. 2020, Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI) at the University Medical Center of the Johannes Gutenberg University Mainz.
5. *Cancer Facts & Figures 2023*. 2023, American Cancer Society: Atlanta.
6. Greaves, M., *A causal mechanism for childhood acute lymphoblastic leukaemia.* Nature Reviews Cancer, 2018. **18**(8): p. 471-484.
7. Dworzak, M.N., et al., *AIEOP-BFM consensus guidelines 2016 for flow cytometric immunophenotyping of pediatric acute lymphoblastic leukemia.* Cytometry Part B-Clinical Cytometry, 2018. **94**(1): p. 82-93.
8. Somasundaram, R., et al., *Transcription factor networks in B-cell differentiation link development to acute lymphoid leukemia.* Blood, 2015. **126**(2): p. 144-152.
9. Morgan, D. and V. Tergaonkar, *Unraveling B cell trajectories at single cell resolution.* Trends in Immunology, 2022. **43**(3): p. 210-229.
10. Schroeder, H.W., A. Radbruch, and C. Berek, *B-cell development and differentiation*, in *Clinical Immunology*. 2019, Elsevier: London. p. 107-118.
11. Max, E.E., J.G. Seidman, and P. Leder, *Sequences of five potential recombination sites encoded close to an immunoglobulin kappa constant region gene.* Proceedings of the National Academy of Sciences of the United States of America, 1979. **76**(7): p. 3450-3454.
12. Notarangelo, L.D., et al., *Human RAG mutations: biochemistry and clinical implications.* Nature Reviews Immunology, 2016. **16**(4): p. 234-246.
13. Zhang, M. and P.C. Swanson, *V(D)J recombinase binding and cleavage of cryptic recombination signal sequences identified from lymphoid malignancies.* Journal of Biological Chemistry, 2008. **283**(11): p. 6717-6727.
14. Papaemmanuil, E., et al., *RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia.* Nature Genetics, 2014. **46**(2): p. 116-125.
15. Murphy, K., C. Weaver, and C. Janeway, *Janeway's immunobiology*. Ninth ed. 2017, New York: Garland Science, Taylor & Francis Group, LLC.
16. Inaba, H. and C.-H. Pui, *Advances in the diagnosis and treatment of pediatric acute lymphoblastic leukemia.* Journal of clinical medicine, 2021. **10**(9): p. 1926.
17. Lampert, F., *Cellulärer DNS-Gehalt und Chromosomenzahl bei der akuten Leukämie im Kindesalter und ihre Bedeutung für Chemotherapie und Prognose.* Klinische Wochenschrift, 1967. **45**(15): p. 763-768.
18. Haas, O.A. and A. Borkhardt, *Hyperdiploidy: the longest known, most prevalent, and most enigmatic form of acute lymphoblastic leukemia in children.* Leukemia, 2022. **36**(12): p. 2769-2783.
19. Paulsson, K., et al., *The genomic landscape of high hyperdiploid childhood acute lymphoblastic leukemia.* Nature Genetics, 2015. **47**(6): p. 672-676.
20. Paulsson, K., et al., *Evidence for a single-step mechanism in the origin of hyperdiploid childhood acute lymphoblastic leukemia.* Genes Chromosomes Cancer, 2005. **44**(2): p. 113-122.

21. Inaba, H. and C.G. Mullighan, *Pediatric acute lymphoblastic leukemia.* Haematologica, 2020. **105**(11): p. 2524-2539.
22. Gu, Z., et al., *PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia.* Nature Genetics, 2019. **51**(2): p. 296-307.
23. Pui, C.-H. and S. Jeha, *New therapeutic strategies for the treatment of acute lymphoblastic leukaemia.* Nature Reviews Drug Discovery, 2007. **6**(2): p. 149-165.
24. Druker, B.J., et al., *Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr–Abl positive cells.* Nature Medicine, 1996. **2**(5): p. 561-566.
25. Kim, J.C., et al., *Transcriptomic classes of BCR-ABL1 lymphoblastic leukemia.* Nature Genetics, 2023.
26. van der Veer, A., et al., *IKZF1 status as a prognostic feature in BCR-ABL1–positive childhood ALL.* Blood, 2014. **123**(11): p. 1691-1698.
27. Andersson, A.K., et al., *The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias.* Nature Genetics, 2015. **47**(4): p. 330-337.
28. Den Boer, M.L., et al., *A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study.* The Lancet Oncology, 2009. **10**(2): p. 125-134.
29. Mullighan, C.G., et al., *Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia.* The New England Journal of Medicine, 2009. **360**(5): p. 470-480.
30. Lilljebjörn, H., et al., *Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia.* Nature Communications, 2016. **7**(1): p. 11790.
31. Yasuda, T., et al., *Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults.* Nature Genetics, 2016. **48**(5): p. 569-574.
32. Gu, Z., et al., *Genomic analyses identify recurrent MEF2D fusions in acute lymphoblastic leukaemia.* Nature Communications, 2016. **7**: p. 13331.
33. Hirabayashi, S., et al., *ZNF384-related fusion genes define a subgroup of childhood B-cell precursor acute lymphoblastic leukemia with a characteristic immunotype.* Haematologica, 2017. **102**(1): p. 118-129.
34. Hormann, F.M., et al., *NUTM1 is a recurrent fusion gene partner in B-cell precursor acute lymphoblastic leukemia associated with increased expression of genes on chromosome band 10p12.31-12.2.* Haematologica, 2019. **104**(10): p. 455-459.
35. Wagener, R., et al., *IG-MYC (+) neoplasms with precursor B-cell phenotype are molecularly distinct from Burkitt lymphomas.* Blood, 2018. **132**(21): p. 2280-2285.
36. Passet, M., et al., *PAX5 P80R mutation identifies a novel subtype of B-cell precursor acute lymphoblastic leukemia with favorable outcome.* Blood, 2019. **133**(3): p. 280-284.
37. Knudson, A.G., *Mutation and cancer: statistical study of retinoblastoma.* Proceedings of the National Academy of Sciences of the United States of America, 1971. **68**(4): p. 820-823.
38. Greaves, M.F. and J. Wiemels, *Origins of chromosome translocations in childhood leukaemia.* Nature Reviews Cancer, 2003. **3**(9): p. 639-649.
39. Wiemels, J.L., et al., *Protracted and variable latency of acute lymphoblastic leukemia after TEL-AML1 gene fusion in utero.* Blood, 1999. **94**(3): p. 1057-1062.
40. Gröbner, S.N., et al., *The landscape of genomic alterations across childhood cancers.* Nature, 2018. **555**(7696): p. 321-327.
41. Greaves, M.F., et al., *Leukemia in twins: lessons in natural history.* Blood, 2003. **102**(7): p. 2321-2333.
42. Clarkson, B.D. and E.A. Boyse, *Possible explanation of the high concordance for acute leukaemia in monozygotic twins.* Lancet, 1971. **1**(7701): p. 699-701.
43. Alpar, D., et al., *Clonal origins of ETV6-RUNX1+ acute lymphoblastic leukemia: studies in monozygotic twins.* Leukemia, 2015. **29**(4): p. 839-846.

44. Gale, K.B., et al., *Backtracking leukemia to birth: identification of clonotypic gene fusion sequences in neonatal blood spots.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(25): p. 13950-13954.
45. Schäfer, D., et al., *Five percent of healthy newborns have an ETV6-RUNX1 fusion as revealed by DNA-based GIPFEL screening.* Blood, 2018. **131**(7): p. 821-826.
46. Ford, A.M., et al., *In utero rearrangements in the trithorax-related oncogene in infant leukaemias.* Nature, 1993. **363**(6427): p. 358-360.
47. Super, H.J.G., et al., *Clonal, nonconstitutional rearrangements of the MLL gene in infant twins with acute lymphoblastic leukemia: in utero chromosome rearrangement of 11q23.* Blood, 1994. **83**(3): p. 641-644.
48. Ford, A.M., et al., *Fetal origins of the TEL-AML1 fusion gene in identical twins with leukemia.* Proceedings of the National Academy of Sciences, 1998. **95**(8): p. 4584-4588.
49. Hein, D., et al., *The preleukemic TCF3-PBX1 gene fusion can be generated in utero and is present in ≈0.6% of healthy newborns.* Blood, 2019. **134**(16): p. 1355-1358.
50. Cazzaniga, G., et al., *Developmental origins and impact of BCR-ABL1 fusion and IKZF1 deletions in monozygotic twins with Ph+ acute lymphoblastic leukemia.* Blood, 2011. **118**(20): p. 5559-5564.
51. Buckley, J.D., et al., *Concordance for childhood cancer in twins.* Medical and Pediatric Oncology, 1996. **26**(4): p. 223-229.
52. Moriyama, T., et al., *Germline genetic variation in ETV6 and risk of childhood acute lymphoblastic leukaemia: a systematic genetic study.* The Lancet Oncology, 2015. **16**(16): p. 1659-1666.
53. Noetzli, L., et al., *Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia.* Nature Genetics, 2015. **47**(5): p. 535-538.
54. Zhang, M.Y., et al., *Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy.* Nature Genetics, 2015. **47**(2): p. 180-185.
55. Shah, S., et al., *A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia.* Nature Genetics, 2013. **45**(10): p. 1226-1231.
56. Auer, F., et al., *Inherited susceptibility to pre B-ALL caused by germline transmission of PAX5 c.547G>A.* Leukemia, 2014. **28**(5): p. 1136-1138.
57. Kuehn, H.S., et al., *Loss of B cells in patients with heterozygous mutations in IKAROS.* The New England Journal of Medicine, 2016. **374**(11): p. 1032-1043.
58. Churchman, M.L., et al., *Germline genetic IKZF1 variation and predisposition to childhood acute lymphoblastic leukemia.* Cancer Cell, 2018. **33**(5): p. 937-948.
59. Qian, M., et al., *TP53 germline variations influence the predisposition and prognosis of B-cell acute lymphoblastic leukemia in children.* Journal of Clinical Oncology, 2018. **36**(6): p. 591-599.
60. Cave, H., et al., *Acute lymphoblastic leukemia in the context of RASopathies.* European Journal of Medical Genetics, 2016. **59**(3): p. 173-178.
61. Brown, A.L., et al., *Inherited genetic susceptibility to acute lymphoblastic leukemia in Down syndrome.* Blood, 2019. **134**(15): p. 1227-1237.
62. Vijayakrishnan, J., et al., *Genome-wide association study identifies susceptibility loci for B-cell childhood acute lymphoblastic leukemia.* Nature Communications, 2018. **9**(1): p. 1340.
63. Ellinghaus, E., et al., *Identification of germline susceptibility loci in ETV6-RUNX1-rearranged childhood acute lymphoblastic leukemia.* Leukemia, 2012. **26**(5): p. 902-909.
64. Steeghs, E.M.P., et al., *Copy number alterations in B-cell development genes, drug resistance, and clinical outcome in pediatric B-cell precursor acute lymphoblastic leukemia.* Scientific Reports, 2019. **9**(1): p. 4634.
65. Mullighan, C.G., et al., *BCR–ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros.* Nature, 2008. **453**(7191): p. 110-114.

66. Doll, R. and R. Wakeford, *Risk of childhood cancer from fetal irradiation.* The British Journal of Radiology, 1997. **70**(830): p. 130-139.
67. Urayama, K.Y., et al., *A meta-analysis of the association between day-care attendance and childhood acute lymphoblastic leukaemia.* International Journal of Epidemiology, 2010. **39**(3): p. 718-732.
68. Marcotte, E.L., et al., *Caesarean delivery and risk of childhood leukaemia: a pooled analysis from the Childhood Leukemia International Consortium (CLIC).* The Lancet Haematology, 2016. **3**(4): p. 176-185.
69. Brown, P., et al., *Pediatric acute lymphoblastic leukemia, version 2.2020, NCCN clinical practice guidelines in oncology.* Journal of the National Comprehensive Cancer Network, 2020. **18**(1): p. 81-112.
70. Cortelazzo, S., et al., *Lymphoblastic lymphoma.* Critical Reviews in Oncology/Hematology, 2017. **113**: p. 304-317.
71. Farber, S. and L.K. Diamond, *Temporary remissions in acute leukemia in children produced by folic acid antagonist, 4-aminopteroyl-glutamic acid.* The New England Journal of Medicine, 1948. **238**(23): p. 787-793.
72. Hunger, S.P. and C.G. Mullighan, *Acute lymphoblastic leukemia in children.* New England Journal of Medicine, 2015. **373**(16): p. 1541-1552.
73. Pui, C.-H., L.L. Robison, and A.T. Look, *Acute lymphoblastic leukaemia.* The Lancet, 2008. **371**(9617): p. 1030-1043.
74. Trausti, O., et al., *Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: prognostic factors, treatment and outcome.* Haematologica, 2016. **101**(1): p. 68-76.
75. Gaudichon, J., et al., *Mechanisms of extramedullary relapse in acute lymphoblastic leukemia: Reconciling biological concepts and clinical issues.* Blood Reviews, 2019. **36**: p. 40-56.
76. Pieters, R. and W.L. Carroll, *Biology and treatment of acute lymphoblastic leukemia.* Pediatric Clinics of North America, 2008. **55**(1): p. 1-20.
77. van der Velden, V.H.J., et al., *New cellular markers at diagnosis are associated with isolated central nervous system relapse in paediatric B-cell precursor acute lymphoblastic leukaemia.* British Journal of Haematology, 2016. **172**(5): p. 769-781.
78. Oskarsson, T., et al., *Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: prognostic factors, treatment and outcome.* Haematologica, 2016. **101**(1): p. 68-76.
79. Terwilliger, T. and M. Abdul-Hay, *Acute lymphoblastic leukemia: a comprehensive review and 2017 update.* Blood Cancer Journal, 2017. **7**(6): p. 577.
80. Buchmann, S., et al., *Remission, treatment failure, and relapse in pediatric ALL: an international consensus of the Ponte-di-Legno Consortium.* Blood, 2022. **139**(12): p. 1785-1793.
81. Silverman, L.B., *Balancing cure and long-term risks in acute lymphoblastic leukemia.* Hematology, 2014. **2014**(1): p. 190-197.
82. Inaba, H., M. Greaves, and C.G. Mullighan, *Acute lymphoblastic leukaemia.* The Lancet, 2013. **381**(9881): p. 1943-1955.
83. Pui, C.-H., M.V. Relling, and J.R. Downing, *Acute lymphoblastic leukemia.* New England Journal of Medicine, 2004. **350**(15): p. 1535-1548.
84. Mori, H., et al., *Chromosome translocations and covert leukemic clones are generated during normal fetal development.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(12): p. 8242-8247.
85. Gandemer, V., et al., *Excellent prognosis of late relapses of ETV6/RUNX1-positive childhood acute lymphoblastic leukemia: lessons from the FRALLE 93 protocol.* Haematologica, 2012. **97**(11): p. 1743-1750.

86. Forestier, E., et al., *Outcome of ETV6/RUNX1-positive childhood acute lymphoblastic leukaemia in the NOPHO-ALL-1992 protocol: frequent late relapses but good overall survival.* British Journal of Haematology, 2008. **140**(6): p. 665-672.
87. Zelent, A., M. Greaves, and T. Enver, *Role of the TEL-AML1 fusion gene in the molecular pathogenesis of childhood acute lymphoblastic leukaemia.* Oncogene, 2004. **23**: p. 4275-4283.
88. Hayashi, Y., Y. Harada, and H. Harada, *Myeloid neoplasms and clonal hematopoiesis from the RUNX1 perspective.* Leukemia, 2022. **36**(5): p. 1203-1214.
89. Wang, L.C., et al., *Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL.* The EMBO Journal, 1997. **16**(14): p. 4374-4383.
90. Chakrabarti, S.R. and G. Nucifora, *The leukemia-associated gene TEL encodes a transcription repressor which associates with SMRT and mSin3A.* Biochemical and Biophysical Research Communications, 1999. **264**(3): p. 871-877.
91. Lopez, R.G., et al., *TEL is a sequence-specific transcriptional repressor.* Journal of Biological Chemistry, 1999. **274**(42): p. 30132-30138.
92. Boccuni, P., et al., *The human L(3)MBT polycomb group protein is a transcriptional repressor and interacts physically and functionally with TEL (ETV6).* Journal of Biological Chemistry, 2003. **278**(17): p. 15412-15420.
93. Nakamura, Y., et al., *TEL/ETV6 binds to corepressor KAP1 via the HLH domain.* International Journal of Hematology, 2006. **84**(4): p. 377-380.
94. Wang, L. and S.W. Hiebert, *TEL contacts multiple co-repressors and specifically associates with histone deacetylase-3.* Oncogene, 2001. **20**(28): p. 3716-3725.
95. Nordentoft, I. and P. Jørgensen, *The acetyltransferase 60 kDa trans-acting regulatory protein of HIV type 1-interacting protein (Tip60) interacts with the translocation E26 transforming-specific leukaemia gene (TEL) and functions as a transcriptional co-repressor.* Biochemical Journal, 2003. **374**(Pt 1): p. 165-173.
96. Wang, L.C., et al., *The TEL/ETV6 gene is required specifically for hematopoiesis in the bone marrow.* Genes & Development, 1998. **12**(15): p. 2392-2402.
97. Hock, H., et al., *Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival.* Genes & Development, 2004. **18**(19): p. 2336-2341.
98. Borst, S., et al., *Study of inherited thrombocytopenia resulting from mutations in ETV6 or RUNX1 using a human pluripotent stem cell model.* Stem Cell Reports, 2021. **16**(6): p. 1458-1467.
99. De Braekeleer, E., et al., *ETV6 fusion genes in hematological malignancies: A review.* Leukemia Research, 2012. **36**(8): p. 945-961.
100. Sroczynska, P., et al., *The differential activities of Runx1 promoters define milestones during embryonic hematopoiesis.* Blood, 2009. **114**(26): p. 5279-5289.
101. Ghozi, M.C., et al., *Expression of the human acute myeloid leukemia gene AML1 is regulated by two promoter regions.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(5): p. 1935-1940.
102. Challen, G.A. and M.A. Goodell, *Runx1 isoforms show differential expression patterns during hematopoietic development but have similar functional effects in adult hematopoietic stem cells.* Experimental Hematology, 2010. **38**(5): p. 403-416.
103. Meyers, S., J.R. Downing, and S.W. Hiebert, *Identification of AML-1 and the (8;21) translocation protein (AML-1/ETO) as sequence-specific DNA-binding proteins: the runt homology domain is required for DNA binding and protein-protein interactions.* Molecular and Cellular Biology, 1993. **13**(10): p. 6336-6345.
104. Gu, T.L., et al., *Auto-inhibition and partner proteins, core-binding factor beta (CBFbeta) and Ets-1, modulate DNA binding by CBFalpha2 (AML1).* Molecular Cell Biology, 2000. **20**(1): p. 91-103.
105. Huang, G., et al., *Dimerization with PEBP2beta protects RUNX1/AML1 from ubiquitin-proteasome-mediated degradation.* The EMBO Journal, 2001. **20**(4): p. 723-733.

106. Wotton, D., et al., *Cooperative binding of Ets-1 and core binding factor to DNA.* Molecular and Cellular Biology, 1994. **14**(1): p. 840-850.
107. Petrovick, M.S., et al., *Multiple functional domains of AML1: PU.1 and C/EBPα synergize with different regions of AML1.* Molecular and Cellular Biology, 1998. **18**(7): p. 3915-3925.
108. Imai, Y., et al., *The corepressor mSin3A regulates phosphorylation-induced activation, intranuclear location, and stability of AML1.* Molecular Cell Biology, 2004. **24**(3): p. 1033-1043.
109. Kitabayashi, I., et al., *Interaction and functional cooperation of the leukemia-associated factors AML1 and p300 in myeloid cell differentiation.* The EMBO Journal, 1998. **17**(11): p. 2994-3004.
110. Wang, Q., et al., *Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(8): p. 3444-3449.
111. Okuda, T., et al., *AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis.* Cell, 1996. **84**(2): p. 321-330.
112. Cai, Z., et al., *Haploinsufficiency of AML1 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo.* Immunity, 2000. **13**(4): p. 423-431.
113. Ichikawa, M., et al., *AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis.* Nature Medicine, 2004. **10**(3): p. 299-304.
114. Niebuhr, B., et al., *Runx1 is essential at two stages of early murine B-cell development.* Blood, 2013. **122**(3): p. 413-423.
115. Lie, A.L.M., et al., *Regulation of RUNX1 dosage is crucial for efficient blood formation from hemogenic endothelium.* Development, 2018. **145**(5).
116. De Braekeleer, E., et al., *RUNX1 translocations and fusion genes in malignant hemopathies.* Future Oncology, 2011. **7**(1): p. 77-91.
117. Sood, R., Y. Kamikubo, and P. Liu, *Role of RUNX1 in hematological malignancies.* Blood, 2017. **129**(15): p. 2070-2082.
118. Song, W.J., et al., *Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia.* Nature Genetics, 1999. **23**(2): p. 166-175.
119. Simon, L., et al., *High frequency of germline RUNX1 mutations in patients with RUNX1-mutated AML.* Blood, 2020. **135**(21): p. 1882-1886.
120. Antony-Debré, I., et al., *Somatic mutations associated with leukemic progression of familial platelet disorder with predisposition to acute myeloid leukemia.* Leukemia, 2016. **30**(4): p. 999-1002.
121. Zuna, J., et al., *ETV6/RUNX1 (TEL/AML1) is a frequent prenatal first hit in childhood leukemia.* Blood, 2011. **117**(1): p. 370-371.
122. Lausten-Thomsen, U., et al., *Prevalence of t(12;21)[ETV6-RUNX1]–positive cells in healthy neonates.* Blood, 2011. **117**(1): p. 186-189.
123. Olsen, M., et al., *RT-PCR screening for ETV6-RUNX1-positive clones in cord blood from newborns in the Danish national birth cohort.* Journal of Pediatric Hematology/Oncology, 2012. **34**(4): p. 301-303.
124. Hong, D., et al., *Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia.* Science, 2008. **319**(5861): p. 336-339.
125. Böiers, C., et al., *A human iPS model implicates embryonic B-myeloid fate restriction as developmental susceptibility to B acute lymphoblastic leukemia-associated ETV6-RUNX1.* Developmental Cell, 2018. **44**(3): p. 362-377.
126. Wray, J.P., et al., *Regulome analysis in B-acute lymphoblastic leukemia exposes Core Binding Factor addiction as a therapeutic vulnerability.* Nature Communications, 2022. **13**(1): p. 7124.

127. Guidez, F., et al., *Recruitment of the nuclear receptor corepressor N-CoR by the TEL moiety of the childhood leukemia–associated TEL-AML1 oncoprotein.* Blood, 2000. **96**(7): p. 2557-2561.
128. Hiebert, S.W., et al., *The t (12; 21) translocation converts AML-1B from an activator to a repressor of transcription.* Molecular and Cellular Biology, 1996.
129. Morrow, M., et al., *TEL-AML1 preleukemic activity requires the DNA binding domain of AML1 and the dimerization and corepressor binding domains of TEL.* Oncogene, 2007. **26**(30): p. 4404-4414.
130. Fuka, G., et al., *The leukemia-specific fusion gene ETV6/RUNX1 perturbs distinct key biological functions primarily by gene repression.* PLoS One, 2011. **6**(10): p. 26348.
131. Teppo, S., et al., *Genome-wide repression of eRNA and target gene loci by the ETV6-RUNX1 fusion in acute leukemia.* Genome Research, 2016. **26**(11): p. 1468-1477.
132. Fuka, G., et al., *Silencing of ETV6/RUNX1 abrogates PI3K/AKT/mTOR signaling and impairs reconstitution of leukemia in xenografts.* Leukemia, 2012. **26**(5): p. 927-933.
133. Jakobczyk, H., et al., *ETV6-RUNX1 and RUNX1 directly regulate RAG1 expression: one more step in the understanding of childhood B-cell acute lymphoblastic leukemia leukemogenesis.* Leukemia, 2022. **36**(2): p. 549-554.
134. Xu, L.S., et al., *ETV6-RUNX1 interacts with a region in SPIB intron 1 to regulate gene expression in pre-B-cell acute lymphoblastic leukemia.* Experimental Hematology, 2019. **73**: p. 50-63.
135. Torrano, V., et al., *ETV6-RUNX1 promotes survival of early B lineage progenitor cells via a dysregulated erythropoietin receptor.* Blood, 2011. **118**(18): p. 4910-4918.
136. Inthal, A., et al., *Role of the erythropoietin receptor in ETV6/RUNX1-positive acute lymphoblastic leukemia.* Clinical Cancer Research, 2008. **14**(22): p. 7196-7204.
137. Dann, E.J., et al., *Lineage specificity of CBFA2 fusion transcripts.* Leukemia Research, 2000. **24**(1): p. 11-17.
138. Borkhardt, A., et al., *Incidence and clinical relevance of TEL/AML1 fusion genes in children with acute lymphoblastic leukemia enrolled in the German and Italian multicenter therapy trials.* Blood, 1997. **90**(2): p. 571-577.
139. Gerr, H., et al., *Acute leukaemias of ambiguous lineage in children: characterization, prognosis and therapy recommendations.* British Journal of Haematology, 2010. **149**(1): p. 84-92.
140. Schindler, J.W., et al., *TEL-AML1 corrupts hematopoietic stem cells to persist in the bone marrow and initiate leukemia.* Cell Stem Cell, 2009. **5**(1): p. 43-53.
141. van der Weyden, L., et al., *Modeling the evolution of ETV6-RUNX1-induced B-cell precursor acute lymphoblastic leukemia in mice.* Blood, 2011. **118**(4): p. 1041-1051.
142. Rodríguez-Hernández, G., et al., *Infection exposure promotes ETV6-RUNX1 precursor B-cell leukemia via impaired H3K4 demethylases.* Cancer Research, 2017. **77**(16): p. 4365-4377.
143. Lilljebjörn, H., et al., *The correlation pattern of acquired copy number changes in 164 ETV6/RUNX1-positive childhood acute lymphoblastic leukemias.* Human Molecular Genetics, 2010. **19**(16): p. 3150-3158.
144. Loncarevic, I.F., et al., *Trisomy 21 is a recurrent secondary aberration in childhood acute lymphoblastic leukemia with TEL/AML1 gene fusion.* Genes, Chromosomes and Cancer, 1999. **24**(3): p. 272-277.
145. Maia, A.T., et al., *Molecular tracking of leukemogenesis in a triplet pregnancy.* Blood, 2001. **98**(2): p. 478-482.
146. Ma, Y., et al., *Developmental timing of mutations revealed by whole-genome sequencing of twins with acute lymphoblastic leukemia.* Proceedings of the National Academy of Sciences, 2013. **110**(18): p. 7429-7433.
147. Fyodorov, D.V., et al., *Emerging roles of linker histones in regulating chromatin structure and function.* Nature Reviews Molecular Cell Biology, 2018. **19**(3): p. 192-206.

148. Huether, R., et al., *The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes.* Nature Communications, 2014. **5**(1): p. 3630.
149. Jaffe, J.D., et al., *Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia.* Nature Genetics, 2013. **45**(11): p. 1386-1391.
150. Oyer, J.A., et al., *Point mutation E1099K in MMSET/NSD2 enhances its methyltranferase activity and leads to altered global chromatin methylation in lymphoid malignancies.* Leukemia, 2014. **28**(1): p. 198-201.
151. Bernt, K.M. and S.A. Armstrong, *Targeting Epigenetic Programs in MLL-Rearranged Leukemias.* Hematology, 2011. **2011**(1): p. 354-360.
152. Mar, B.G., et al., *Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia.* Nature Communications, 2014. **5**: p. 3469.
153. Mullighan, C.G., et al., *CREBBP mutations in relapsed acute lymphoblastic leukaemia.* Nature, 2011. **471**(7337): p. 235-239.
154. Malinowska-Ozdowy, K., et al., *KRAS and CREBBP mutations: a relapse-linked malicious liaison in childhood high hyperdiploid acute lymphoblastic leukemia.* Leukemia, 2015. **29**(8): p. 1656-1667.
155. Larsen, F., et al., *CpG islands as gene markers in the human genome.* Genomics, 1992. **13**(4): p. 1095-1107.
156. Milani, L., et al., *DNA methylation for subtype classification and prediction of treatment outcome in patients with childhood acute lymphoblastic leukemia.* Blood, 2010. **115**(6): p. 1214-1225.
157. Figueroa, M.E., et al., *Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia.* The Journal of Clinical Investigation, 2013. **123**(7): p. 3099-3111.
158. Nordlund, J., et al., *Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia.* Genome Biology, 2013. **14**(9): p. 105.
159. Gabriel, A.S., et al., *Epigenetic landscape correlates with genetic subtype but does not predict outcome in childhood acute lymphoblastic leukemia.* Epigenetics, 2015. **10**(8): p. 717-726.
160. Lee, S.-T., et al., *Epigenetic remodeling in B-cell acute lymphoblastic leukemia occurs in two tracks and employs embryonic stem cell-like signatures.* Nucleic Acids Research, 2015. **43**(5): p. 2590-2602.
161. Wahlberg, P., et al., *DNA methylome analysis of acute lymphoblastic leukemia cells reveals stochastic de novo DNA methylation in CpG islands.* Epigenomics, 2016. **8**(10): p. 1367-1387.
162. Nordlund, J., et al., *DNA methylation-based subtype prediction for pediatric acute lymphoblastic leukemia.* Clinical Epigenetics, 2015. **7**(1): p. 11.
163. Figueroa, M.E., et al., *An Integrative Genomic and Epigenomic Approach for the Study of Transcriptional Regulation.* PLoS One, 2008. **3**(3): p. 1882.
164. Janczar, K., et al., *Preserved global histone H4 acetylation linked to ETV6-RUNX1 fusion and PAX5 deletions is associated with favorable outcome in pediatric B-cell progenitor acute lymphoblastic leukemia.* Leukemia Research, 2015. **39**(12): p. 1455-1461.
165. Barnett, K.R., et al., *Epigenomic mapping reveals distinct B cell acute lymphoblastic leukemia chromatin architectures and regulators.* Cell Genomics, 2023. **3**(12): p. 100442.
166. Oudet, P., M. Gross-Bellard, and P. Chambon, *Electron microscopic and biochemical evidence that chromatin structure is a repeating unit.* Cell, 1975. **4**(4): p. 281-300.
167. Brown, D.T., T. Izard, and T. Misteli, *Mapping the interaction surface of linker histone H10 with the nucleosome of native chromatin in vivo.* Nature Structural & Molecular Biology, 2006. **13**(3): p. 250-255.

168. Allan, J., et al., *The structure of histone H1 and its location in chromatin.* Nature, 1980. **288**(5792): p. 675-679.
169. Herrera, J.E., et al., *Histone H1 is a specific repressor of core histone acetylation in chromatin.* Molecular and Cellular Biology, 2000. **20**(2): p. 523-529.
170. Hirano, R., et al., *Structural basis of RNA polymerase II transcription on the chromatosome containing linker histone H1.* Nature Communications, 2022. **13**(1): p. 7287.
171. Healton, S.E., et al., *H1 linker histones silence repetitive elements by promoting both histone H3K9 methylation and chromatin compaction.* Proceedings of the National Academy of Sciences of the United States of America, 2020. **117**(25): p. 14251-14258.
172. Kim, J.-M., et al., *Linker histone H1.2 establishes chromatin compaction and gene silencing through recognition of H3K27me3.* Scientific Reports, 2015. **5**(1): p. 16714.
173. Cao, K., et al., *High-resolution mapping of h1 linker histone variants in embryonic stem cells.* PLoS Genetics, 2013. **9**(4): p. 1003417.
174. Zhou, B.-R., et al., *Distinct Structures and Dynamics of Chromatosomes with Different Human Linker Histone Isoforms.* Molecular Cell, 2021. **81**(1): p. 166-182.
175. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.
176. Sehnal, D., et al., *Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures.* Nucleic Acids Research, 2021. **49**(W1): p. 431-437.
177. Hergeth, S.P. and R. Schneider, *The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle.* EMBO Reports, 2015. **16**(11): p. 1439-1453.
178. Clausell, J., et al., *Histone H1 subtypes differentially modulate chromatin condensation without preventing ATP-dependent remodeling by SWI/SNF or NURF.* PLoS One, 2009. **4**(10): p. 0007243.
179. Mayor, R., et al., *Genome distribution of replication-independent histone H1 variants shows H1.0 associated with nucleolar domains and H1X associated with RNA polymerase II-enriched regions.* Journal of Biological Chemistry, 2015. **290**(12): p. 7474-7491.
180. Izzo, A., et al., *Dynamic changes in H1 subtype composition during epigenetic reprogramming.* Journal of Cell Biology, 2017. **216**(10): p. 3017-3028.
181. Pehrson, J.R. and R.D. Cole, *Histone H1 subfractions and H10 turnover at different rates in nondividing cells.* Biochemistry, 1982. **21**(3): p. 456-60.
182. Gorka, C., J.J. Lawrence, and S. Khochbin, *Variation of H1(0) content throughout the cell cycle in regenerating rat liver.* Experimental Cell Research, 1995. **217**(2): p. 528-533.
183. Morales Torres, C., et al., *The linker histone H1.0 generates epigenetic and functional intratumor heterogeneity.* Science, 2016. **353**(6307).
184. Panyim, S. and R. Chalkley, *A new histone found only in mammalian tissues with little cell division.* Biochemical and Biophysical Research Communications, 1969. **37**(6): p. 1042-1049.
185. Gjerset, R., et al., *Developmental and hormonal regulation of protein H1 degrees in rodents.* Proceedings of the National Academy of Sciences of the United States of America, 1982. **79**(7): p. 2333-2337.
186. Piña, B., et al., *Differential kinetics of histone H1o accumulation in neuronal and glial cells from rat cerebral cortex during postnatal development.* Biochemical and Biophysical Research Communications, 1984. **123**(2): p. 697-702.
187. Valiron, O. and C. Gorka, *Histone H1(0) expression is restricted to progenitor cells during human hematopoiesis.* European Journal of Cell Biology, 1997. **72**(1): p. 39-45.
188. Sirotkin, A.M., et al., *Mice develop normally without the H1(0) linker histone.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(14): p. 6434-6438.

189. Gabrilovich, D.I., et al., *H1° histone and differentiation of dendritic cells. A molecular target for tumor-derived factors.* Journal of Leukocyte Biology, 2002. **72**(2): p. 285-296.
190. Fan, Y., et al., *H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo.* Molecular and Cellular Biology, 2003. **23**: p. 4559-4572.
191. Pan, C. and Y. Fan, *Role of H1 linker histones in mammalian development and stem cell differentiation.* Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2016. **1859**(3): p. 496-509.
192. Mannironi, C., et al., *Histone H1° is synthesized by human lymphocytic leukemia cells but not by normal lymphocytes.* Blood, 1987. **70**(4): p. 1203-1207.
193. Happel, N., et al., *H1 subtype expression during cell proliferation and growth arrest.* Cell Cycle, 2009. **8**(14): p. 2226-2232.
194. Morales Torres, C., et al., *Selective inhibition of cancer cell self-renewal through a Quisinostat-histone H1.0 axis.* Nature Communications, 2020. **11**(1): p. 1792.
195. Concordet, J.-P. and M. Haeussler, *CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens.* Nucleic Acids Research, 2018. **46**(W1): p. 242-245.
196. Vogt, M., et al., *Co-targeting HSP90 alpha and CDK7 overcomes resistance against HSP90 inhibitors in BCR-ABL1+ leukemia cells.* Cell Death & Disease, 2023. **14**(12): p. 799.
197. Davis, M.W. and E.M. Jorgensen, *ApE, a plasmid editor: A freely available DNA manipulation and visualization program.* Frontiers in Bioinformatics, 2022. **2**: p. 818619.
198. Xing, Y., et al., *An expectation-maximization algorithm for probabilistic reconstructions of full-length isoforms from splice graphs.* Nucleic Acids Research, 2006. **34**(10): p. 3150-3160.
199. Heberle, H., et al., *InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams.* BMC Bioinformatics, 2015. **16**(1): p. 169.
200. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome Research, 2003. **13**(11): p. 2498-2504.
201. Reimand, J., et al., *Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap.* Nature Protocols, 2019. **14**(2): p. 482-517.
202. Booeshaghi, A.S., et al., *Depth normalization for single-cell genomics count data.* bioRxiv, 2022.
203. Townes, F.W., et al., *Feature selection and dimension reduction for single-cell RNA-Seq based on a multinomial model.* Genome Biology, 2019. **20**(1): p. 295.
204. Tirosh, I., et al., *Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq.* Science, 2016. **352**(6282): p. 189-196.
205. Jardine, L., et al., *Blood and immune development in human fetal bone marrow and Down syndrome.* Nature, 2021. **598**(7880): p. 327-331.
206. Downing, J.R., et al., *The Pediatric Cancer Genome Project.* Nature Genetics, 2012. **44**(6): p. 619-622.
207. McLeod, C., et al., *St. Jude Cloud: A Pediatric Cancer Genomic Data-Sharing Ecosystem.* Cancer Discovery, 2021. **11**(5): p. 1082-1099.
208. Kohlmann, A., et al., *An international standardization programme towards the application of gene expression profiling in routine leukaemia diagnostics: the Microarray Innovations in LEukemia study prephase.* British Journal of Haematology, 2008. **142**(5): p. 802-807.
209. Polak, R., et al., *Autophagy inhibition as a potential future targeted therapy for ETV6-RUNX1-driven B-cell precursor acute lymphoblastic leukemia.* Haematologica, 2019. **104**(4): p. 738-748.

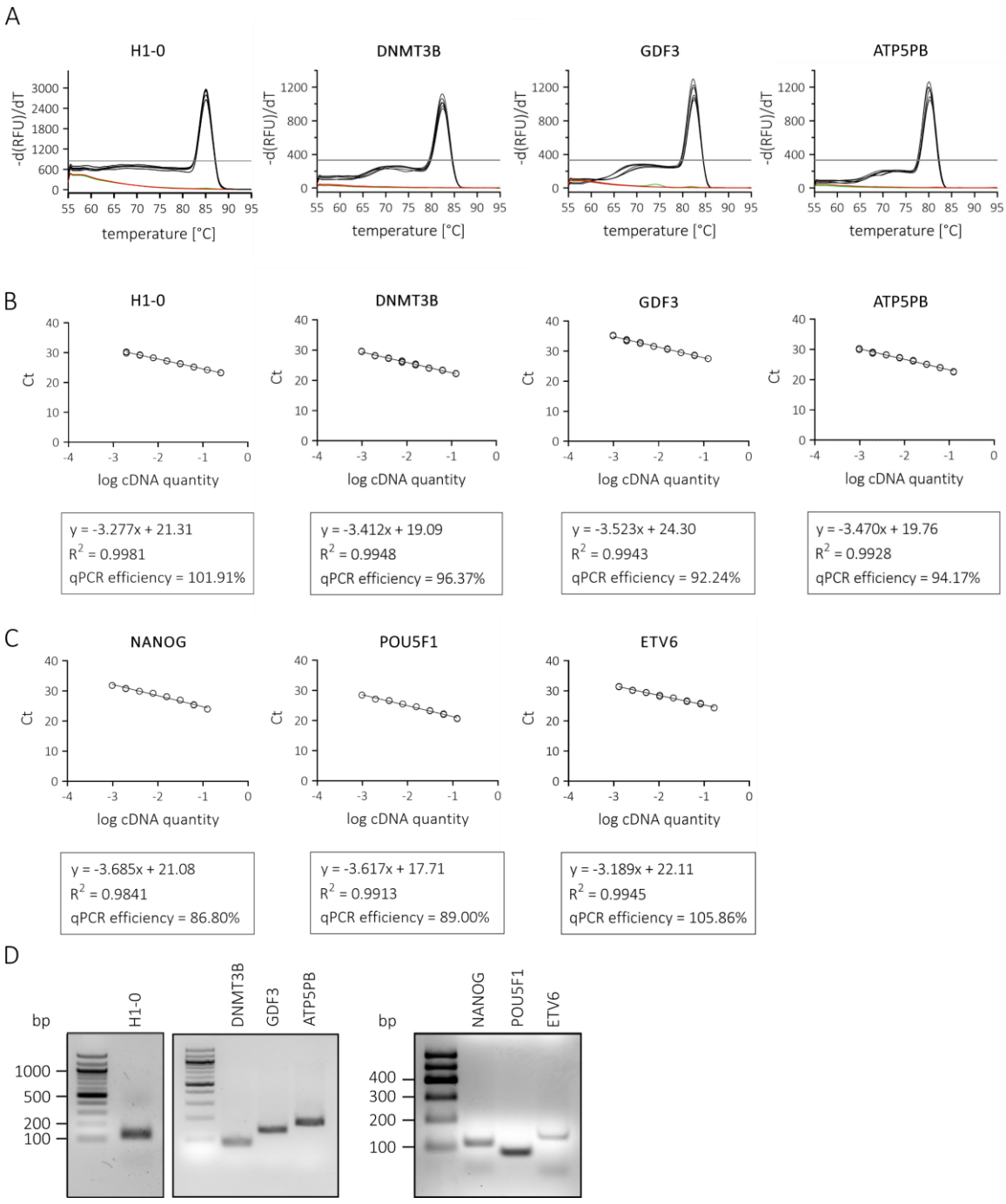
210. Jerchel, I.S., et al., *High PDGFRA expression does not serve as an effective therapeutic target in ERG-deleted B-cell precursor acute lymphoblastic leukemia.* Haematologica, 2018. **103**(2): p. 73-77.
211. Novershtern, N., et al., *Densely interconnected transcriptional circuits control cell states in human hematopoiesis.* Cell, 2011. **144**(2): p. 296-309.
212. Black, K.L., et al., *Aberrant splicing in B-cell acute lymphoblastic leukemia.* Nucleic Acids Res, 2018. **46**(21): p. 11357-11369.
213. Mehtonen, J., et al., *Single cell characterization of B-lymphoid differentiation and leukemic cell states during chemotherapy in ETV6-RUNX1-positive pediatric leukemia identifies drug-targetable transcription factor activities.* Genome Medicine, 2020. **12**(1): p. 99.
214. Popescu, D.-M., et al., *Decoding human fetal liver haematopoiesis.* Nature, 2019. **574**(7778): p. 365-371.
215. Jakobczyk, H., et al., *Reduction of RUNX1 transcription factor activity by a CBFA2T3-mimicking peptide: application to B cell precursor acute lymphoblastic leukemia.* Journal of Hematology & Oncology, 2021. **14**(1): p. 47.
216. Robinson, J.T., et al., *Integrative genomics viewer.* Nature Biotechnology, 2011. **29**(1): p. 24-26.
217. Tsuzuki, S. and M. Seto, *TEL (ETV6)-AML1 (RUNX1) initiates self-renewing fetal pro-B cells in association with a transcriptional program shared with embryonic stem cells in mice.* Stem Cells, 2013. **31**(2): p. 236-247.
218. Bernardin, F., et al., *TEL-AML1, Expressed from t(12;21) in Human Acute Lymphocytic Leukemia, Induces Acute Leukemia in Mice1.* Cancer Research, 2002. **62**(14): p. 3904-3908.
219. Rodríguez-Hernández, G., et al., *The second oncogenic hit determines the cell fate of ETV6-RUNX1 positive leukemia.* Frontiers in Cell and Developmental Biology, 2021. **9**: p. 704591.
220. Bae, S., J. Park, and J.S. Kim, *Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases.* Bioinformatics, 2014. **30**(10): p. 1473-1475.
221. Linka, Y., et al., *The impact of TEL-AML1 (ETV6-RUNX1) expression in precursor B cells and implications for leukaemia using three different genome-wide screening methods.* Blood Cancer Journal, 2013. **3**(10): p. 151.
222. Li, J., et al., *Discrepant mRNA and protein expression in immune cells.* Current Genomics, 2020. **21**(8): p. 560-563.
223. Kress, H., R. Tönjes, and D. Doenecke, *Butyrate induced accumulation of a 2.3 kb polyadenylated H1(0) histone mRNA in HeLa cells.* Nucleic Acids Research, 1986. **14**(18): p. 7189-7197.
224. Girardot, V., et al., *Relationship between core histone acetylation and histone H10 gene activity.* European Journal of Biochemistry, 1994. **224**(3): p. 885-892.
225. Kaindl, U., et al., *Blocking ETV6/RUNX1-induced MDM2 overexpression by Nutlin-3 reactivates p53 signaling in childhood leukemia.* Leukemia, 2014. **28**(3): p. 600-608.
226. Chen, D., et al., *RAG1 co-expression signature identifies ETV6-RUNX1-like B-cell precursor acute lymphoblastic leukemia in children.* Cancer Medicine, 2021. **10**(12): p. 3997-4003.
227. Leo, I.R., et al., *Integrative multi-omics and drug response profiling of childhood acute lymphoblastic leukemia cell lines.* Nature Communications, 2022. **13**(1): p. 1691.
228. Aswad, L. and R. Jafari, *FORALL: an interactive shiny/R web portal to navigate multi-omics high-throughput data of pediatric acute lymphoblastic leukemia.* Bioinformatics Advances, 2023. **3**(1).
229. Pui, C.H., et al., *Biology, risk stratification, and therapy of pediatric acute leukemias: an update.* Journal of Clinical Oncology, 2011. **29**(5): p. 551-565.

230. Kodgule, R., et al., *ETV6 deficiency unlocks ERG-dependent microsatellite enhancers to drive aberrant gene activation in B-lymphoblastic leukemia.* Blood Cancer Discovery, 2023. **4**(1): p. 34-53.
231. Ruiz, J.P., et al., *Robust generation of erythroid and multilineage hematopoietic progenitors from human iPSCs using a scalable monolayer culture system.* Stem Cell Research, 2019. **41**: p. 101600.
232. Tsuzuki, S., et al., *Modeling first-hit functions of the t(12;21) TEL-AML1 translocation in mice.* Proceedings of the National Academy of Sciences, 2004. **101**(22): p. 8443-8448.
233. Eldeeb, M., et al., *Ontogeny shapes the ability of ETV6::RUNX1 to enhance hematopoietic stem cell self-renewal and disrupt early lymphopoiesis.* Leukemia, 2024. **38**(2): p. 455-459.
234. Signer, R.A.J., et al., *Haematopoietic stem cells require a highly regulated protein synthesis rate.* Nature, 2014. **509**(7498): p. 49-54.
235. Gulati, G.S., et al., *Single-cell transcriptional diversity is a hallmark of developmental potential.* Science, 2020. **367**(6476): p. 405-411.
236. Khan, A.O., et al., *Human bone marrow organoids for disease modeling, discovery, and validation of therapeutic targets in hematologic malignancies.* Cancer Discovery, 2023. **13**(2): p. 364-385.
237. Frenz-Wiessner, S., et al., *Generation of complex bone marrow organoids from human induced pluripotent stem cells.* Nature Methods, 2024. **21**(5): p. 868-881.
238. Suzuki, N., et al., *Generation of engraftable hematopoietic stem cells from induced pluripotent stem cells by way of teratoma formation.* Molecular Therapy, 2013. **21**(7): p. 1424-1431.
239. Sugimura, R., et al., *Haematopoietic stem and progenitor cells from human pluripotent stem cells.* Nature, 2017. **545**(7655): p. 432-438.
240. Nafria, M., et al., *Expression of RUNX1-ETO rapidly alters the chromatin landscape and growth of early human myeloid precursor cells.* Cell Reports, 2020. **31**(8).
241. Ebinger, S., et al., *Characterization of rare, dormant, and therapy-resistant cells in acute lymphoblastic leukemia.* Cancer Cell, 2016. **30**(6): p. 849-862.
242. Linka, Y., et al., *Identification of TEL-AML1 (ETV6-RUNX1) associated DNA and its impact on mRNA and protein output using ChIP, mRNA expression arrays and SILAC.* Genomics Data, 2014. **2**: p. 85-88.
243. Chabanas, A., et al., *Cell cycle regulation of histone H1O in CHO cells: a flow cytofluorimetric study after double staining of the cells.* The EMBO Journal, 1983. **2**(6): p. 833-837.
244. Kalashnikova, A.A., et al., *Linker histone H1.0 interacts with an extensive network of proteins found in the nucleolus.* Nucleic Acids Research, 2013. **41**(7): p. 4026-4035.
245. Szklarczyk, D., et al., *The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest.* Nucleic Acids Research, 2023. **51**(D1): p. 638-646.
246. Thorslund, T., et al., *Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage.* Nature, 2015. **527**(7578): p. 389-393.
247. Murga, M., et al., *Global chromatin compaction limits the strength of the DNA damage response.* Journal of Cell Biology, 2007. **178**(7): p. 1101-1108.
248. Zhang, Y., et al., *Current treatment strategies targeting histone deacetylase inhibitors in acute lymphocytic leukemia: a systematic review.* Frontiers in Oncology, 2024. **14**.
249. Ho, T.C.S., A.H.Y. Chan, and A. Ganesan, *Thirty years of HDAC inhibitors: 2020 insight and hindsight.* Journal of Medicinal Chemistry, 2020. **63**(21): p. 12460-12484.
250. Arts, J., et al., *JNJ-26481585, a novel second-generation oral histone deacetylase inhibitor, shows broad-spectrum preclinical antitumoral activity.* Clinical Cancer Research, 2009. **15**(22): p. 6841-6851.

251. Venugopal, B., et al., *A Phase I Study of Quisinostat (JNJ-26481585), an Oral Hydroxamate Histone Deacetylase Inhibitor with Evidence of Target Modulation and Antitumor Activity, in Patients with Advanced Solid Tumors.* Clinical Cancer Research, 2013. **19**(15): p. 4262-4272.
252. Araki, H., et al., *Chromatin-modifying agents permit human hematopoietic stem cells to undergo multiple cell divisions while retaining their repopulating potential.* Blood, 2006. **109**(8): p. 3570-3578.
253. Young, J.C., et al., *Inhibitors of histone deacetylases promote hematopoietic stem cell self-renewal.* Cytotherapy, 2004. **6**(4): p. 328-336.
254. Deleu, S., et al., *Bortezomib alone or in combination with the histone deacetylase inhibitor JNJ-26481585: Effect on myeloma bone disease in the 5T2MM murine model of myeloma.* Cancer Research, 2009. **69**(13): p. 5307-5311.
255. Bastian, L., et al., *Synergistic activity of bortezomib and HDACi in preclinical models of B-cell precursor acute lymphoblastic leukemia via modulation of p53, PI3K/AKT, and NF-κB.* Clinical Cancer Research, 2013. **19**(6): p. 1445-1457.
256. Wang, T., et al., *Sequential CRISPR gene editing in human iPSCs charts the clonal evolution of myeloid leukemia and identifies early disease targets.* Cell Stem Cell, 2021. **28**(6): p. 1074-1089.
257. Hasle, H., et al., *Germline GATA1s-generating mutations predispose to leukemia with acquired trisomy 21 and Down syndrome-like phenotype.* Blood, 2022. **139**(21): p. 3159-3165.
258. Hock, H. and A. Shimamura, *ETV6 in hematopoiesis and leukemia predisposition.* Seminars in Hematology, 2017. **54**(2): p. 98-104.
259. Richardson, S.E., et al., *In vitro differentiation of human pluripotent stem cells into the B lineage using OP9-MS5 co-culture.* STAR Protocols, 2021. **2**(2): p. 100420.

8. APPENDIX

8.1 Supplementary Figures



Supplementary Figure 8.1: Testing RT-qPCR primer specificities and efficiencies. (A) Melt curve analyses of RT-qPCR primer pairs for detection of *H1-0*, *DNMT3B*, *GDF3* and *ATP5BP* expression. No-reverse transcriptase controls are indicated in red, no-template controls are indicated in green. **(B,C)** 1:2 serial dilution standard curves for RT-qPCR primer pairs detecting (B) *H1-0*, *DNMT3B*, *GDF3* and *ATP5BP* or (C) probe-containing RT-qPCR assays detecting *NANOG*, *PU5F1* and *ETV6* expression. Regression line formulas, R^2 values and efficiencies are indicated. **(D)** Assessment of RT-qPCR product lengths and purity using agarose gel electrophoresis.

8.2 Supplementary Tables

Supplementary Table 8.1: Gene set enrichment (GSEA) results of canonical pathways in REH cells treated with non-targeting compared to *H1-0*-targeting siRNA. Gene sets are filtered for NES<-1.5. NES: normalized enrichment score, FDR: false discovery rate.

gene set name	gene set size	ES	NES	FDR q-value
REACTOME_HDACS_DEACETYLATE_HISTONES	46	-0.87	-2.30	<0.001
REACTOME_POSITIVE_EPIGENETIC_REGULATION_OF_ RRNA_EXPRESSION	65	-0.79	-2.27	<0.001
REACTOME_MEIOSIS	63	-0.78	-2.25	<0.001
REACTOME_MITOTIC_PROPHASE	98	-0.72	-2.25	<0.001
REACTOME_TRANSCRIPTIONAL_REGULATION_BY_SMALL _RNAS	65	-0.80	-2.23	<0.001
REACTOME_REPRODUCTION	69	-0.76	-2.23	<0.001
REACTOME_HATS_ACETYLATE_HISTONES	92	-0.76	-2.23	<0.001
REACTOME_ERCC6_CSB_AND_EHMT2_G9A_POSITIVELY_ REGULATE_RRNA_EXPRESSION	35	-0.86	-2.21	<0.001
REACTOME_FORMATION_OF_THE_BETA_CATENIN_TCF_ TRANSACTIVATING_COMPLEX	46	-0.82	-2.19	<0.001
REACTOME_PRC2_METHYLATES_HISTONES_AND_DNA	32	-0.87	-2.19	<0.001
REACTOME_B_WICH_COMPLEX_POSITIVELY_REGULATES_ RRNA_EXPRESSION	50	-0.82	-2.19	<0.001
REACTOME_CONDENSATION_OF_PROPHASE_ CHROMOSOMES	32	-0.87	-2.19	<0.001
REACTOME_RNA_POLYMERASE_I_TRANSCRIPTION	69	-0.75	-2.18	<0.001
REACTOME_SIRT1_NEGATIVELY_REGULATES_RRNA_ EXPRESSION	27	-0.88	-2.17	<0.001
REACTOME_GENE_SILENCING_BY_RNA	84	-0.73	-2.16	<0.001
REACTOME_NEGATIVE_EPIGENETIC_REGULATION_OF_ RRNA_EXPRESSION	68	-0.76	-2.16	<0.001
REACTOME_MEIOTIC_RECOMBINATION	41	-0.84	-2.15	<0.001
REACTOME_ASSEMBLY_OF_THE_ORC_COMPLEX_AT_THE _ORIGIN_OF_REPLICATION	28	-0.87	-2.15	<0.001
REACTOME_MEIOTIC_SYNAPSIS	38	-0.81	-2.14	<0.001
REACTOME_BASE_EXCISION_REPAIR	60	-0.75	-2.14	<0.001
REACTOME_RUNX1_REGULATES_TRANSCRIPTION_OF_ GENES_INVOLVED_IN_DIFFERENTIATION_OF_HSCS	81	-0.74	-2.12	<0.001
REACTOME_NONHOMOLOGOUS_END_JOINING_NHEJ	42	-0.80	-2.11	<0.001
REACTOME_EPIGENETIC_REGULATION_OF_GENE_ EXPRESSION	106	-0.69	-2.10	<0.001
REACTOME_RNA_POLYMERASE_I_PROMOTER_ESCAPE	50	-0.80	-2.10	<0.001
REACTOME_DEPOSITION_OF_NEW_CENPA_CONTAINING _NUCLEOSOMES_AT_THE_CENTROMERE	43	-0.79	-2.10	<0.001
REACTOME_DNA_REPLICATION_PRE_INITIATION	118	-0.68	-2.10	<0.001
REACTOME_DNA_METHYLATION	23	-0.89	-2.10	8.34×10^{-5}
REACTOME_PROCESSING_OF_DNA_DOUBLE_STRAND_ BREAK_ENDS	70	-0.74	-2.08	8.04×10^{-5}
REACTOME_DISEASES_OF_PROGRAMMED_CELL_DEATH	58	-0.75	-2.06	7.76×10^{-5}
REACTOME_G2_M_DNA_DAMAGE_CHECKPOINT	67	-0.73	-2.05	7.50×10^{-5}

REACTOME_INHIBITION_OF_DNA_RECOMBINATION_AT_TELOMERE	37	-0.82	-2.05	$7.26x10^{-5}$
REACTOME_E3_UBIQUITIN_LIGASES_UBIQUITINATE_TARGET_PROTEINS	48	-0.77	-2.05	$1.07x10^{-4}$
REACTOME_DNA_DAMAGE_TELOMERE_STRESS_INDUCED_SENESCENCE	44	-0.77	-2.04	$1.39x10^{-4}$
REACTOME_TELOMERE_MAINTENANCE	81	-0.68	-2.04	$1.35x10^{-4}$
REACTOME_DNA_REPLICATION	143	-0.65	-2.03	$1.65x10^{-4}$
REACTOME_ACTIVATED_PKN1_STIMULATES_TRANSCRIPTION_OF_AR_ANDROGEN_RECEPTOR_REGULATED_GENES_KLK2_AND_KLK3	23	-0.89	-2.03	$1.91x10^{-4}$
REACTOME_DNA_DOUBLE_STRAND_BREAK_RESPONSE	50	-0.76	-2.03	$1.86x10^{-4}$
REACTOME_BASE_EXCISION_REPAIR_AP_SITE_FORMATION	32	-0.81	-2.02	$3.36x10^{-4}$
REACTOME_OXIDATIVE_STRESS_INDUCED_SENESCENCE	78	-0.69	-2.01	$4.17x10^{-4}$
REACTOME_HCMV_LATE_EVENTS	64	-0.71	-2.01	$4.07x10^{-4}$
REACTOME_TRANSCRIPTIONAL_REGULATION_OF_GRANULOPOIESIS	45	-0.77	-2.00	$4.55x10^{-4}$
REACTOME_PROTEIN_UBIQUITINATION	68	-0.71	-2.00	$4.44x10^{-4}$
REACTOME_ESTROGEN_DEPENDENT_GENE_EXPRESSION	98	-0.66	-1.99	$4.62x10^{-4}$
REACTOME_ACTIVATION_OF_ANTERIOR_HOX_GENES_IN_HINDBRAIN_DEVELOPMENT_DURING_EARLY_EMBRYOGENESIS	67	-0.71	-1.99	$4.77x10^{-4}$
REACTOME_SENESCENCE_ASSOCIATED_SECRETORY_PHENOTYPE_SASP	64	-0.69	-1.98	$8.56x10^{-4}$
REACTOME_HCMV_EARLY_EVENTS	76	-0.67	-1.97	0.0011
REACTOME_HOMOLOGY_DIRECTED_REPAIR	111	-0.62	-1.93	0.0023
REACTOME_CHROMOSOME_MAINTENANCE	108	-0.62	-1.93	0.0025
REACTOME_RECOGNITION_AND_ASSOCIATION_OF_DNA_GLYCOSYLASE_WITH_SITE_CONTAINING_AN_AFFECTED_PURINE	25	-0.82	-1.93	0.0026
REACTOME_RMTS_METHYLATE_HISTONE_ARGININES	40	-0.75	-1.92	0.0027
REACTOME_PKMTS_METHYLATE_HISTONE_LYSINES	46	-0.72	-1.92	0.0028
REACTOME_RUNX1_REGULATES_GENES_INVOLVED_IN_MEGAKARYOCYTE_DIFFERENTIATION_AND_PLATELET_FUNCTION	48	-0.72	-1.90	0.0045
REACTOME_RHO_GTPASES_ACTIVATE_PKNS	46	-0.71	-1.89	0.0050
REACTOME_G2_M_CHECKPOINTS	138	-0.59	-1.85	0.0096
REACTOME_CHROMATIN_MODIFYING_ENZYMES	209	-0.56	-1.81	0.0182
REACTOME_PRE_NOTCH_EXPRESSION_AND_PROCESSING	64	-0.64	-1.81	0.0194
KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS	48	-0.66	-1.80	0.0208
REACTOME_METALLOPROTEASE_DUBS	24	-0.73	-1.76	0.0372
REACTOME_HDMS_DEMETHYLATE_HISTONES	25	-0.74	-1.75	0.0389
REACTOME_AMYLOID_FIBER_FORMATION	47	-0.66	-1.74	0.0432
REACTOME_CELLULAR_SENESCENCE	139	-0.55	-1.74	0.0440
REACTOME_HCMV_INFECTION	97	-0.57	-1.71	0.0595
REACTOME_DNA_DOUBLE_STRAND_BREAK_REPAIR	137	-0.54	-1.71	0.0620
WP_EFFECT_OF_PROGERIN_ON_GENES_INVOLVED_IN_HUTCHINSONGILFORD_PROGERIA_SYNDROME	26	-0.71	-1.70	0.0646
KEGG_HEDGEHOG_SIGNALING_PATHWAY	25	-0.67	-1.68	0.0828

REACTOME_ESR_MEDIATED_SIGNALING	148	-0.52	-1.63	0.1246
WP_COMPLEMENT_SYSTEM	32	-0.65	-1.62	0.1396
REACTOME_INWARDLY_RECTIFYING_K_CHANNELS	15	-0.74	-1.60	0.1589
REACTOME_SIGNALING_BY_NUCLEAR_RECEPTORS	192	-0.48	-1.59	0.1656
WP_SPLICING_FACTOR_NOVA_REGULATED_SYNAPTIC_ PROTEINS	23	-0.70	-1.58	0.1752
KEGG_BETA_ALANINE_METABOLISM	16	-0.73	-1.58	0.1823
REACTOME_DEGRADATION_OF_THE_EXTRACELLULAR_ MATRIX	47	-0.59	-1.56	0.1991
REACTOME_UB_SPECIFIC_PROCESSING_PROTEASES	160	-0.48	-1.56	0.2108
PID_INTEGRIN1_PATHWAY	19	-0.69	-1.55	0.2212
PID_HNF3B_PATHWAY	18	-0.69	-1.52	0.2596
WP_HISTONE_MODIFICATIONS	42	-0.58	-1.52	0.2647
PID_PI3KCI_PATHWAY	45	-0.57	-1.51	0.2802

Supplementary Table 8.2: Gene set enrichment (GSEA) results of canonical pathways in REH cells treated with *H1-0*-targeting compared to non-targeting siRNA. Gene sets are filtered for NES>1.5. NES: normalized enrichment score, FDR: false discovery rate.

gene set name	gene set size	ES	NES	FDR q-value
REACTOME_REGULATION_OF_RUNX1_EXPRESSION_AND _ACTIVITY	16	0.87	1.69	1
PID_ECADHERIN_STABILIZATION_PATHWAY	32	0.78	1.67	1
REACTOME_ABERRANT_REGULATION_OF_MITOTIC_G1_S _TRANSITION_IN_CANCER_DUE_TO_RB1_DEFECTS	15	0.83	1.64	1
BIOCARTA_CELLCYCLE_PATHWAY	19	0.81	1.62	1
REACTOME_ASSOCIATION_OF_TRIC_CCT_WITH_TARGET _PROTEINS_DURING_BIOSYNTHESIS	36	0.73	1.62	1
REACTOME_SARS_COV_1_ACTIVATES_MODULATES_ INNATE_IMMUNE_RESPONSES	39	0.73	1.61	1
WP_INTERACTIONS_BETWEEN_IMMUNE_CELLS_AND_ MICRORNAS_IN_TUMOR_MICROENVIRONMENT	15	0.83	1.60	1
REACTOME_ASYMMETRIC_LOCALIZATION_OF_PCP_ PROTEINS	60	0.67	1.58	1
WP_MITOCHONDRIAL_IMMUNE_RESPONSE_TO_ SARSCOV2	26	0.76	1.58	1
REACTOME_COOPERATION_OF_PREFOLDIN_AND_TRIC_ CCT_IN_ACTIN_AND_TUBULIN_FOLDING	24	0.75	1.58	1
WP_SPINAL_CORD_INJURY	57	0.68	1.58	0.9917
BIOCARTA_NGF_PATHWAY	18	0.78	1.56	1
REACTOME_DEGRADATION_OF_DVL	55	0.66	1.55	1
WP_STING_PATHWAY_IN_KAWASAKILIKE_DISEASE_AND_ COVID19	19	0.78	1.55	1
PID_ARF6_PATHWAY	23	0.74	1.54	1
BIOCARTA_TOLL_PATHWAY	24	0.75	1.54	1
REACTOME_SIGNALING_BY_CSF1_M_CSF_IN_MYELOID_ CELLS	28	0.72	1.53	1

SA_PTEN_PATHWAY	17	0.78	1.53	1
REACTOME_CYCLIN_D_ASSOCIATED_EVENTS_IN_G1	42	0.68	1.53	1
REACTOME_FORMATION_OF_TUBULIN_FOLDING_INTERMEDIATES_BY_CCT_TRIC	17	0.78	1.52	1
REACTOME_NEPHRIN_FAMILY_INTERACTIONS	18	0.76	1.52	1
PID_MAPK_TRK_PATHWAY	31	0.69	1.52	1
WP_MEASLES_VIRUS_INFECTION	100	0.60	1.51	1
REACTOME_DAP12_SIGNALING	24	0.73	1.51	1
REACTOME_PROTEIN_FOLDING	75	0.63	1.51	1
REACTOME_DISEASES_OF_MITOTIC_CELL_CYCLE	36	0.68	1.50	1
BIOCARTA_ETS_PATHWAY	16	0.79	1.50	0.9994

Supplementary Table 8.3: Ingenuity Pathway Analysis (IPA, QIAGEN) of upstream regulators in siH1-0_1 versus siCtrl REH cells. Depicted are filtered upstream regulators with p<0.05.

upstream regulator	activation z-score	p-value of overlap	target molecules in dataset
ETV6-RUNX1	-4.885	$3.73x10^{-16}$	ABCG1,ARHGAP4,CACNA1A,CALN1,CCL5,CD70,CD72, CHI3L2,CXXC5,DDIT4,IRF7,ISG15,LGALS1,MME,MPO, NPY,PCP4,PDGFRB,PTGDR,S100A4,SMIM3,SNAP91,SPIB, TLR4,TNFRSF10B,TNFSF9,ZMAT3
SOX2	-1.448	0.00508	BHLHE40,CCL5,CD70,CD9,CDKN1A,CSF1R,DPY30, EFEMP1,IRS1,NCKAP1,PLD1,TCL1A,TEC
NFE2L2	-0.726	0.0197	ARHGEF3,CCL5,CDKN1A,ELN,GSTA4,HBEGF,ISG15, MMP14,NCKAP1,SRGN,STING1,TMED2
AHR	-0.566	0.0121	ACTA2,CCL5,CD84,CDKN1A,EFEMP1,ELN,FAS,ITGAM, LCK,PCK1,PDGFRB,TUBA8
IL4	-0.350	0.00000499	ACTA2,ADAP1,ARID5A,ASS1,BHLHE40,CCL5,CCND2, CD163L1,CD72,CDKN1A,CNR2,CSF1R,DDIT4,DUSP10 ,FAS,FCGR2A,H1-0,HBEGF,HGF,HLX,IRF7,ISG15,LGALS1,LSR,MIR155HG, MSR1,NDRG1,PINK1,PKP3,PLD1,SEMA7A,TLR4,ZBTB46
PRDM1	-0.323	0.0000353	ATP9A,DDIT4,FNDC5,IRF7,ITGAM,LGALS1,METTL7A, MIR155HG,MS4A4A,NPAS1,SPIB,SULT1A3/SULT1A4, TNFSF9
IL10	0.091	0.000000467	ACTA2,CCL5,CCND2,CD163L1,CD72,CDKN1A,CHI3L2, CSF1R,CTSH,FAS,FCGR2A,HGF,ITGAM,KLRC4-KLRK1/KLRK1,LGALS1,LILRB2,MMP14,MSR1,TLR4,TLR7
TCF12	0.138	0.000202	CDKN1A,CSF1R,LCK,NGFR,RXRA,SEMA7A,SPIB,SULT1A3/ SULT1A4
IRF8	0.179	0.0000141	CCL5,CDKN1A,CSF1R,DRD5,FAS,GABRR2,IRF7,ISG15, ITGAM,MSR1,TLR4
PTEN	0.568	0.000166	ABHD4,AKR1C3,BBC3,CCL5,CDKN1A,FAS,GALK1,GMDS, HGF,ISG15,LGR6,MME,MMP14,NDRG1,NGFR,P2RY14, PDGFRB,PERP,PINK1,RAB43,SULF2
APP	0.704	$7.35x10^{-10}$	ABCG1,ACTA2,ADAP1,BAD,BBC3,CCL5,CCND2,CD82, CDKN1A,CSF1R,ELN,FCGR2A,GALK1,HECW1,HGF,IRF7, ITGAM,KCNE3,MGAT3,MME,MMP14,MPO,MSR1,NGFR, NPY,PCP4,PDGFRB,PKP3,PRKACA,TLR4,TLR7,TNFRSF10B, TP53INP1,TUBA8,TUBB2A,VASP

IL15	0.713	0.00000188	BBC3,CCL5,CCND2,CD84,CDKN1A,FAS,GLDC,HBEGF,HLX, IL1RAP,ISG15,KLRC4-KLRK1/KLRK1,LCK,LSR,PLD1,PRKACA,SRGN,SYNPO, TNFRSF10B
TCF3	0.955	0.00322	ARHGAP4,CCND2,CDKN1A,CSF1R,H1-0,LCK,MXD3,OPTN, RXRA,SEMA7A,SPIB,SULT1A3/SULT1A4
IKBKB	0.971	0.00246	ACTA2,CCL5,CDKN1A,CSF1R,FAS,HGF,IGFBP7,ISG15, LIPE,S100A4
HMG20A	1.134	0.000265	CCL5,CD9,ELN,HBEGF,IRF7,MMP14,PDGFRB
ESR1	1.279	0.000418	AASS,ACTA2,ASS1,ATF5,C1QTNF4,CAP2,CCDC85B, CDKN1A,DDIT4,DUSP10,EFEMP1,ELN,FAS,FMN1,HBEGF, IRS1,KLRC4-KLRK1/KLRK1,LCK,LGALS1,METTL7A,MME,NCKAP1, NDUFB7,NGFR,NPY,PCP4,SCAND1,SYNPO,TNFSF9, TRPM2,TSPO
TNFSF11	1.332	0.0015	CCL5,CDKN1A,CSF1R,FAS,ITGAM,LSR,PHLDA3,PLD1, SPIB,TLR4
ITPR2	1.342	0.00149	CCL5,CSF1R,PCK1,SEMA7A,TLR7
BHLHE40	1.667	0.0119	ACTA2,CD82,CD9,FAS,HBEGF,MMP14,MSR1,NDRG1, P2RY14
POU5F1	1.674	0.00217	ATP6V0A4,BAD,CCL5,CD70,CDKN1A,EFEMP1,FAS,GAA, IRS1,NCKAP1,TCL1A,TNFRSF10B
CEBPA	1.796	0.00000476	AKR1C3,ASS1,CCND2,CDKN1A,CSF1R,H1-10,HGF,IRS1, ISG15,ITGAM,KCNMB1,LCK,LGALS1,MPO,PCK1,SULT1A3 /SULT1A4,TUBB2A,ZNF296
CTNNB1	1.946	0.000107	ACTA2,BBC3,BHLHE40,CCL5,CCND2,CDKN1A,CTHRC1, DPEP1,FAS,GRAMD4,HBEGF,IRS1,LCK,LGR6,LIPE, MECOM,MME,MMP14,NGFR,NME3,PCK1,PKP3,RPS27L, S100A4,TCL1A,VASP
VEGFA	1.988	0.0259	ACTA2,CDKN1A,HBEGF,HLX,LIPE,MVP,PDGFRB,TSPO
TNF	2.137	0.000000729	ACTA2,APOBEC3B,ASS1,BBC3,BCYRN1,BHLHE40,CCL5, CCND2,CD163L1,CD70,CD82,CDKN1A,CHI3L2,CMBL, CNR2,CSF1R,DUSP10,EGFL7,FAS,HBEGF,HGF,IRF7,IRS1, ISG15,ITGAM,KIAA1671,LIPE,LRIG1,MECOM,MMP14, MPO,MSR1,MVP,NGFR,NME3,OPTN,PCDH10,PCK1, PCP4,RXRA,SKI,SYNPO,TLR4,TLR7,TNFRSF10B,TNFSF9, TP53INP1,VASP
IL2	2.272	0.00000283	ARHGEF3,ATF5,BBC3,BHLHE40,CCL5,CCND2,CD9, CDKN1A,CSF1R,DDIT4,DRD5,FAS,HBEGF,ISG15,KLRC4-KLRK1/KLRK1,LCK,LSR,MME,NDRG1,RXRA,S100A4,SLA, TNFRSF10B,TNFSF9,TP53INP1
IFNG	2.368	0.0000875	ACTA2,ADA2,ASS1,BBC3,BCYRN1,CCL5,CCND2,CD163L1, CD72,CDKN1A,CHI3L2,CSF1R,CTSH,ELN,FAS,FCGR2A, H6PD,HAAO,HBEGF,IRF7,IRS1,ISG15,ITGAM,KLRC4-KLRK1/KLRK1,MSR1,OPTN,P2RY14,PCK1,PCTP,PLD1, SEMA7A,STING1,TLR4,TLR7,TNFRSF10B
AGT	2.394	0.0000744	ACTA2,ATP6V0A4,BHLHE40,CCL5,CCND2,CDKN1A, CTHRC1,DRD5,EFEMP1,ELN,FDXR,GAA,HBEGF,HGF, IGFBP7,IRF7,IRS1,MSR1,NDRG1,NPY,PDGFRB,PTGDR, RGMA,SRGAP3,TMIE
ZBTB10	2.449	0.00911	ARID5A,CCL5,ISG15,OPTN,SEMA7A,STING1
IL18	2.557	0.0000952	ADA2,CCL5,DDIT4,FAS,HLX,ITGAM,KLRC4-KLRK1/KLRK1, MPO,NPY,TLR4

TP53	4.365	0.000000172	ABHD4,ACTA2,ASS1,ASTN2,ATXN1,BAD,BBC3,BHLHE40, CCL5,CCND2,CD70,CD82,CDKN1A,CMBL,CSF1R,CTSH, DDIT4,DGCR6/LOC102724770,DPEP1,FAS,FDXR,HBEGF, HGF,IGFBP7,IRF7,IRS1,ISG15,LAPTM4A,MIR155HG,MVP, NDRG1,P2RY14,PCK1,PDGFRB,PERP,PHLDA3,PINK1, PRAG1,RPS27L,S100A4,SRGAP3,SULF2,TCL1A,TLR7, TNFRSF10B,TNFSF9,TP53INP1,TSPO,ZMAT3

Supplementary Table 8.4: Ingenuity Pathway Analysis (IPA, QIAGEN) of upstream regulators in siH1-0_2 versus siCtrl REH cells. Depicted are filtered upstream regulators with p<0.05.

upstream regulator	activation z-score	p-value of overlap	target molecules in dataset
PRDM1	-1.673	0.000131	ADGRG1,FTH1,MAP7D2,RGS16,SPIB,ZAP70
NFE2L2	-1.423	0.0137	CCL5,ELN,FTH1,NCKAP1,STING1
ETV6-RUNX1	-1.414	$3.54x10^{-11}$	CCL5,FTH1,MS4A1,PCP4,PDGFRB,PECAM1,PTGDR,RGS16, SAMHD1,SPIB,TSC22D3,ZBTB20
SOX2	-1.131	0.0442	CCL5,CSF1R,NCKAP1,PECAM1
IL4	-0.758	0.0163	CCL5,CSF1R,H1-0,HLX,MS4A1,SEMA7A,TLN1,ZAP70
AHR	-0.025	0.0104	CCL5,ELN,PDGFRB,TUBA8,ZAP70
POU5F1	0.055	0.0216	BAD,CCL5,NCKAP1,PECAM1
VEGFA	0.106	0.00933	CTSS,HLX,PDGFRB,PECAM1
APP	0.214	0.00000162	BAD,CCL5,CSF1R,ELN,FTH1,HECW1,HSPE1,MGAT3,NGFR, PCP4,PDGFRB,TLR7,TUBA8
PTEN	0.218	0.000857	CCL5,FTH1,LGR6,NGFR,P2RY14,PDGFRB,PECAM1,TSC22D3
IL10	0.31	0.000238	CCL5,CSF1R,CTSS,KLRC4-KLRK1/KLRK1,RGS16,TLR7,TSC22D3
IFNG	0.439	0.0000104	ADGRG1,CCL5,CSF1R,CTSS,ELN,FTH1,KLRC4-KLRK1/KLRK1, P2RY14,PECAM1,SAMHD1,SEMA7A,SNHG16,STING1,TLR7, TSC22D3
IRF8	0.453	0.000164	CCL5,CSF1R,CTSS,GABRR2,MS4A1
TNF	0.489	0.0163	CCL5,CSF1R,CTSS,FTH1,NGFR,PCP4,PECAM1,RGS16,SNHG16, TLR7,TSC22D3
AGT	0.689	0.0378	CCL5,CTSS,ELN,PDGFRB,PECAM1,PTGDR
TCF12	0.739	0.000525	CSF1R,NGFR,SEMA7A,SPIB
TP53	0.917	0.00184	BAD,CCL5,CSF1R,FTH1,P2RY14,PDGFRB,PECAM1,PHLDA3, RGS16,TLN1,TLR7,TSC22D3,ZAP70
HMG20A	1	0.000211	CCL5,ELN,PDGFRB,RGS16
BHLHE40	1	0.0107	CTSS,P2RY14,RGS16,SAMHD1
CEBPA	1.103	0.0321	ARL4C,CSF1R,PTPN3,TSC22D3
IL2	1.235	0.0164	CCL5,CSF1R,KLRC4-KLRK1/KLRK1,PECAM1,PTPN3,RGS16
ESR1	1.39	0.0127	CTSS,ELN,KLRC4-KLRK1/KLRK1,NCKAP1,NGFR,PCP4,RGS16, SAMHD1,TSC22D3
TCF3	1.446	0.00511	CSF1R,H1-0,RGS16,SEMA7A,SPIB
CTNNB1	1.633	0.018	CCL5,HSPE1,LGR6,NDUFB2,NGFR,PECAM1,SNHG16
IKBKB	1.91	0.00844	CCL5,CSF1R,RGS16,TLN1
TNFSF11	1.949	0.00000969	CCL5,CSF1R,PHLDA3,RGS16,SIGLEC15,SPIB,TLN1
IL15	1.951	0.00797	CCL5,HLX,HSPE1,KLRC4-KLRK1/KLRK1,RGS16
IL18	1.96	0.00192	CCL5,HLX,KLRC4-KLRK1/KLRK1,SAMHD1

ITPR2	2	0.0000437	CCL5,CSF1R,SEMA7A,TLR7
ZBTB10	2.236	0.0000608	CCL5,SAMHD1,SEMA7A,STING1,ZBTB20

8.3 *RUNX1* HDR template sequence (5' → 3')

GCAGCGGCCCCTCCGGTCCCCCCTGGACAACATGATCCGCCGCCTCTCCCCGGCTGAGAGAGCTCAGGGACCCAGGCCGCACCAGGAG
AACAACCACCAGGAGTCCTACCCTCTGTCAGTGTCTCCCATGGAGAATAATCACTGCCCAGCGTCCTCCGAGTCCCACCCGAAGCCAT
CCAGCCCCCGGCAGGAGAGCACACGCGTGATCCAGCTGATGCCCAGCCCCATCATGCACCCTCTGATCCTGAACCCCCGGCACTCCGT
GGATTTCAAACAGTCCAGGCTCTCCGAGGACGGGCTGCATAGGGAAGGGAAGCCCATCAACCTCTCTCATCGGGAAGACCTGGCTTAC
ATGAACCACATCATGGTCTCTGTCTCCCCGCCTGAAGAGCACGCCATGCCCATTGGGAGAATAGCAGAATGCATACTTGGAATGAATC
CTTCTAGAGACGTCCACGATGCCAGCACGAGCCGCCGCTTCACGCCGCCTTCCACCGCGCTGAGCCCAGGCAAGATGAGCGAGGCGTT
GCCGCTGGGCGCCCCGGACGCCGGCGCTGCCCTGGCCGGCAAGCTGAGGAGCGGCGACCGCAGCATGGTGGAGGTGCTGGCCGACCAC
CCGGGCGAGCTGGTGCGCACCGACAGCCCCAACTTCCTCTGCTCCGTGCTGCCTACGCACTGGCGCTGCAACAAGACCCTGCCCATCG
CTTTCAAGGTGGTGGCCCTAGGGGATGTTCCAGATGGCACTCTGGTCACTGTGATGGCTGGCAATGATGAAAACTACTCGGCTGAGCT
GAGAAATGCTACCGCAGCCATGAAGAACCAGGTTGCAAGATTTAATGACCTCAGGTTTGTCGGTCGAAGTGGAAGAGGGAAAAGCTTC
ACTCTGACCATCACTGTCTTCACAAACCCACCGCAAGTCGCCACCTACCACAGAGCCATCAAAATCACAGTGGATGGGCCCCGAGAAC
CTCGAAGACATCGGCAGAAACTAGATGATCAGACCAAGCCCGGGAGCTTGTCCTTTTCCGAGCGGCTCAGTGAACTGGAGCAGCTGCG
GCGCACAGCCATGAGGGTCAGCCCACACCACCCAGCCCCCACGCCCAACCCTCGTGCCTCCCTGAACCACTCCACTGCCTTTAACCCT
CAGCCTCAGAGTCAGATGCAGGATACAAGGCAGATCCAACCATCCCCACCGTGGTCCTACGATCAGTCCTACCAATACCTGGGATCCA
TTGCCTCTCCTTCTGTGCACCCAGCAACGCCCATTTCACCTGGACGTGCCAGCGGCATGACAACCCTCTCTGCAGAACTTTCCAGTCG
ACTCTCAACGGCACCCGACCTGACAGCGTTCAGCGACCCGCGCCAGTTCCCCGCGCTGCCCTCCATCTCCGACCCCCGCATGCACTAT
CCAGGCGCCTTCACCTACTCCCCGACGCCGGTCACCTCGGGCATCGGCATCGGCATGTCGGCCATGGGCTCGGCCACGCGCTACCACA
CCTACCTGCCGCCGCCCTACCCCGGCTCGTCGCAAGCGCAGGGAGGCCCGTTCCAAGCCAGCTCGCCCTCCTACCACCTGTACTACGG
CGCCTCGGCCGGCTCCTACCAGTTCTCCATGGTGGGCGGCGAGCGCTCGCCGCCGCGCATCCTGCCGCCCTGCACCAACGCCTCCACC
GGCTCCGCGCTGCTCAACCCCAGCCTCCCGAACCAGAGCGACGTGGTGGAGGCCGAGGGCAGCCACAGCAACTCCCCCACCAACATGG
CGCCCTCCGCGCGCCTGGAGGAGGCCGTGTGGAGGCCCTACTGAAATAAATGGTCAATGTTATTCTTTTTCACTATTATTGCTATTTA
TGTGCctcgagttcaaaattttatcgatactagtggatCTGCGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAG
TCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACGGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTA
CTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCC
GCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCCTTCACGCGCCCGCCGCCCTACCTGAGGCCGCCATCCACGCCGGTTGA
GTCGCGTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCT
TTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCCTGCTTGCTCAACTCTACGTCTTTGTT
TCGTTTTCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCCTACtctagagctagcgccaccggccgcacctgcagcccaag
cttaccATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCAGGGCCGTACGCACCCTCGCCGCCGCGTTCG
CCGACTACCCCGCCACGCGCCACACCGTCGATCCGGACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGT
CGGGCTCGACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCG
GTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGC
ACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCT
CCCCGGAGTGGAGGCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTC
GGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGAgcggccgcaa
agctagcttcgatcCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGT
GAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTACAACAACAATTGCATTCATTTTATGTTTCA
GGTTCAGGGGGAGGTGTGGGAGGTTTTTTtgatcaGATGGAGTAGTTAATGAGCCTCAGAAATGTTAAGAAACAAATGTCCTACGTCC
AGCTTACAAGGAGAGTCACATCAGAATCAAGGCTAAGCGAAAACATTTAAAATAAAAGGTTTATGAGCATGCTAATGGCCCTGTCCTG
AAGTTTCCAGCAGCTAATTAATGACCAGACACAGCATAAAGAAGCTTTGTCTCAGATTCAGGCCTGTAATTCCATCCTGTGAGCGGTA
ACGCCATGCAGCCTTCACCTTAGGAGCGGGTAGGAGAGGAAAACAGGATTATGGTATTGGAGGCAGGTGTTGCTGGGCCATTTGGACA
GAGGAAACCACTCCCAGATTCTCTTTCATTTATTATAAGAAGCCCAAATTTGCTTACTTAAGGGAAGAAACCAGTGGAAACCAGTGGG
AAAAATATCTACACATTGTCTGTTTCCTAGGATGCTCATAATTTACATGAACTCATTCTTCAGAAGCTTAATCCTCTAACCATGTGAC
AAGGTTA

5' HA (hg38; chr12: 11869550-11869969, + strand)

RUNX1 exons 1-8

poly(A) tail (RUNX1 exon 7A / SV40 poly(A))

EF1α promoter

puromycin resistance

3' HA (hg38; chr12: 11870158-11873658, + strand)

LIST OF ABBREVIATIONS

%	percent
°C	degree Celsius
µg	microgram
µl	microliter
µM	micromolar
AML	acute myeloid leukemia
ATAC-seq	Assay for Transposase-Accessible Chromatin using sequencing
BCP-ALL	B cell precursor acute lymphoblastic leukemia
BCR	B cell receptor
bp	base pairs
BSA	bovine serum albumin
C	cytosine
CBF	core binding factor
CD	cluster of differentiation
cDNA	complementary DNA
CGI	CpG island
ChIP	chromatin immunoprecipitation
CLL	chronic lymphoblastic leukemia
CLP	common lymphoid progenitor
cm	centimeter
CML	chronic myeloid leukemia
CNA	copy number alteration
CNS	central nervous system
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DDR	DNA damage response
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
dNTP	deoxynucleotide triphosphates
E	embryonic day
E::R	ETV6::RUNX1
ECL	electrochemiluminescence
ETS	erythroblast transformation specific domain
FC	fold change
FDA	Food and Drug Administration
FDR	false discovery rate
FISH	fluorescence *in situ* hybridization
FPDMM	familial platelet disorder with associated myeloid malignancy
fw	forward
G	guanine
g	g force
GSEA	gene set enrichment analysis
H	heavy
h	hour
HA	homology arm

HAT	histone acetyl transferase
HDACi	histone deacetylase inhibitor
HDR	homology-directed repair
hiPSC	human induced pluripotent stem cell
HLH	helix-loop-helix
HPC	hematopoietic precursor cell
HRP	horse radish peroxidase
HSC	hematopoietic stem cell
Ig	immunoglobulin
IPA	Ingenuity Pathway Analysis
K	lysine
kb	kilobase pairs
kDa	kilodalton
KDM	histone lysine demethylase
KMT	histone lysine methyltransferase
L	light
LMPP	lymphoid-primed multipotent progenitor
LSC	leukemia stem cell
min	minute
ml	milliliter
mM	millimolar
MPP	multipotent progenitor
NES	normalized enrichment score
ng	nanogram
NGS	next-generation sequencing
NHEJ	non-homologous end joining
NLS	nuclear localization sequence
nm	nanometer
nM	nanomolar
NRT	no-reverse-transcriptase control
NTC	no-template control
OD	optical density
OS	overall survival
PAM	protospacer adjacent motif
PBMC	peripheral blood mononuclear cell
PBS	phosphate buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
PDX	patient-derived xenograft
PNT	pointed domain
R	arginine
RAG	recombination-activating gene
rev	reverse
RHD	Runt homology domain
rpm	rounds per minute
RT	room temperature
RT-PCR	reverse transcription polymerase chain reaction

s	second
SAHA	suberanilohydroxamic acid
SB	Sleeping Beauty transposase
scRNA-seq	single-cell RNA sequencing
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sDSS	selective drug sensitivity score
sgRNA	single guide ribonucleic acid
siCtrl	non-targeting siRNA
siH1-0	*H1-0*-targeting siRNA
SNP	single-nucleotide polymorphism
SNV	single-nucleotide variant
ssDNA	single stranded DNA
STR	short tandem repeat
TAD	transactivation domain
TDF	tumor differentiation factor
TdT	terminal deoxynucleotidyl transferase
TSA	Trichostatin A
TSS	transcription start site
UMAP	Uniform Manifold Approximation and Projection
UMI	unique molecular identifier
V	volt
w/v	weight per volume
WT	wild-type
ZIP	zero interaction potency score

LIST OF FIGURES

LIST OF TABLES

PUBLICATIONS

Publication included in this thesis:

Jepsen V.H., Hanel A., Picard D., Mehtonen J., Hasselmann R., Schliehe-Diecks J., Scharov K., Tu J.-W., Bhave R., Kameri E., Qin N., Wang H., Zhang Z., Wagener R., Blümel L., Lautwein T., Hein D., Kögler G., Remke M., Bhatia S., Heinäniemi M., Borkhardt A., Fischer U. **H1-0 is a specific mediator of the repressive ETV6::RUNX1 transcriptional landscape in preleukemia and B cell acute lymphoblastic leukemia.** Submitted.

Available as a preprint at BioRxiv (DOI: 10.1101/2024.06.28.601221)

Other publications:

Gebing P., Loizou S., Hänsch S., Schliehe-Diecks J., Spory L., Stachura P., **Jepsen V.H.**, Vogt M., Pandyra A.A., Wang H., Zhuang Z., Zimmermann J., Schrappe M., Cario G., Alsadeq A., Schewe D.M., Borkhardt A., Lenk L., Fischer U., Bhatia S. **A brain organoid/ALL co-culture model reveals the AP-1 pathway as critically associated with CNS involvement of BCP-ALL.** *Blood advances*, 2024. Online ahead of print. Impact factor: 7.4

DOI: 10.1182/bloodadvances.2023011145

Blümel L., Bernardi F., Picard D., Torrejon Diaz J., **Jepsen V.H.**, Hasselmann R., Schliehe-Diecks J., Bartl J., Qin N., Bornhauser B., Bhatia S., Marovka B., Marsaud V., Dingli F., Loew D., Stanulla M., Bourquin J.-P., Borkhardt A., Remke M., Ayrault O., Fischer U. **Proteogenomic profiling uncovers differential therapeutic vulnerabilities between *TCF3::PBX1* and *TCF3::HLF* translocated B cell acute lymphoblastic leukemia**. *Haematologica*, 2024. 109(7). Impact factor: 8.2

DOI: 10.3324/haematol.2023.283928

Rüchel N., Oldenburg M., Janssen S., Pandyra A.A., Liu W., Vasileiou E., Hein D., **Jepsen V.H.**, Fischer U., Picard D., Kögler G., Hauer J., Auer F., Beer A., Adams O., MacKenzie C., Jaeger M., Netea M.G., Borkhardt A., Gössling K.L. **Cytokine hyperresponsiveness in children with *ETV6::RUNX1*-positive acute lymphoblastic leukemia after challenge with common pathogens.** *Hemasphere*, 2023. 7(2): e835. Impact factor: 12.1

DOI: 10.1097/HS9.0000000000000835

Rüchel N.*, **Jepsen V.H.***, Hein D., Fischer U., Borkhardt A., Gössling K.L. ***In utero* development and immunosurveillance of B cell acute lymphoblastic leukemia.** *Current Treatment Options in Oncology*, 2022. 23(4):543-561. Impact factor: 3.8

DOI: 10.1007/s11864-022-00963-3

*authors contributed equally

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Prof. Arndt Borkhardt for the opportunity to work on this challenging project and his support throughout my PhD journey. Additionally, I would like to thank Prof. Sebastian Wesselborg for taking on the role as second supervisor.

Gratitude is owed to Prof. Merja Heinäniemi for the fruitful collaboration, her bioinformatic expertise on single-cell transcriptomics and the opportunity to participate in the summer school in Kuopio.

A big thank you goes out to Dr. Ute Fischer for her continuous support, her kindness and unwavering belief in my work, even during times when I was overly critical.

Thank you to everybody who contributed to my PhD project: Daniel Hein for supervision and molecular cloning know-how, Lisa and Nan for the introduction into iPSC culture, Andrea for her bioinformatic skills to analyze my single-cell RNA-seq data, Rabea for retrieving DNA methylation data, Julian and Rebecca for performing and analyzing drug screenings, Rig for his expertise on flow cytometry, and especially Daniel Picard for bioinformatic support first in Düsseldorf and then from Homburg.

Thank you to all the current and former members of my office (Dani, Nadine, Ersen, Ciro, Laura, Jasmin) for the great atmosphere, chats and coffee breaks. To everybody from the KMT lab I would like to express my deepest gratitude for making me feel welcome from day one. Thank you for the countless lunch breaks, laughs and taking my mind off the stresses of research. A special thanks goes out to Silke for bringing order to the chaos of the KMT lab.

I would like to thank my lab bestie Danielle for sharing all the ups and downs of our PhD journeys. You made even the difficult times more bearable and I am so happy that I gained you as a friend.

Thank you to Laura for gym sessions, hot chocolate, playground time with Mats and for listening to all things scientific or personal. Thank you for being the wonderful person that you are.

I want to thank Alexander for his patience with me and his unwavering support. I know that my focus on this project has often been demanding, and I am deeply grateful for your understanding and encouragement during this time.

Finally, I would like to thank my parents for always being there for me and for their unconditional love. You have taught me the importance of perseverance, and I am incredibly grateful for everything you have given me.