Department of Biomedical Sciences
University of Veterinary Medicine Vienna
St. Anna Children's Cancer Research Institute

Institute of Medical Biochemistry
(Head: Univ.-Prof. Dr. rer.nat. Florian Grebien)
Genetics of Leukemia group
(Head: Dr. Sabine Strehl)

Modelling KMT2A::MLLT3-driven leukemia in human induced pluripotent stem cells

Master Thesis

For obtaining the degree

Master of Science (MSc)

University of Veterinary Medicine Vienna

submitted by

Edwin Rzepa, BSc

Vienna, December 2023

Supervisor: Dr. Sabine Strehl (external), Dr. Klaus Fortschegger (external), Univ.-Prof. Dr. rer.nat. Florian Grebien (internal)

Reviewer:

Acknowledgements

First of all, I want to thank Dr. Sabine Strehl for giving me the opportunity to conduct my master thesis and for giving me guidance and advice throughout my time at the St. Anna Children's Cancer Research Institute and Univ.-Prof. Dr. rer.nat. Florian Grebien for the internal supervision.

In particular, I want to thank Klaus Fortschegger for the superb supervision, patience, help, support and advice he has given me during my time conducting this master thesis.

Additionally, I want to thank all the members of the Strehl Group for providing a welcoming and kind working environment.

Lastly, I want to thank my family for all the great support and my girlfriend, who has always encouraged and motivated me.

Table of content

1. Introduction

1.1 Development of the hematopoietic system

Already in the early stages of human life, the production of blood cells is essential to ensure the transport of oxygen to embryonic tissues. Therefore, the hematopoietic system is one of the first complex tissues, which develops inside the human embryo, giving rise to hematopoietic stem cells (HSCs), which differentiate, through a process called hematopoiesis, into all types of blood cells (Dzierzak & Speck, 2008). As for all human tissues, the development of the hematopoietic system follows a strict pattern of developmental sequences that are guided by regulatory networks, which ensure the correct differentiation into specific cell types. Moreover, hematopoiesis occurs at different sites during embryonic and fetal organogenesis, while adult hematopoiesis occurs predominantly in the bone marrow and to some extent in the liver, spleen, and lymph nodes. Therefore, one distinguishes between embryonic/fetal and adult hematopoiesis.

1.1.1 Embryonic and fetal development of the mammalian hematopoietic system

Embryonic or fetal hematopoiesis is divided into three consecutive waves, which were first described in mice. The onset of mammalian hematopoietic development is marked by the first wave, also termed primitive hematopoiesis, in which transient HSCs of mesodermal origin called hemangioblasts residing in the yolk sac produce primitive erythroid progenitors, macrophages and megakaryocytes required for oxygenation, tissue remodeling and vascular maintenance (Dzierzak & Speck, 2008). Shortly after the formation of primitive blood cells the second wave called pre-definitive hematopoiesis takes place, which is marked by the emergence of myeloid and lymphoid progenitors with adult type features (Gao et al., 2018). This wave is transiently replaced by the third wave, termed definitive hematopoiesis. It takes place inside the aorta-gonad-mesonephros (AGM), an embryonic tissue of mesodermal origin, which contains the early dorsal aorta containing the hemogenic endothelium, the genital ridge and the mesonephros (Pietilä & Vainio, 2005). Inside the hemogenic endothelium, cells of vascular origin give rise to definite, multipotent HSCs via a process called endothelial-to-hematopoietic transition (Bertrand et al., 2010; Gritz & Hirschi, 2016). These engraftable HSCs then colonize the fetal liver, thymus, spleen, and bone marrow, the latter becoming the

predominant site of hematopoiesis after birth (**Figure 1**) (Hirakawa et al., 2022; Orkin & Zon, 2008).

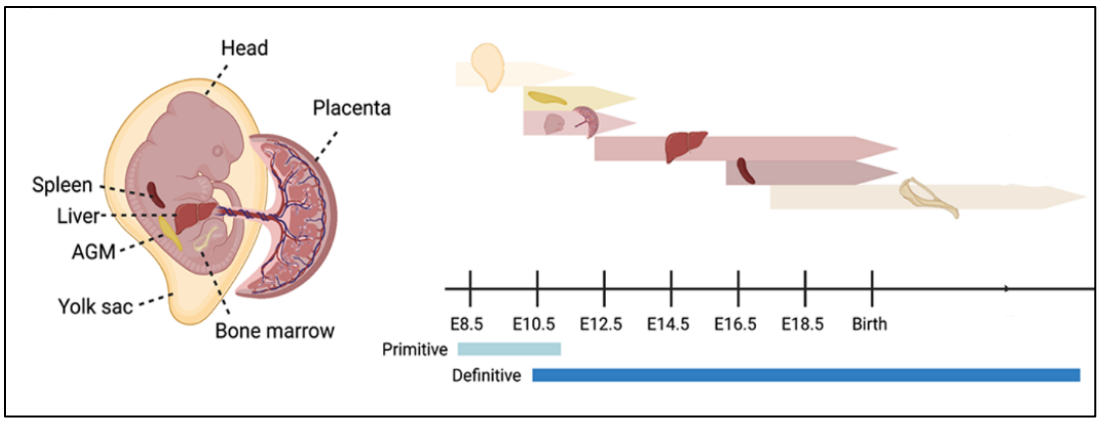


Figure 1: Timeline of the hematopoietic development (modified from Hirakawa & Ding, 2022). Depicted are the different sites of hematopoiesis during fetal and embryonic development, as well as the timeline of primitive and definitive hematopoiesis in days of embryonic development (E). Aorta gonad mesonephros (AGM).

1.1.2 Adult hematopoietic differentiation

After birth, bone-marrow-resident HSCs either remain in a quiescent state or undergo differentiation. To better understand this differentiation process, various models have been developed over the years. One of the most prominent but outdated models is the classical hierarchical model, which resembles an inverted tree. In this model, the root or stem contains quiescent, pluripotent HSCs, mainly long-term HSCs (LT-HSCs) transitioning into short-term HSCs (ST-HSCs), which then differentiate into multipotent progenitors (MPPs). These MPPs further differentiate into either common myeloid progenitors (CMPs) or common lymphoid progenitors (CLPs), the former giving rise to the myeloid lineage, mainly granulocyte-macrophage progenitors (GMPs) and megakaryocyte-erythrocyte progenitors (MEPs), the latter to progenitors of the lymphoid lineage. Therefore, GMPs give rise to granulocytes and monocytes, MEPs to megakaryocytes and erythrocytes, and CLPs to B- and T-cells as well as to natural killer (NK) cells (**Figure 2**) (Seita & Weissman, 2010; Watcham et al., 2019).

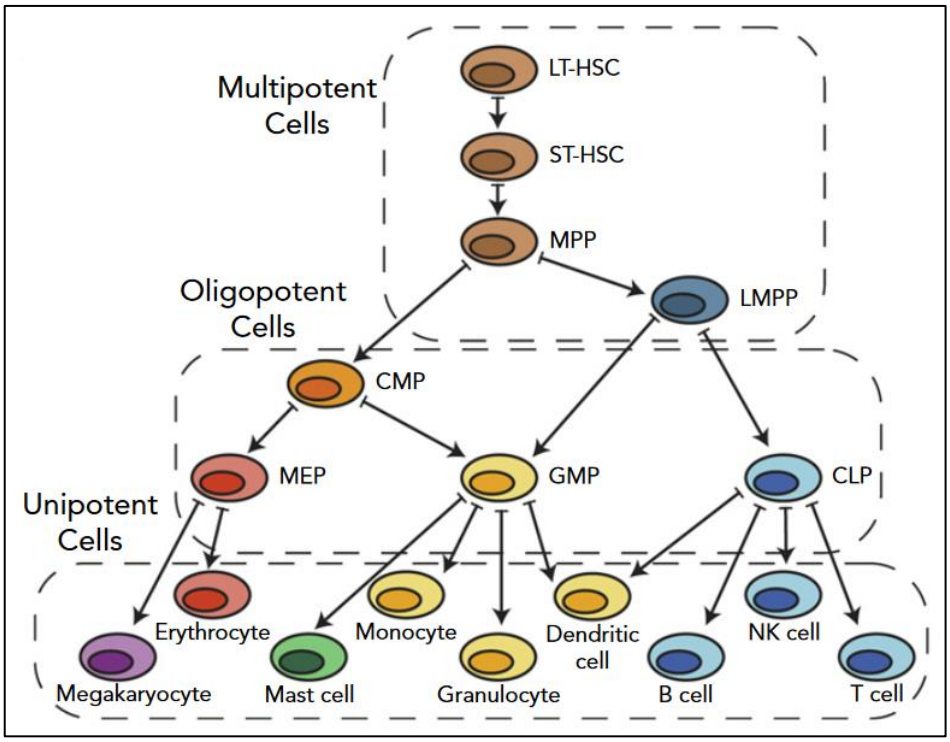


Figure 2: The hematopoietic tree (modified from Watcham et al., 2019). Long-term hematopoietic stem cell (LT-HSC), short-term hematopoietic stem cell (ST-HSC), multipotent progenitor (MPP), lymphoid primed multipotent progenitor (LMPP), common lymphoid progenitor (CLP), natural killer cell (NK-cell), common myeloid progenitor (CMP), megakaryocyte-erythroid progenitor (MEP), granulocyte-monocyte progenitor (GMP).

One important caveat of the classical model is that it is based on data from bulk cell analysis of surface marker expression using flow cytometry (Liggett & Sankaran, 2020). Consequently, this model assumes that each differentiation state comprises a homogenous cell population, which oversimplifies the hematopoietic hierarchy. Moreover, it proposes that differentiation occurs in a stepwise manner, in which HSCs transition through discrete progenitor states, thereby acquiring a lineage bias (Cheng et al., 2020). These assumptions have been challenged by recent results from single-cell and barcoding technologies such as single-cell RNA-sequencing (scRNA-seq) (Pei et al., 2017; Ranzoni et al., 2021). It became apparent that differentiation occurs continuously, and that stem cells gradually acquire a lineage bias along multiple directions without transitioning through discrete progenitor stages (Cheng et al., 2020). Moreover, new subpopulations both in the HSC and progenitor compartments have been identified. Extensive analysis of the multipotent HSC/MPP pool demonstrated that the border between HSCs and progenitor cells is much more diffuse than previously thought and is composed of different subpopulations, which produce diverse myeloid and lymphoid outputs (Laurenti & Göttgens, 2018). In this context, it has been shown that the MPP pool of mice consists of at least four subpopulations, namely MPP1, MPP2, MPP3, and MPP4, each having

its own lineage-biased potential (Pietras et al., 2015). These findings also indicate that lineage priming can already occur in the multipotent stem cell state. Moreover, several subpopulations within the myeloid progenitor compartment and at least 18 different subtypes within the CMP and MEP pools have been identified by scRNA-seq (Watcham et al., 2019).

Although the classical model proved useful to understand hematopoiesis in a simplified way, it is nowadays outdated, and the newest insights suggest that the hematopoietic tree shall be seen as a hierarchical, organized collection of heterogenous populations that gradually transition from one to another.

1.1.3 Regulatory mechanisms of hematopoietic differentiation

Even though hematopoietic differentiation gives rise to various cell populations with distinct functions, the mechanisms that control this process are conserved. A variety of external cues such as physical tension, soluble factors and cell-cell contacts are provided by the so-called hematopoietic niche, a tightly controlled local environment inside the bone marrow consisting of bone marrow resident cells such as osteoblasts, adipocytes, macrophages, nerve, endothelial, and perivascular cells, which act on HSC quiescence, proliferation, and differentiation (Gao et al., 2018). One of the most important regulators of HSC and progenitor differentiation as well as survival are cytokines, which are produced in the bone marrow and other organs such as the kidneys or lungs (Metcalf, 2008). Over the years, two models have been established to describe the role of cytokines in hematopoietic differentiation. The instructive model proposes that cytokines direct multipotent cells into a specific lineage, while the permissive model suggests that lineage commitment is determined intrinsically, and cytokines only provide permissive growth and survival signals (Robb, 2007). Nevertheless, their importance for hematopoiesis is undisputed, and the mechanism behind cytokine function is conserved. Cytokines are secreted and bind to receptors on HSCs and progenitor cells, which then activate signaling pathways inside the cell, subsequently leading to proliferation and/or differentiation (Jafari et al., 2019; Staerk & Constantinescu, 2012).

The hematopoietic cytokines are comprised of the colony stimulating factors (CSFs), interleukins (ILs), thrombopoietin (TPO), erythropoietin (EPO), FMS-like tyrosine kinase 3 ligand (FLT3L), and stem cell factor (SCF), the latter is mainly required for HSCs maintenance. Colony formation assays have revealed that there a four CSFs, granulocyte-macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), and macrophage colony stimulating factor (M-CSF) as well as the multipotential colony-stimulating

factor also known as interleukin-3 (IL-3) (Sachs, 1996). Together these CSFs act in combination with other cytokines on proliferation, lineage-commitment, cell-survival, maturation, and function of myeloid progenitor cells to give rise to granulocytes and macrophages. Other examples that highlight the importance of cytokines in hematopoiesis are the production of erythrocytes upon stimulation by renally secreted EPO, the role of interleukin 7 (IL-7) in lymphoid cell development and the role of TPO in megakaryopoiesis (Bunn, 2013; Chen et al., 2021; Geddis, 2010). Many hematopoietic cytokine receptors belong to a group termed receptor tyrosine kinases, which mediate the cytokine-induced signal into the cell and activate signaling pathways such as the JAK-STAT and PI3K-AKT pathways (Jafari et al., 2019; Staerk & Constantinescu, 2012).

As a result, large gene regulatory networks, consisting of a plethora of DNA-binding proteins called transcription factors (TFs) and *cis* regulatory elements in the genome such as promotors, enhancers and silencers, are activated. Depending on the interplay of different TFs and epigenetic modifications, individual genes or entire gene clusters are transcriptionally up- or downregulated dependent on cofactor interaction. As an intrinsic regulatory component, TFs are essential to repress or promote the differentiation of hematopoietic cells into a specific lineage. This is generally accomplished by transcriptional activation of lineage-specific genes and simultaneous transcriptional repression of lineage-inappropriate genes and those regulating self-renewal.

It has been demonstrated that the myeloid TF CCAAT/enhancer binding protein alpha (C/EBPα) is essential for myeloid gene expression during cell fate decision by binding to promotors of myeloid-related genes such as cytokine receptors *CSF3R* and *IL6R*, or other TFs including *CEBPE*, *GFI1* and *KLF5* (Avellino & Delwel, 2017). Other examples are the runt-related (RUNX) family of pioneer TFs in the development of the murine hematopoietic system as well as in embryonic, fetal, and adult hematopoiesis, where they change the epigenetic landscape to expose new *cis* regulatory elements and promote expression of lineage-specific genes such as *Il7r* required for T-lymphocyte differentiation (De Bruijn & Dzierzak, 2017; Voon et al., 2015). The family *of* GATA-binding (GATA) transcription factors is required for erythroid and megakaryocytic differentiation, interacting with other TFs to, on the one hand, transcriptionally upregulate erythroid genes such as *Kit*, *Lyl1*, and *Spi1* and, on the other hand, repress genes that drive self-renewal (Doré & Crispino, 2011; Katsumura et al., 2017). Transcription factors also play an indispensable role in governing B-cell development, as

exemplified by the pivotal contributions of PAX5, EBF1, TCF3, and IKZF1 in orchestrating lineage commitment and immunoglobulin rearrangement (Somasundaram et al., 2015).

It is evident that TFs rely on epigenetic modifiers to fully unfold their regulatory function in the chromatin context. During hematopoiesis chromatin remodeling is an important regulatory mechanism, which controls DNA accessibility for TFs, thereby influencing the transcription of lineage-specific genes or exposing *cis* regulatory elements, as well as being responsible for epigenetic lineage priming of HSCs (Rodrigues et al., 2021; Yu et al., 2016). Moreover, TFs need to recruit and interact with chromatin remodelers to increase or decrease the transcription of their target genes (Gore & Weinstein, 2016). Chromatin is a highly organized DNA-protein complex: negatively charged DNA is tightly packed around nucleosomes, which are multiprotein complexes consisting of four positively charged histone subunits H2A, H2B, H3, and H4, each containing long amino acid tails that can be post-translationally modified by histone-modifying enzymes. These reversible modifications, also known as histone marks, include acetylation, methylation, ubiquitination, sumoylation, and phosphorylation, which can decrease or increase DNA interaction with nucleosomes. Additionally, DNA can be directly modified by cytosine methylation to regulate transcriptional activity without altering the DNA sequence (Gore et al., 2016). Moreover, these epigenetic marks are recognized by a set of chromatin-associated proteins known as epigenetic readers. These readers selectively bind to histone marks or other forms of post-translational modifications and orchestrate the formation of large nuclear protein complexes comprised of a variety of DNA associated proteins, that modulate gene expression dependent on the epigenetic context (Andrews et al., 2016). Therefore, histone methylations either increase or decrease transcription depending on the degree of methylation and the specific methylated residues. Active methylation marks include H3K4me2/3 and H3K9me0, while repressive methylation marks include H3K27me2/3 and H3K9me2/3 (Goyama & Kitamura, 2017). Histone methylations are added to lysine residues by the enzyme family of histone methyltransferases (KMTs) such as SETD2 and KMT2A (Husmann & Gozani, 2019). It has been determined that KMTs regulate the expression of hematopoietic genes important for HSC self-renewal and lineage priming, and mutations in genes encoding these enzymes may lead to altered differentiation and the development of hematopoietic malignancies (Goyama & Kitamura, 2017).

1.2 Pediatric acute myeloid leukemia

Pediatric acute myeloid leukemia (AML) is a hematologic malignancy that is characterized by the abnormal growth and aberrant differentiation of myeloid progenitors driven by perturbations of gene regulatory networks caused by genetic alterations, mainly gene rearrangements and point mutations. As a result, immature myeloblasts accumulate in the bone marrow and extramedullary tissues, impeding the terminal differentiation of normal blood cells such as erythrocytes, megakaryocytes, and granulocytes, leading to symptoms like anemia, thrombocytopenia and secondary immunodeficiencies.

According to the leukemia stem cell (LSC) theory, these malignant blasts originate from early hematopoietic progenitors that evolved into LSCs, a subset of mutated and preleukemic HSCs that can propagate the disease upon xenotransplantation (Marchand & Pinho, 2021). These LSCs continually replenish the bulk of leukemic cells, in part resembling physiological hematopoietic differentiation (Marchand & Pinho, 2021). Twin studies have shown that pediatric AML can be initiated already *in utero*, giving rise to preleukemic cells, in contrast to the development of adult AML, in which the disease arises as a result of over time accumulated mutations, a process called clonal hematopoiesis of indeterminate potential (Marcotte et al., 2021). Moreover, adult AML often develops from prior non-malignant hematological diseases, while almost all pediatric AML cases are *de novo.*

Pediatric AML is a rare disease, affecting seven out of one million children below the age of 14 years. It represents 15-20 % of all pediatric leukemias while up to 80 % comprise various subtypes of B-cell and T-cell acute lymphoblastic leukemia (ALL). AML is a very heterogenous disease, which can affect all myeloid lineages. Therefore, two classification systems have been established. The French-American-British (FAB) classification of AML has already been developed in 1970 and is merely based on morphological assessment of the affected myeloid cell type or lineage and their level of maturity (Ladines-Castro et al., 2016) (**Figure 3**). The AML FAB classification distinguishes nine subtypes:

- M0: Undifferentiated acute myeloblastic leukemia
- M1: Acute myeloblastic leukemia with minimal maturation
- M2: Acute myeloblastic leukemia with maturation
- M3: Acute promyelocytic leukemia
- M4: Acute myelomonocytic leukemia
- M4eo: Acute myelomonocytic leukemia with eosinophilia

- M5: Acute monocytic leukemia
- M6: Acute erythroid leukemia
- M7: Acute megakaryoblastic leukemia (AMKL)

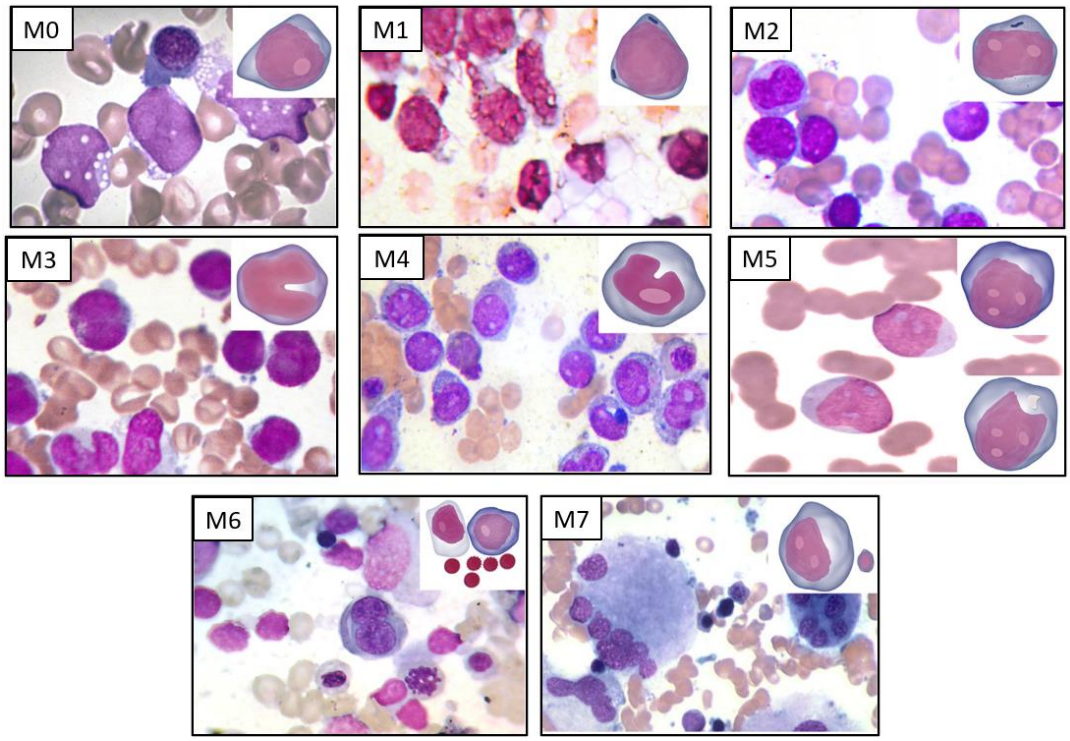


Figure 3: Morphology of different FAB-AML subtypes (modified from Ladines-Castro et al., 2016).

The second classification system is the World Health Organization (WHO) classification, which is continuously updated according to novel findings. This system is based on clinical, morphologic, immunophenotypic and genetic features, many of which are known to affect prognosis (Alaggio et al., 2022). Although the WHO classification delineates a much higher number of AML subtypes, the FAB morphology is still used as a primary diagnostic tool and may guide further subclassification. Since age is a contributing prognostic factor, the WHO classifications specified infant AML as an additional subtype. Infant AML is defined by the occurrence of AML below the age of one and has unique biological and clinical properties, including a high frequency of specific chromosomal aberrations often associated with AMKL (FAB M7), such as histone-lysine N-methyltransferase 2A (*KMT2A*) rearrangements (*KMT2Ar*) making up to 40-60 % of cases followed by core binding factor rearrangements (CBF) being found in 5 % of cases (Calvo et al., 2021).

AML is usually detected by abnormal white blood cell counts and the presence of myeloblasts in the blood. Subsequently, bone marrow samples are used for morphological assessments and immunophenotyping, as well as genetic analyses by karyotyping, fluorescence *in situ* hybridization (FISH), reverse transcription-polymerase chain reaction (RT-PCR), and mutation analysis. Nowadays, optimally targeted next-generation sequencing (NGS) approaches, which allow for the simultaneous detection of gene rearrangements and mutations, are employed. Generally, AML is diagnosed if myeloblasts comprise more than 20 % of total bone marrow cells, however, some AML types with defining genetic abnormalities no longer require this criterion according to the new WHO classification (Alaggio et al., 2022). Importantly, the outcome is strongly dependent on the underlying cytogenetic and molecular abnormalities. Therefore, their detection is essential for risk stratification and optimization of treatment conditions.

Owing to risk stratification and therapy optimization over the last decades, survival rates have increased from merely 20 % to up to 75 % provided that comprehensive diagnostics, adjusted and intensive therapy, as well as effective supportive care, are applied (Reinhardt et al., 2022). However, treatment has not changed much over the last years, and chemotherapy is still the standard treatment for pediatric AML. Induction therapy comprises three days of anthracycline and seven to ten days of cytarabine administrations, reducing bone marrow blast cell counts below 5 %, which is considered complete morphological remission, in 80 % of children and adolescents (Creutzig et al., 2012). Subsequent consolidation therapy consists of additional rounds of chemotherapy using high dosages of non-cross-resistant chemotherapeutics to maintain remission and avoid resistance. Alternatively, allogenic, or autologous hematopoietic stem cell transplantation is performed for patients with a high risk of relapse (Kassim & Savani, 2017). Nevertheless, there is still an alarming inequality of survival rates between standard treatment protocols, ranging from 50-80 % (Reinhardt et al., 2022; Rubnitz & Inaba, 2012). Moreover, standard chemotherapy protocols are plagued by high rates of relapse, reaching up to 30% in pediatric cases after first remission, and treatment-related side effects such as anthracycline-induced cardiomyopathies (Reinhardt et al., 2022).

To improve survival rates and find less toxic treatment options, intensive research was conducted to develop targeted therapies such as tyrosine kinase inhibitors (TKIs), antibodies and cellular immunotherapy to specifically target cancer cells. For example, the TKIs Midostaurin and Sorafenib, which inhibit aberrant signaling caused by an internal tandem duplication (ITD) of the FMS-like tyrosine kinase 3 (FLT3), in combination with chemotherapy,

improved treatment outcomes in clinical phase 3 trials (Pommert & Tarlock, 2022). Targeting of the myeloid surface marker CD33 with the antibody-drug conjugate gemtuzumab ozogamicin showed single-agent efficacy and improved patient outcome (Pommert & Tarlock, 2022). Encouraged by the success of immunotherapies in ALL, new immunotherapeutic options are now also already available or under development for AML, including CAR-T and NK-cell therapies (Koedijk et al., 2021). However, many of these therapies have been developed for adult AML, and there is a lack of therapies that target drivers specific to pediatric AML. Moreover, the use and development of such therapies require detailed knowledge and understanding of the underlying genetic abnormalities and molecular disease mechanisms.

1.3 Common genetic alterations in pediatric acute myeloid leukemia

Genetic alterations found in AML are generally divided into two subgroups, and usually both are required for the development of AML. Type I abnormalities are responsible for the abnormal growth of myeloblasts and comprise somatic, activating mutations in signaling pathways, which govern self-renewal, survival, and proliferation (de Rooij et al., 2015). Type II abnormalities impair differentiation and comprise mainly rearrangements of genes encoding hematopoietic transcription factors, resulting in the expression of oncogenic fusion proteins (de Rooij et al., 2015). Notably, fusion genes are found in 80 % of pediatric AML patients, while the somatic mutational burden is significantly higher in adults (Aung et al., 2021). Therefore, pediatric and adult AML are considered as distinct biological entities (Aung et al., 2021).

The somatic mutational burden is very low in pediatric AML, and point mutations are only found in 5-15 % of the cases. The mutational landscape, frequency, and co-mutations in pediatric AML differ significantly from those in adult AML. For example, in children, genes involved in signaling processes such as the *RAS* family (e.g. *NRAS*, *KRAS*), *KIT* and *CBL* are often mutated, while in adults mutations are commonly found in epigenetic regulators including *TET2*, *DNMT3A*, *ASXL1*, *NPM1* or *IDH2* (Aung et al., 2021; Bolouri et al., 2018) (**Figure 4**).

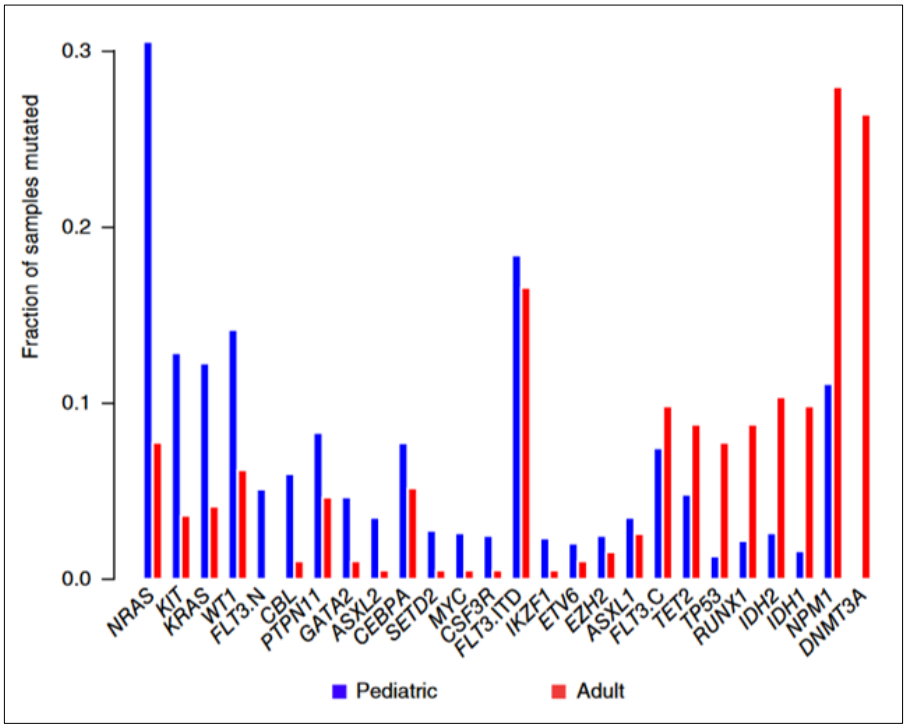


Figure 4: Frequency of driver mutations in pediatric and adult AML (modified from Bolouri et al., 2018). Age-dependent differences in the prevalence of mutations. The y-axis depicts the fraction of mutated samples (n=684) of the corresponding gene on the x-axis.

In addition, pediatric AML is characterized by a high frequency of cytogenetically visible genomic aberrations, which are found in 80 % of cases but only in 55 % of adult AML (Shiba, 2023). These cytogenetic aberrations are mainly nonrandom, balanced translocations or inversions and are considered important diagnostic and prognostic factors. Chromosomal rearrangements account for the majority of genetic entities in the category of “AML with defining genetic abnormalities” in the current WHO classification (Alaggio et al., 2022). On the one hand, chromosomal rearrangements can lead to the relocation of enhancer regions into the proximity of proto-oncogenes, resulting in their increased expression. For example, the upregulation of the stem-cell regulator *MECOM* in AML is induced by the repositioning of a distal *GATA2* enhancer by chromosome 3 inversions (Gröschel et al., 2014). On the other hand, chromosomal rearrangements may result in the fusion of two different genes, leading to the expression of chimeric fusion proteins, some of which are exclusively found in pediatric patients. In fact, many hematopoietic TFs have been discovered through chromosomal translocations and the subsequent characterization of the resulting novel transcripts (Sive & Göttgens, 2014). By fusion of functional protein domains to each other, these fusion proteins acquire new oncogenic properties and interfere with signaling pathways, transcriptionally deregulate genes or alter the epigenetic landscapes (Alcalay et al., 2003). An overview of recurrent genetic abnormalities in pediatric AML is depicted in **Figure 5** (Quessada et al., 2021).

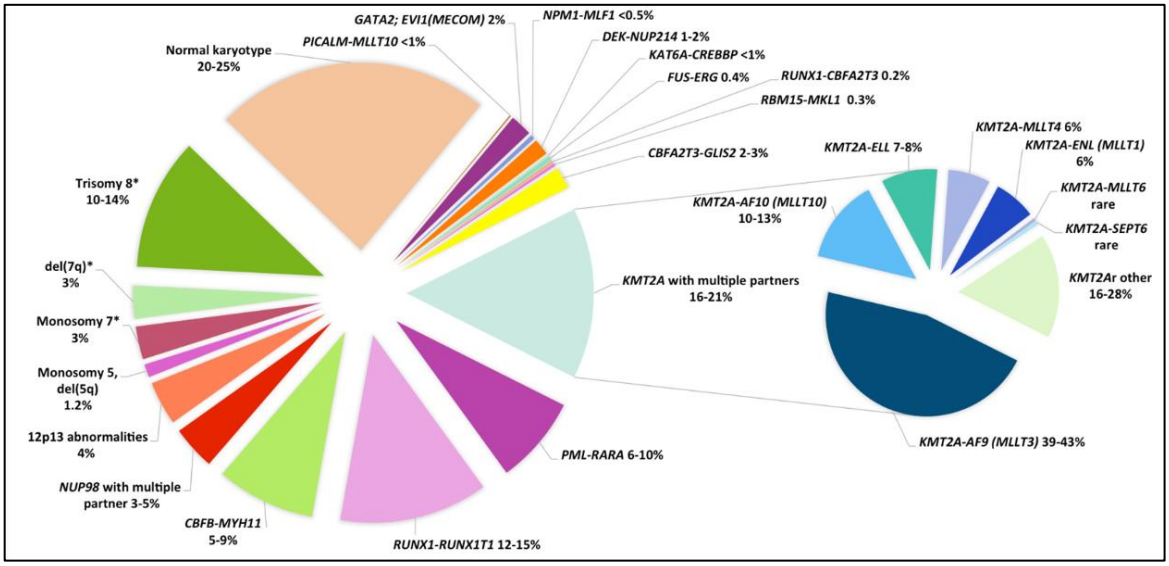


Figure 5: Frequencies of cytogenetic aberrations in pediatric AML (modified from Quessada et al., 2021). Left pie chart shows distribution of cytogenetic abnormalities found in pediatric AML. Right pie chart shows distribution of *KMT2A* fusion partners.

The largest pediatric subgroup is represented by CBF leukemias, which make up 25 % of all pediatric AML cases, with a median age of 8-9 years. It refers to AML that exhibits either a t(8;21)(q22;q22) resulting in the *RUNX1::RUNX1T1* fusion or an inv(16)(p13q22)/*CBFB::MYH11* rearrangement accounting for 12-15 % and 7-11 % of the cases, respectively. The expression of CBF fusion proteins alone does not cause full-blown leukemia and is therefore always accompanied by secondary mutations or additional cytogenetic abnormalities. AML with *RUNX1::RUNX1T1* is mainly associated with the FAB subtype M2, and *CBFB::MYH11* with M4eo. Overall, CBF leukemia is associated with a good prognosis and has an overall survival rate of over 80 % (Quessada et al., 2021).

The *PML::RARA* fusion is generated by a t(15;17)(q24;q21) and is found in 5-10 % of pediatric AML cases. It is causative for the development of acute promyelocytic leukemia (APL, FAB M3), which is characterized by a differentiation block in the granulocytic lineage leading to the expansion of promyelocytes. APL occurs at a median age of 12 years and is a very aggressive form of AML. However, treatment with all-trans retinoic acid (ATRA) and arsenic trioxide (ATO), which induce terminal differentiation of promyeloblast has drastically improved overall survival rates (Quessada et al., 2021; Zhang et al., 2018).

Chromosomal rearrangements involving the *KMT2A* gene are commonly found in pediatric AML, accounting for about 20 % of the cases while being rare in adults with a frequency of only 2 % (Aung et al., 2021; Quessada et al., 2021). *KMT2Ar* are also found in B-ALL, T-ALL, myelodysplastic syndromes and secondary AML. Infant AML shows the highest frequency, with 40-60 % of patients harboring *KMT2Ar*. *KMT2Ar* AML mainly shows a monocytic phenotype and is therefore associated with FAB subtypes M4 and M5 in roughly 70 % of cases, but *KMT2Ar* are also found in AMKL (FAB M7) (Quessada et al., 2021).

KMT2A is located on chromosome 11 (11q23), and the encoded protein facilitates the transfer of N-methyl groups onto lysine 4 histone 3 (H3K4). The large nuclear protein initially weighs around 500 kda and consists of an N-terminal menin binding domain, a domain that is essential for its catalytic function, followed by three AT hooks and a CXXC domain, which are required for DNA binding. Histone tail recognition and binding are established by a plant homology domain (PHD) and a bromodomain (BRD). KMT2A is cleaved by Taspase-1 into an N-terminal (~320 kda) and a C-terminal part (~180 kda) (Krivtsov et al., 2017). Both parts still interact via the FYRN and FYRC domains, allowing the protein to assemble into large multiprotein

complexes. The C-terminal Su(var)3–9, Ezh2, Trithorax (SET) domain catalyzes the mono-, di or trimethylation of H3K4 (Krivtsov et al., 2017) (**Figure 6**).

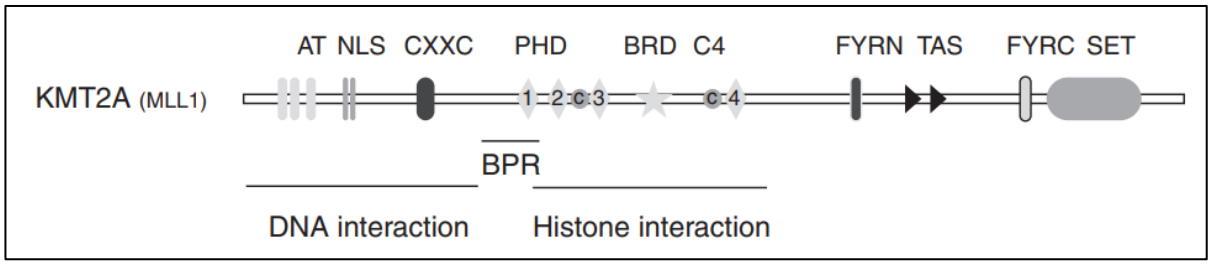


Figure 6: Protein domains of KMT2A (modified from Krivtsov et al., 2017). AT hook (AT), nuclear localization signal (NLS), CXXC motif (CXXC), breakpoint region (BPR), plant homology domain (PHD), bromodomain (BRD), zinc finger (C), Taspase-1 cute site (TAS), (Su(var)3–9, Ezh2, Trithorax methyltransferase (SET).

KMT2A is an important transcriptional regulator, which is enriched at promotors and enhancers. Via its enzymatic activity and interaction with other nuclear proteins, KMT2A facilitates transcription of *HOX* genes, including *HOXA9*, *HOXA7,* and *HOXA8* (Winters & Bernt, 2017). Knock-out studies revealed that KMT2A is required for proper axial skeletal formation and specification of functional HSCs during embryogenesis, as well as the maintenance of adult HSCs (Antunes & Ottersbach, 2020).

The translocation breakpoints in the *KMT2A* gene cluster between exons 8 and 14 lead to the fusion of the 5' N-terminal part of KMT2A to the 3' C-terminal part of the fusion partner. The majority of KMT2A fusion partners are part of the super elongation complex (SEC), which is involved in transcriptional regulation by modulating mRNA elongation via interaction with RNA-Polymerase II (Pol II) (Luo et al., 2012). Thereby, Pol II is directly recruited to KMT2A-defined target genes, which induces global changes in gene expression and aberrant epigenetic signatures (Mercher & Schwaller, 2019) (**Figure 7**). As a result, KMT2A fusions enhance the expression of *HOX* genes such as *HOXA9,* which is overexpressed up to 8-fold in *KMT2Ar* AML compared to healthy controls (Collins & Hess, 2016). HOXA9 is an important hematopoietic transcription factor involved in adult hematopoiesis, and its overexpression leads to upregulation of proliferative genes and impaired differentiation, contributing to the subsequent transformation of the affected cells (Collins & Hess, 2016).

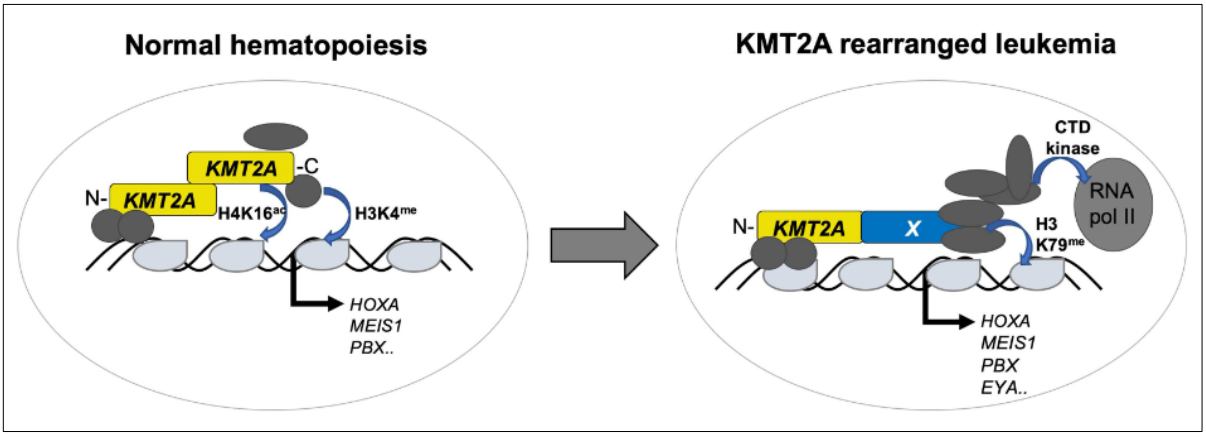


Figure 7: KMT2A function in normal hematopoiesis and KMT2A rearranged leukemia (modified from Mercher & Schwaller, 2019). In normal hematopoiesis (left), the N-terminal and C-terminal part of KMT2A interact to regulate expression of self-renewal genes via H4K16 acetylation and H3K4 methylation. In KMT2A rearranged leukemia (right), KMT2A fusion proteins recruit the transcriptional machinery to aberrantly drive expression of target genes, resulting in leukemogenesis.

The prognosis and outcome of patients with *KMT2A*r AML are highly dependent on the fusion partner, additional cytogenetic aberrations, and secondary mutations, although *KMT2Ar* AML shows a very low mutational burden compared to other genetic subtypes (Quessada et al., 2021; Yuen et al., 2023). To date, more than 100 *KMT2A* translocation partners are known ,and the most frequent ones in pediatric AML include *MLLT3* [t(9;11)(p22;q23)], *MLLT10* [t(10;11)(p12;q23)], *ELL* [t(11;19)(q23;p13)], and *AFDN* [t(6;11)(q27;q23)], of which *KMT2A::MLLT3* is the most frequent translocation being found in about 40 % of cases (Meyer et al., 2018, 2023; Quessada et al., 2021) (**Figure 5**). Other more frequent *KMT2A* fusion partners include *MLLT1* [t(11;19)(q23;p13.3)], *EPS15* [t(1;11)(p32;q23)], and *AFF1* [t(4;11)(q21;q23)] (Meyer et al., 2018, 2023; Quessada et al., 2021).

Like for other disease entities, the outcome of *KMT2Ar* AML is affected by the presence of coinciding mutations, which are mainly found in genes involved in signaling such as *NRAS*, *FLT3* and *PTPN11*, with *NRAS* being the most frequently mutated gene (Bolouri et al., 2018; Yuen et al., 2023). *NRAS*-mutated pediatric AML is associated with a poor prognosis, although its prognostic impact in *KMT2Ar* AML is still debated (Yuen et al., 2023). Like all members of the *RAS* gene family, *NRAS* codes for a guanine nucleotide-binding factor, which is a central signaling protein in many cellular pathways that govern proliferation, survival, and apoptosis. Its activity is regulated by the hydrolysis of GTP to GDP by GTPase-activating proteins and the replacement of GDP to GTP via guanine nucleotide exchange factors (Karnoub & Weinberg, 2008; Mercher & Schwaller, 2019). *NRAS* mutations are found in 20 % of human cancers and frequently cluster as single point mutations in exon 2 around glycine codons 12

or 13, for example, yielding the pathogenic NRAS G12D protein (Liu et al., 2019; Yuen et al., 2023). This mutation impairs the GTPase activity of NRAS, leading to the accumulation of GTP-bound active NRAS inside the cell, which results in aberrant signaling and fuels uncontrolled proliferation.

1.4 KMT2A::MLLT3 leukemia

The *KMT2A::MLLT3* fusion is predominantly found in pediatric AML, with the highest incidence in infants (Meyer et al., 2018, 2023; Quessada et al., 2021) and is associated with an intermediate prognosis and an overall favorable outcome compared to other *KMT2A* fusion partners (Yuen et al., 2023) (**Figure 8**).

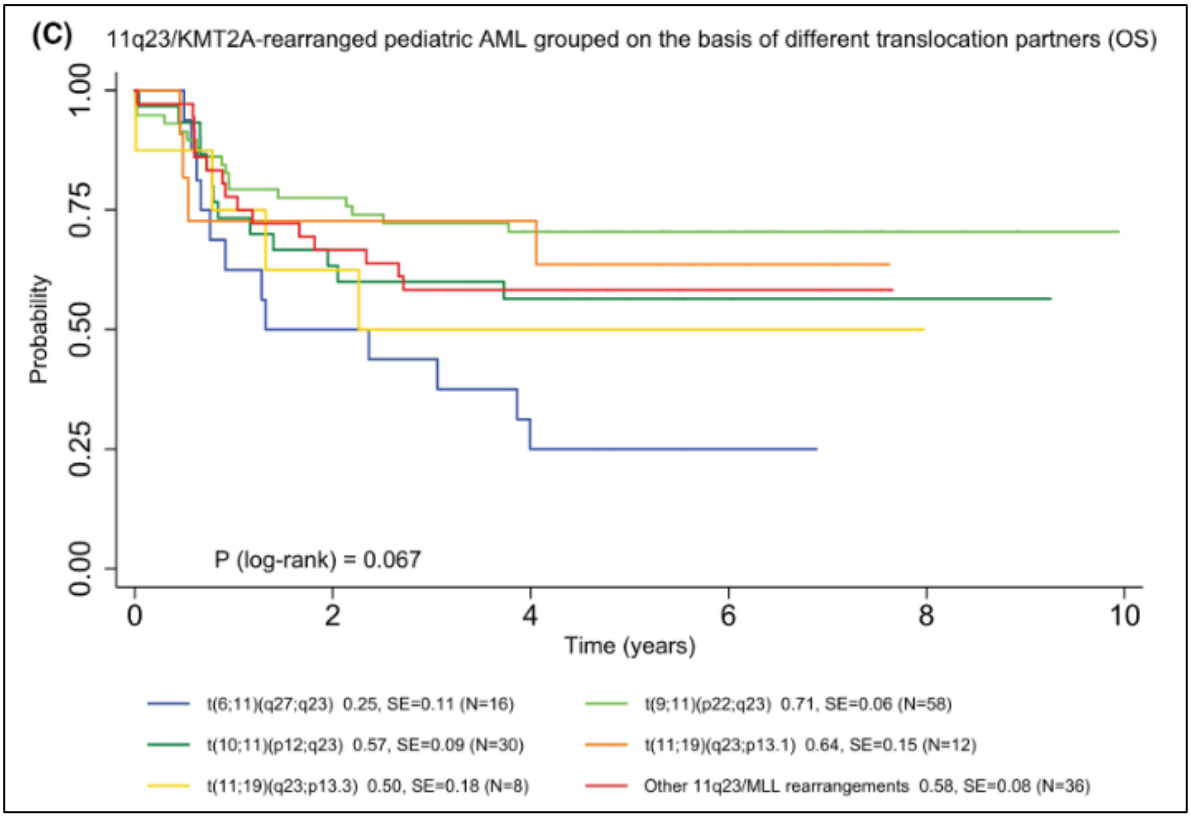


Figure 8: Probability of overall survival of patients with different *KMT2A* translocation partners (modified from Yuen et al., 2023). *KMT2A::AFDN* t(6;11)(q27;q23), *KMT2A::MLLT10* t(10;11)(p12;q23), *KMT2A::ELL* t(11;19)(q23;p13), *KMT2A::MLLT3* t(9;11)(p22;q23), overall survival (OS), standard error (SE).

The Mixed-Lineage Leukemia Translocated To 3 protein (MLLT3) is an important epigenetic reader as part of the SEC. MLLT3 belongs to the YEATS family of proteins, which is characterized by the presence of a YEATS domain that can bind to acetylated and crotonylated lysines (Kabra & Bushweller, 2022). MLLT3 is a critical component of the SEC and regulates the self-renewal of HSCs in hematopoiesis by preserving HSC gene expression via DOT1L recruitment (Calvanese et al., 2019). In KMT2A::MLLT3 fusions, the N-terminal part of KMT2A

is fused to the C-terminal part of MLLT3, which results in the recruitment of the SEC to KMT2A target genes (**Figure 7** and **Figure 9**). As a consequence, paused Pol II is prematurely released, skipping the transcription elongation checkpoint and resulting in transcriptional upregulation of target genes such as *HOXA9* and *HOXA10* (Luo et al., 2012). While this process is considered as the major transforming event of many different *KMT2A* fusions, additional oncogenic mechanisms unique to specific fusions, which could explain the heterogeneity in patient outcomes, remain widely unknown.

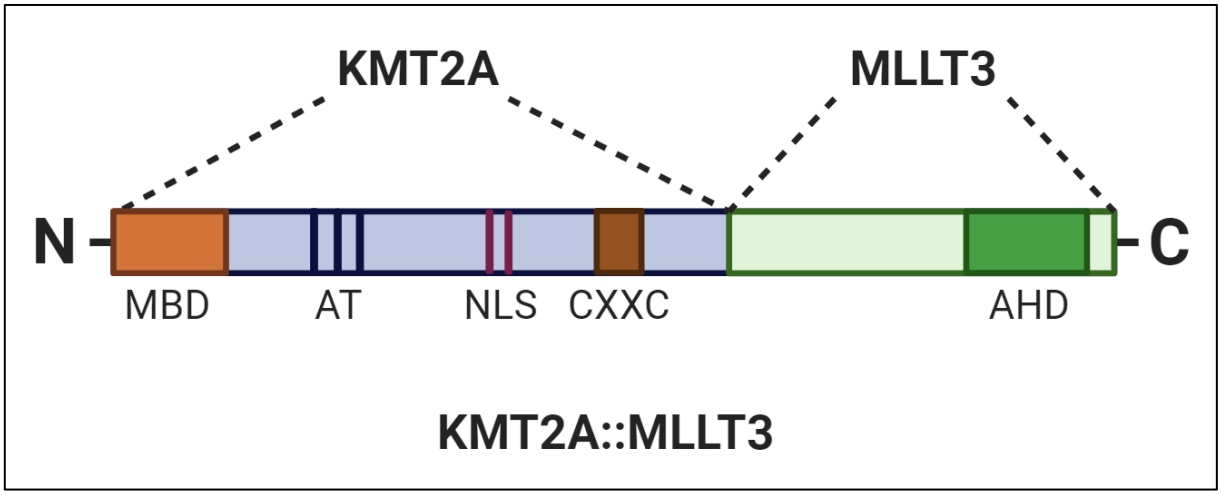


Figure 9: Structure of the KMT2A::MLLT3 fusion protein. Menin binding domain (MBD), AT hook (AT), nuclear localization signal (NLS), CXXC motif (CXXC), ANC1 homology domain (AHD). Created in Biorender.com.

1.4.1 KMT2A::MLLT3 leukemia models

Since patient material only recapitulates the late stages of AML, model systems to better understand the oncogenic mechanisms driving malignant transformation are needed. Over the years, a variety of *in vitro* and *in vivo* models to study both normal hematopoiesis and leukemogenesis have been developed. With the advent of novel gene editing tools such as CRISPR/Cas9, which greatly facilitate knock-in, knock-out or mutation of genes, the availability of model systems has skyrocketed.

In addition, advances in DNA sequencing technologies have allowed large-scale genomic studies of pediatric leukemias through extensive sequencing of large numbers of tumor samples. Landmark projects such as the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) study provide a detailed multiomics-based molecular landscape of AML, which may be used to uncover oncogenic mechanisms suitable for targeted therapy (Milan et al., 2019). Importantly, these studies generated large publicly available datasets, which can be used to compare model systems to patient data in terms of gene expression profiles and chromatin accessibility (Milan et al., 2019).

In vivo models have the advantage of recapitulating the hematopoietic niche, allowing the study of environmental factors such as the interactions between the immune system and cancer cells, and have substantially contributed to the understanding of AML pathogenesis and biology. In this context, a variety of AML mouse models have been developed, including transgenic mice as well as transplantation and xenograft models, in which *in vitro* gene edited HSCs or patient-derived AML cells are transplanted into immunocompromised or irradiated recipients (Almosailleakh & Schwaller, 2019). Mouse models have also been used to study a variety of recurrent AML-associated fusion genes. In fact, the KMT2A::MLLT3 fusion was the first KMT2A fusion, which was modeled using gene-edited knock-in mice (Corral et al., 1996; Milne, 2017). Using a murine KMT2A::MLL3 transplantation model, it has been determined that both the age of the hematopoietic microenvironment and the cell of origin influence the leukemia phenotype, with a neonatal environment promoting the generation of mixed-lineage lymphocytic/myeloid acute leukemia, while an adult microenvironment promotes the generation of AML (Rowe et al., 2019). Although mouse models are highly useful, they still have phylogenetic differences in terms of gene location and function, as well as a higher susceptibility to mutation-induced transformation compared to humans (Perlman, 2016). Moreover, mouse models pose ethical concerns related to potential animal suffering, the importance of humane endpoints, and the need for strict oversight to ensure the welfare of research subjects. Additionally, mouse models are very expensive, requiring high amounts of time and resources, limiting affordable sample numbers, and making them unsuitable for large-scale drug screens.

Zebrafish are another useful model organism, which shares many conserved hematopoietic mechanisms with humans and are easy to maintain. Zebrafish leukemia xenograft models exist, and because of their small size and transparency during early development fish larvae can be cultivated in almost unlimited numbers and used for large-scale drug screens (Molina et al., 2021). Furthermore, KMT2A::MLLT3 expression in transgenic zebrafish embryos resulted in aberrant myeloid expansion and *Hox* gene upregulation, with this effect being reversed upon treatment with menin inhibitors (Tan et al., 2018). Nevertheless, zebrafish are anatomically very different from humans and the introduction of fusion proteins usually relies on ectopic expression techniques, which may potentiate an inaccurate phenotype.

In vitro models such as primary cancer cells or cancer cell lines are a cost-effective alternative and allow to study the effects of fusion proteins on self-renewal and proliferation in a controlled setting. Primary cancer cells derived from tumorous tissue such as bone marrow or peripheral

blood retain important disease characteristics, but are also very heterogenous due to the different patient genotypes and are prone to senescence *in vitro* (Richter et al., 2021). In contrast, cancer cell lines derived by extended culture adaptation of primary cancer cells, can be indefinitely expanded, and several cell lines of *KMT2Ar* AML, including KMT2A::MLLT3-expressing ones, have been established. At least seven cell lines, such as THP-1, NOMO-1 and MOLM-13, harbor a *KMT2A::MLLT3* fusion with an FAB M5 phenotype (Drexler et al., 2004). However, such conventional cancer cell lines often display secondary mutations and can only model terminal stages of AML, rendering them unsuitable for studying early steps of malignant transformation.

Other useful cellular models are cord blood-derived HSCs and progenitor cells derived from adult healthy donors. These cell types have the main advantage of overcoming the genetic heterogeneity of patient samples, since all cells are derived from one individual with a well-defined genetic background and normal isogenic controls at hand. Transformation via viral transduction of isolated CD34+ cord blood-derived progenitor cells with constructs encoding KMT2A::MLLT3, generated transplantable leukemic cells that recapitulated the gene expression changes of AML, also showing that in this cellular context the fusion protein can transform cells without any additional mutations (Barabé et al., 2017). The major drawbacks are that gene editing of cord blood-derived HSCs is difficult to achieve due to its low efficiency, low expandability, and decreased potency upon prolonged culture (Klaver-Flores et al., 2021).

In recent years, new and promising *in vitro* disease models based on human induced pluripotent stem cells (hiPSCs) have been developed, which will be described in the next chapter.

1.5 Human induced pluripotent stem cells

In 2007, Takahashi et al. succeeded in reprogramming somatic cells into hiPSCs by retroviral transduction of four reprogramming factors, namely octamer-binding transcription factor 4 (OCT4), SRY (sex determining Y)-box 2 (SOX2), krüppel-like factor 4 (KLF4), and MYC into adult dermal fibroblasts (Takahashi et al., 2007). Through the expression of these factors, a transcriptional autoregulatory loop is established, which promotes the upregulation of an endogenous pluripotency gene network while simultaneously silencing genes that drive differentiation (Omole & Fakoya, 2018). HiPSCs are characterized by their ability to differentiate into (almost) all cell types of the three germ layers, which makes them an excellent alternative to the ethically controversial embryonic stem cells (ESCs) and attractive for disease modeling, drug discovery and regenerative medicine.

Since their discovery, additional reprogramming factors and safer delivery methods have been discovered, resulting in acceptable levels of reprogramming efficiencies and allowing footprint-free reprogramming of HSCs, hematopoietic progenitors and even primary AML cells (Chen et al., 2015; Omole & Fakoya, 2018; Salci et al., 2015). Hematopoietic differentiation of hiPSCs or ESCs has been achieved by co-culture with stromal cells, embryoid body formation and recently developed monolayer-based hematopoietic differentiation protocols, allowing to study hematopoietic cell development *in vitro* (Dang et al., 2002; Kardel & Eaves, 2012; Nakano et al., 1996; Ruiz et al., 2019). Although hiPSC differentiation into HSCs holds great promise for leukemia modeling and therapy, further differentiation into certain blood cell types shows highly variable efficiencies (Lim et al., 2013). For instance, differentiation of hiPSCs into B-cells is limited by a myeloid differentiation bias, and engraftment of hiPSC-derived HSCs and progenitors is mostly inefficient (Papapetrou, 2019; Slukvin, 2013).

Successful reprogramming of cancer cells allows to produce unlimited quantities of patient-derived hiPSCs. Subsequent differentiation into affected cell types can be used to gain mechanistic insights into disease development, which is highly valuable if patient material is scarce. HiPSCs also offer a great opportunity to study the effects of disease drivers on early hematopoiesis, since *in vitro* hematopoietic differentiation shares many similarities with fetal hematopoiesis and HSC emergence *in vivo* (Donada et al., 2020; Kardel & Eaves, 2012). In addition, cellular reprogramming resets the epigenome, which allows to study genetic mutations without interference of the epigenetic memory and to track the establishment of expression signatures induced by oncogenic fusion proteins. Reprogramming of patient material further allows to investigate and characterize clonal composition, since hiPSCs are

generated from single cells and are passaged as clones (Chao et al., 2017). These features render hiPSCs a valuable model system for the study of clonal evolution and for drug resistance screening.

For example, *in vitro* differentiation of reprogrammed AML cells, harboring the *KMT2A::MLLT3* rearrangement and an additional *KRAS* mutation, into HSCs resulted in the reacquisition of leukemia-associated methylation signatures and gene expression patterns that had been erased during reprogramming. Furthermore, targeted therapy with DOT1L and MEK inhibitors revealed differential sensitivity and cytarabine resistance in subclones (Chao et al., 2017). Another more recent study generated hiPSCs derived from 15 AML patients, capturing different genetic subgroups of AML, including *KMT2A::MLLT3*. These AML-derived hiPSCs displayed hallmark AML features, including the ability to engraft in mice and faithfully reconstituted the clonal hierarchy (Kotini et al., 2023). However, reprogramming efficiencies of cancer, foremost leukemia cells remain low since they can harbor genetic alterations that impair the establishment of pluripotency (Papapetrou, 2019).

To circumvent this problem, gene editing of hiPSC cell lines derived from healthy donors can be used to model AML in a well-defined genetic background using isogenic controls (Bertuccio et al., 2022) (**Figure 10**)**.** The stepwise introduction of three driver mutations using sequential CRISPR gene editing of a healthy hiPSC line reflected all stages of leukemia development, with each subsequent mutation promoting a more malignant phenotype, resulting in the formation of transplantable AML-hiPSCs (Wang et al., 2021). By investigating underlying molecular changes, inflammatory signaling has been revealed as an early targetable vulnerability, and inhibition thereof in AML-hiPSCs resulted in decreased hematopoietic progenitor colony formation (Wang et al., 2021).

Taken together, these findings confirm that *in vitro* hematopoietic differentiation of genome-edited hiPSCs represents a promising and powerful tool to gain insights into the early stages of leukemia development and to perform drug screens for targetable lesions.

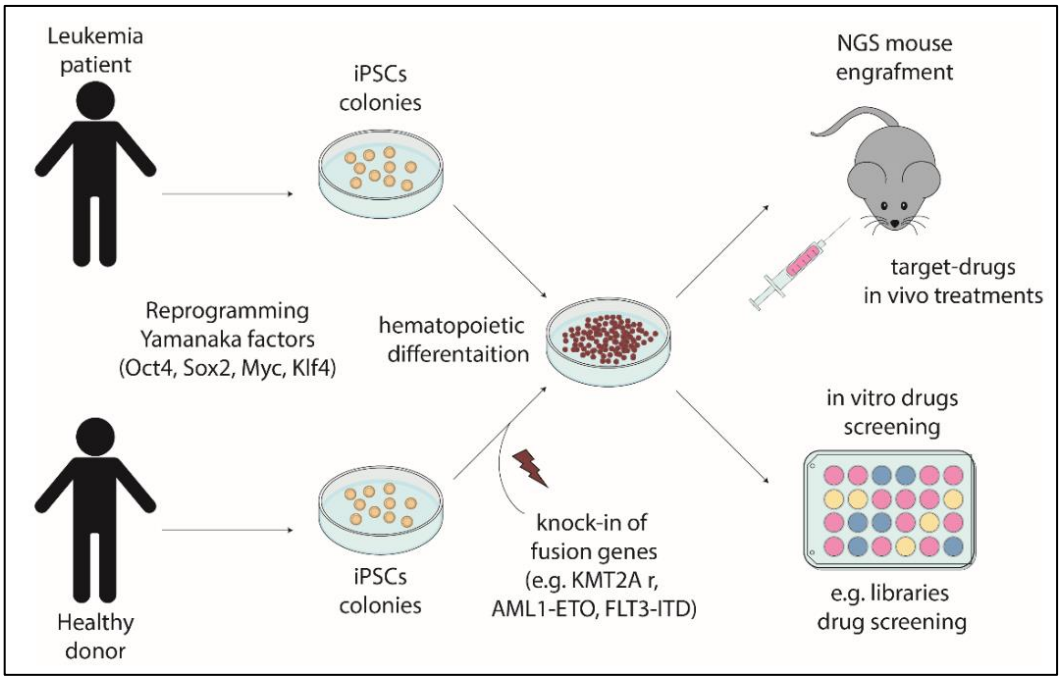


Figure 10: Modeling pediatric leukemia with hiPSC technology (modified from Bertuccio et al., 2022). Induced pluripotent stem cells (iPSCs), *KMT2A* rearrangements (*KMT2A* r).

1.6 CRISPR/Cas9 genome editing

The discovery and employment of the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 system for precise genome editing marked the beginning of a new gene editing era, which allowed the development of new gene editing tools, many of which were used to gain new insights into cancer biology and to generate unprecedented cancer models (Cong et al., 2013; Jinek et al., 2012; Katti et al., 2022). Originally, the CRISPR/Cas9 system belongs to a group of RNA-based defense systems in bacteria, facilitating adaptive immunity against foreign invading plasmids and viruses (Terns & Terns, 2011). It utilizes the bacterial endonuclease CRISPR-associated protein 9 (Cas9), which is recruited by a guide RNA (gRNA) consisting of a CRISPR RNA (crRNA) facilitating complementary DNA-RNA base-pairing to target loci and a trans activating CRISPR (tracrRNA) required for gRNA stability and Cas9-gRNA complex formation (Jinek et al., 2012). After precise binding of the Cas9 enzyme to the protospacer target sequence and its protospacer-adjacent motif (PAM), it induces double-stranded DNA breaks (DSBs) (Jinek et al., 2012). These DSBs activate one of two DNA repair mechanisms: homology-directed repair (HDR), which can be used to knock-in specific DNA sequences, or error-prone non-homologous end joining (NHEJ), which allows the disruption of coding or non-coding gene regions by small insertions or deletions (Cong et al., 2013; Jinek et al., 2012; Katti et al., 2022).

In recent years, many new Cas9 variants have been discovered or developed to improve and expand the usability of CRISPR-based systems. The original Cas9 enzyme, which is derived from *Streptococcus pyogenes* (*Sp*Cas9), is restricted to genomic loci that contain PAM sites with an NGG motif (Jinek et al., 2012). The use of Cas9 orthologs from different bacterial species, such as *Staphylococcus aureus* (*Sa*Cas9), recognizing different PAMs, allowed to broaden the target range of CRISPR/Cas9-based genome-editing systems (Katti et al., 2022). Additionally, protein engineering has allowed the generation of synthetic Cas mutants such as the “enhanced specificity” *Sp*Cas9 variant, which showed reduced off-target effects compared to the conventional *Sp*Cas9 enzyme (Slaymaker et al., 2016).

Nevertheless, Cas9-induced off-target DSBs are frequent, which often affect regions, that show high similarity to the target sequence, leading to undesirable genetic changes in the form of off-target knock-outs, random rearrangements or chromosomal instability, limiting its use to unambiguous targets (Zhang et al., 2015). To further reduce off-target effects, Cas9 nickases like the Cas9^{D10A} nickase have been developed. The Cas9^{D10A} enzyme harbors an alanine substitution (D10A) in the RuvC I domain, converting the *Sp*Cas9 endonuclease into an enzyme, which produces single-strand DNA breaks (SSBs), also called nicks, instead of DBSs at target loci (Cong et al., 2013). DNA nicks are noncanonical substrates for NHEJ but can be repaired by seamless base excision repair or non-canonical HDR, which results in lesser off-target effects and allows the precise knock-in of genes with high efficiencies (Chen et al., 2017).

Previously, gene fusions have been introduced into mice or cell lines by transfection or via viral transduction of respective vectors to be ectopically expressed from exogenous promotors. The downsides of this approach are the unphysiological expression levels of fusion proteins, the lack of regulation by surrounding genetic elements, and the absence of fusion partner haploinsufficiency (Alonso & Dow, 2021).

Knock-in of partner genes into endogenous gene loci offers a practicable alternative because it allows endogenous regulatory elements to control the expression of fusion proteins and partly recapitulates haploinsufficiency. In fact, knock-in via homologous recombination of the *MLLT3* 3’-terminal portion into exon 8 of *KMT2A* in murine embryonic stem cells resulted in expression of the KMT2A::MLLT3 fusion protein, leading to the development of AML in chimeric mice (Corral et al., 1996). Furthermore, a CRISPR-based gene editing strategy called *in trans* paired nicking has been developed, which utilizes the Cas9^{D10A} nickase to nick both

the genomic target locus and donor construct to facilitate highly efficient HDR-mediated integration of donor sequences into target loci (Chen et al., 2017). Utilizing such a gene editing strategy, our group introduced a *RUNX1::JAK2* fusion gene into hiPSCs via knock-in of the *JAK2* sequence into the *RUNX1* locus, which resulted in correct expression of the fusion protein upon hematopoietic differentiation (Fortschegger et al., 2021).

However, the described methods still do not fully recapitulate an actual translocation. The CRISPR/Cas9 system also enables the generation of chromosomal rearrangements both *in vivo* and *in vitro*. Previous approaches tried to model rearrangements using the Cre-loxP system, but were ineffective in generating translocations (Alonso & Dow, 2021). CRISPR-based systems offer the ability to target two distinct loci at once, inducing DSBs on each site, which can be joined via NHEJ, allowing to create a wide variety of chromosomal aberrations (Alonso & Dow, 2021). For example, in 2017, Vanoli et al. introduced the t(11;22)(p13;q12) translocation, resulting in the expression of an *EWSR1::WT1* fusion frequently found in sarcomas, into human mesenchymal cells, by inducing intronic DBSs in both partner genes with two different gRNAs. A donor plasmid containing two homology arms and a floxed puromycin cassette as a removable selection marker served as a repair template for CRISPR/Cas9-induced HDR, which resulted in expression of the EWSR1::WT1 fusion after transient puromycin selection and Cre recombination (Vanoli et al., 2017). These findings demonstrate the feasibility and effectiveness of CRIPSR/Cas-based systems in creating cancer models expressing oncogenic fusion proteins.

Nowadays, several genome editing tools are available for the introduction of pathogenic single nucleotide variants or small deletions or insertions, which may promote or drive tumor development. Base editing is a CRISPR/Cas-based gene editing technology, which generates precise point mutations through enzymatically induced base transitions or transversions in DNA or RNA without the creation of DSBs (Rees & Liu, 2018). Furthermore, the development of prime editors allows the introduction of small deletions or insertions (Anzalone et al., 2019). Prime editors consist of a Cas9 nickase fused to a reverse transcriptase (RT), which associates with a prime editing guide RNA (pegRNA). The pegRNA consists of a gRNA that has a 3' extension containing an RT template, which encodes the desired base edit, followed by a primer binding site. Upon DNA-nicking, the exposed 3'-hydroxyl group primes the RT to reverse transcribe the pegRNA-encoded RT template with the desired base edit into the target site (Anzalone et al., 2019). Through a process called flap equilibration the introduced edit also becomes incorporated on the other strand, allowing for both introduction or repair of mutations

with minimal off-target effects (Anzalone et al., 2019) (**Figure 11**). To improve editing efficiencies of the original prime editor system, several novel, more efficient versions have been developed that utilize additional proteins inhibiting DNA mismatch repair, or modified pegRNAs, which contain an additional 3' RNA structural motif, which increases prime editing efficiencies (Chen et al., 2021) (**Figure 11**). The ability of prime editors to accurately introduce point mutations, small insertions or deletions holds great promise for the introduction of oncogenic mutations into model systems.

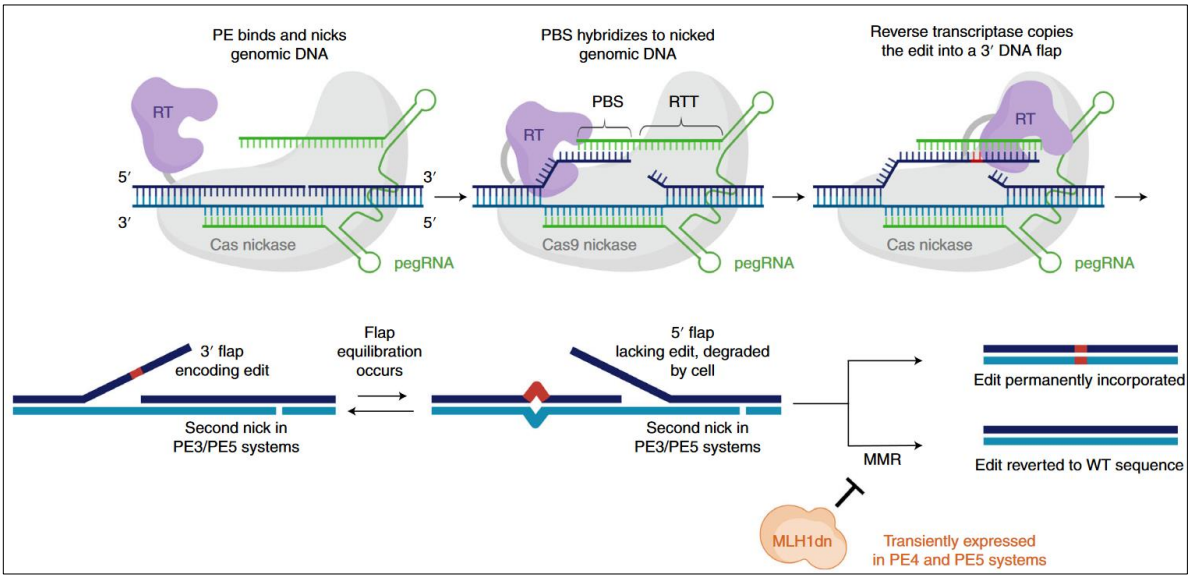


Figure 11: Prime editor mechanisms and components (modified from Anzalone et al., 2019). Guided by a pegRNA the Cas nickase induces a single strand break at the target sequence. The now exposed 3' DNA flap hybridizes with the PBS, activating the RT, which reverse transcribes the edit from the RTT into the 3' DNA flap. Through a process called flap equilibration the edit gets integrated into the opposite strand, establishing permanent incorporation of the edit. PE 4 and 5 systems additionally utilize the MLH1dn protein to reduce MMR and prevent repair/removal of the edit. Prime editor (PE), reverse transcriptase (RT), primer binding site (PBS), reverse transcriptase template (RTT), mismatch repair (MMR), wild-type (WT).

1.7 Hypothesis and aims of this study

Based on the oncogenic capacity of the KMT2A::MLLT3 fusion protein in other models, we hypothesized that its expression in hiPSCs upon hematopoietic *in vitro* differentiation phenocopies pediatric acute myeloid leukemia. Using state-of-the-art gene editing technologies, we aimed to establish hiPSCs models of KMT2A::MLLT3-driven leukemia, which would allow us to uncover early mechanisms of malignant transformation and to assess the impact of this fusion on hematopoietic differentiation, ultimately aiming to exploit the model for drug screening. In summary, the aims of this study were:

- Introduction of the KMT2A::MLLT3 fusion into a healthy donor hiPSC line using a CRIPSR/Cas9-mediated knock-in approach.
- Introduction of a secondary constitutive NRAS G12D mutation into KMT2A::MLLT3-expressing hiPSCs via prime editing or *in trans* paired nicking.
- Introduction of a genuine *KMT2A::MLLT3* translocation into a normal hiPSC line using CRIPSR/Cas9 technology.
- *In vitro* hematopoietic differentiation of the above-mentioned hiPSC lines and subsequent functional characterization as well as targeted gene expression analysis.

2. Material and Methods

2.1 DNA and RNA isolation for cloning and genotyping

DNA isolation was performed using the QIAamp Blood Mini Kit (Qiagen) following the manufacturer's instructions. RNA isolation was performed using TRIzol reagent (Invitrogen) following the manufacturer's instructions with minor changes. In brief, 1-5 million cells were resuspended in 1 ml of TRIzol reagent and incubated for 5 minutes. Next, 200 µl of chloroform (Merck) was added, the suspension mixed, and incubated for 2-3 minutes at room temperature followed by centrifugation for 15 minutes at 12 000 x *g* at 4 °C. The upper aqueous phase was transferred into a new tube and 20 µg of RNase-free glycogen (ThermoFisher Scientific) was added as coprecipitate. Subsequently, 0.5 ml of isopropanol was added, followed by incubation at room temperature for 10 minutes. Next, a centrifugation step at 12 000 x *g* at 4°C for 10 minutes was performed. The supernatant was discarded, and the precipitate containing the RNA was washed with 1 ml 75 % ethanol. In the last step, the RNA pellet was air dried for 5-10 minutes and resuspended in 20-50 µl RNase-free water (Ambion). RNA and DNA yields were measured using a DS-11 spectrophotometer (DeNovix).

2.2 Cloning of KMT2A::MLLT3 knock-in constructs

We employed an *in trans* paired nicking approach to introduce the *KMT2A::MLLT3* fusion into hiPSCs via knock-in of exons 6-11 of the *MLLT3 g*ene into the *KMT2A* locus. Furthermore, we incorporated a C-terminal degradation tag (dTAG; FKBP12 F36V), a hemagglutinin epitope-tag (HA), an internal ribosome entry site (IRES) driving the fluorescent reporter monomeric Kusabira-Orange2 (mKO2), and a puromycin resistance cassette (PuroR) flanked by two loxP sites (floxed). The required HDR donor construct, termed K-M_dTAG_Nick2_pUC57_donor for this approach, had already been designed, cloned, and tested prior to this thesis. The gRNA target site and the donor vector are shown in **Table 1** and **Supplementary Figure 1**, respectively.

Table 1: CRISPR/Cas9 gRNA target sites and applications

gRNA/crRNA name	gRNA target	Target Sequence 5' → 3' (/PAM)	Application
KMT2Ain9ex10 crRNA	*KMT2A* intron 9 exon 10	TCTGCTTCACAAT CCTCCTG/TGG	In trans paired nicking (*KMT2A::MLLT3*) knock-in
NRASex2 gRNA1	*NRAS* exon 2	GAGTACAAACTGG TGGTGGT/TGG	Prime editing (*NRAS* G12D)

NRASex2 crRNA2	*NRAS* exon 2	TGGTTGGAGCAGG TGGTGTT/GGG	In trans paired nicking (*NRAS* G12D) knock-in
KMT2Ain9 ***Sa*** gRNA1	*KMT2A* intron 9	GAATAAGAACTCC CATTAGCAG/GTG GGT	*KMT2A::MLLT3* translocation
MLLT3in5 ***Sp*** gRNA2	*MLLT3* intron 5	ACCTGTATAGCAG TAAACGT/AGG	*KMT2A::MLLT3* translocation

2.3 Cloning of *NRAS* G12D mutation and knock-in constructs

To introduce an *NRAS* G12D mutation, three gene editing approaches were conducted. The first approach utilized the PE4 prime editing system. For this, we first designed a pegRNA targeting exon 2 of *NRAS* and encoding the *NRAS* G12D mutation as well as a NarI cut site. The gRNA part of the pegRNA was designed using the UCSC genome browser. The target site and corresponding designed pegRNA sequences are shown in **Table 1** and **Table 2**, respectively.

Table 2: pegRNA sequence

pegRNA name	**Sequence 5' → 3' (gRNA, scaffold, 3' extension)**
pegRNA1 NRAS G12D	GAGTACAAACTGGTGGTGGTGTTTTAGAGCTAGAAATAGCA AGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGG CACCGAGTCGGTGCACCATCGGCGCCAACCACCACCAGTT TGTACTC

Cloning of the pegRNA vector was adapted from a previous publication with minor changes (Anzalone et al., 2019). The conducted cloning steps are depicted in **Figure 12** and the oligos for pegRNA cloning are provided in **Table 3**. The pegRNA spacer and 3' extension oligos were designed with overhangs specific for the two flanks of the BsaI digested pegRNA acceptor vector. After oligo annealing, the pegRNA vector was assembled via conventional ligation using the T4 ligase DNA ligase (Promega).

Table 3: Oligos for pegRNA cloning (all from Microsynth)

Oligo	**Sequence 5' → 3'**	**pegRNA part**
NRAS_G12pegRNA-F1	caccGAGTACAAACTGGTGGTGGTG TTTT	(Proto)spacer (gRNA)
NRAS_G12pegRNA-R1	CTCTAAAACACCACCACCAGTTTG TACTC	(Proto)spacer (gRNA)

Scaffold1-F	AGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTG AAAAAGTGGCACCGAGTCG	Scaffold
Scaffold1-F	GCACCGACTCGGTGCCACTTTTTC AAGTTGATAACGGACTAGCCTTAT TTTAACTTGCTATTTCTAG	Scaffold
NRAS_G12D_3ext-F1	GTGCACCATCGGCGCCAACCACC ACCAGTTTGTACTC	3' extension
NRAS_G12D_3ext-R1	aaaaGAGTACAAACTGGTGGTGGTT GGCGCCGATGGT	3' extension

In brief, 2 µl of 1 µM annealed 5' phosphorylated pegRNA Scaffold1 oligodeoxynucleotides were mixed with 0.5 µl T4 DNA ligase (10 units), 1 µl 10x T4 DNA ligase buffer, and 2.9 µl ddH_2O. Next, the mixture was incubated for 1 hour and 30 minutes at 16 °C for ligation. Then 1 µl of 1 µM annealed pegRNA spacer oligos was added, followed by incubation for 1 hour at 16 °C. After incubation, 1 µl of 1 µM annealed pegRNA 3'extension oligos was added, and incubated for 1 hour at 16 °C. In the last step, 0.6 µl (30 ng total) BsaI digested pegRNA acceptor vector was added followed by incubation for 1 hour and 30 minutes at 16 °C.

All ligations used for cloning were precipitated with glycogen. For this, 1/10 of total volume 3 M sodium acetate pH 5.2, 1 µl glycogen (20 µg) (Roche) and 2.5x of total volume 96 % ethanol were added, followed by incubation at -20 °C for one hour. Next, a centrifugation step at 4 °C for 30 minutes at 13 000 x *g* was performed. The supernatant was discarded, and the remaining DNA pellet was washed with 2x of total volume 75 % ethanol. The pellet was air dried and resuspended in 5-10 µl ddH_2O. The assembled pegRNA sequence and pegRNA vector are depicted in **Supplementary Figure 2**. The pegRNA acceptor vector as well as the prime editor plasmid encoding the Cas9 nickase were purchased from Addgene (**Supplementary Figure 2** and **Supplementary Figure 3**).

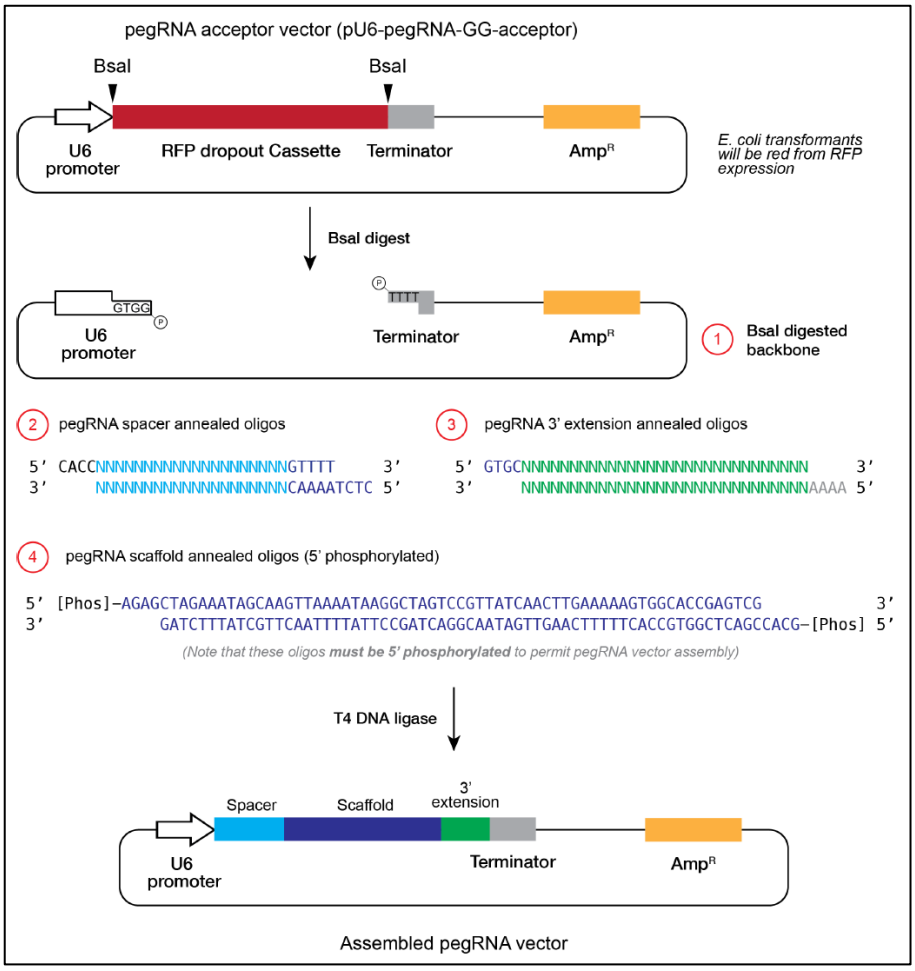


Figure 12: pegRNA vector cloning steps (modified from Anzalone et al., 2019). Ampicillin resistance (Amp^R), red fluorescent protein (RFP).

Due to inefficient prime editing, we next tested an approach in which we intended to introduce the NRAS G12D mutation by *in trans* paired nicking knock-in. First, the genomic region including exon 2 of *NRAS* was PCR amplified using primers that encode the NRAS G12D mutation and a NarI cut site to create the *NRAS* HDR donor construct depicted in **Supplementary Figure 4**. The integration of a synonymous NarI restriction site served to facilitate screening for correctly edited cells and prevent unintended nicking within the donor *NRAS* sequence. Next, genomic DNA was extracted from wild-type (WT) hiPSCs and PCR was performed with primers NRAS_proto_ex1-F1 and NRAS_G12D_NarI_R to amplify the 5' part and primers NRAS_G12D_NarI_F and NRAS_proto_in2-R1 for the 3' part of the *NRAS* homology (**Table 4**). For the vector backbone, 2 µg K-M_dTAG_Nick2_pUC57_donor was digested with 15-20 U of restriction enzymes NotI and SgsI in 1x buffer R, all from ThermoFisher Scientific. In general, restriction digests for cloning were carried out using 2 µg of vector DNA or 50-100 ng of DNA for smaller fragments/inserts, 10-20 U of enzyme, and

buffers recommended in the manual or by the DoubleDigest Calculator tool from ThermoFisher Scientific. After purification of the backbone and the *NRAS* homology parts, Gibson Assembly was performed using the NEBuilder HiFi DNA Assembly Cloning Kit (NEB) to obtain the NRAS_G12D_NarI_donor_trans vector (**Supplementary Figure 4**; gRNA target sites are listed in **Table 1**).

As we encountered again unsatisfactory gene editing efficiency, we next chose a selection-based approach similar to the *KMT2A::MLLT3* knock-in, in which we aimed to insert exons 2-7 of *NRAS* harboring the G12D mutation, a NarI site, a polyadenylation site and a PuroR into exon 2 of *NRAS*. For cloning of the respective donor, total RNA was extracted from WT hiPSCs and reverse transcription (RT) followed by PCR was performed using primers NRAS_G12D_NarI_F (containing the NRAS G12D mutation and NarI restriction site) and NRASex7-R1_GA to retrieve *NRAS* exons 2-7 (**Table 4**). The 5' homology arm (5'HA) containing *NRAS* exon 1 and intron 1 was generated by PCR amplification of hiPSC WT DNA using the primers NRAS_proto_ex-F1 and NRAS_G12D_NarI_R, the former containing a restriction site for NotI (**Table 4**). The 3' homology arm (3'HA) containing *NRAS* exon 2 and intron 2 was generated by PCR amplification of hiPSC WT DNA using the primers NRASex2-F2_XhoI and NRAS_proto_in2_R1 containing XhoI and SgsI restriction sites, respectively (**Table 4**). The backbone was obtained by digestion of the K-M_dTAG_Nick2_pUC57_donor (**Supplementary Figure 1**) with NotI and EcoRI to first allow insertion of the 5'HA and mutated *NRAS* exons 2-7 via Gibson assembly using the NEBuilder HiFi DNA Assembly Cloning Kit (NEB). Integration of the 3'HA was achieved by digestion of both the 3'HA and the assembled intermediate with SgsI and XhoI followed by ligation using T4 DNA ligase (Promega). The NRASex2 crRNA2 target site is also present outside of both homology arms so that the nicked donor can act as an HDR template (**Supplementary Figure 5**; gRNA target sites are listed in **Table 1**). Restriction enzymes and buffers were purchased from ThermoFisher Scientific.

Table 4: List of primers for cloning of *NRAS* G12D donor constructs (all from Microsynth)

Primer	Sequence 5' → 3'
NRAS_G12D_NarI_F	GTACAAACTGGTGGTGGTTGGcGCcGaTGGTGTTGGGAAAAG CGCAC
NRASex7-R1_GA	ATTCTCCTGAGACGAATTCAAGTCAGGACCAGGGTGTC
NRAS_proto_ex1-F1	gtcgagatgcggccgcTGGTTGGAGCAGGTGGTGTTGGGATTTTTC CCGGCTGTGGTCC
NRAS_G12D_NarI_R	GTGCGCTTTTCCCAACACCAtCgGCgCCAACCACCACCAGTTT GTAC

NRASex2-F2_XhoI	TACTCGAGGTTGGGAAAAGCGCACTGAC
NRAS_proto_in2-R1	gcatccgatGGCGCGCCCAACACCACCTGCTCCAACCAACAGTC TCGCTACTATGGCCT

2.4 Cloning of t(9;11)(p22;q23) translocation constructs

To induce a translocation between chromosomes 9 and 11, and consequently a *KMT2A::MLLT3* fusion gene in hiPSCs we used two orthogonal Cas9 enzymes from different species, namely *Staphylococcus aureus* (*Sa*) and *Streptococcus pyogenes* (*Sp*), associating with the *Sa*Cas9 gRNA1 targeting intron 9 of *KMT2A* and the *Sp*Cas9 gRNA2 targeting intron 5 of *MLLT3*. Because of low basal interchromosomal translocation efficiency, we switched to a selection-based approach adapted from Vanoli et al., in which we knocked-in a removable, LoxP-flanked puromycin resistance cassette at the translocation breakpoint (Vanoli et al., 2017).

Both plasmids containing the respective Cas9 and gRNA were designed and ordered from Vectorbuilder (**Supplementary Figure 6** and **Supplementary Figure 7**; gRNA target sites are listed in **Table 1**). The donor construct encoding the LoxP-flanked puromycin resistance cassette is depicted in **Supplementary Figure 8**. As a starting point for cloning, the K-M_dTAG_Nick2_pUC57_donor was used as a backbone (**Supplementary Figure 1**). The 5'HA spanning intron 8, exon 9 and intron 9 of *KMT2A* as well as the 3'HA spanning intron 5, exon 6 and intron 6 of *MLLT3* were generated by PCR amplification of WT hiPSC DNA using primers KMT2A_5HA_trans_F and KMT2A_5HA_trans_R for the former and primers MLLT3_3HA_trans_F and MLLT3_3HA_trans_R for the latter (**Table 5**). Note that both homology arms are flanked by a respective gRNA target site, and cleavage of the donor results in a linearized HDR template. Digest of the backbone and 5'HA was performed with restriction enzymes NotI and XbaI followed by DNA ligation using T4 ligase (Promega). Subsequently, digestion of the intermediate vector and 3'HA was performed with restriction enzymes XhoI and SgsI followed by DNA ligation.

Table 5: List of primers for cloning of t(9;11)(p22;q23) translocation constructs (all from Microsynth)

Primer	**Sequence 5' → 3'**
KMT2A_5HA_trans_F	tgcggccgcGAATAAGAACTCCCATTAGCAGGTGGGTCTTTGTG GCCCCACATGTTC
KMT2A_5HA_trans_R	cctctagaTAATGGGAGTTCTTATTCATTCCCC
MLLT3_3HA_trans_F	tctcgagTAGGAGTGAAATCAAAGCCAGG

MLLT3_3HA_trans_R	tGGCGCGCCTACGTTTACTGCTATACAGGTaCTAGAGCCTTG CCTCGTACA

2.5 Bacterial transformation

Bacterial transformation was performed by heat shock according to standard laboratory protocols. In brief, thawed One Shot Stbl3 bacteria (ThermoFisher Scientific) or NEB 5-alpha bacteria (NEB) were shortly incubated with corresponding plasmid ligations on ice. Next, bacteria were heat shocked at 42 °C for 45 seconds, cooled and subsequently recovered in S.O.C medium (ThermoFisher Scientific) for 30 min at 37 °C. Transformed bacteria were next spread on LB-agar-Ampicillin plates and incubated at 37 °C overnight. Colonies were picked, expanded for one day in LB containing Ampicillin, and plasmid DNA was isolated using the EndoFree Plasmid Maxi kit (QIAGEN) following the manufacturer's instructions.

2.6 Cell culture

For genome editing and downstream experiments, a commercial Gibco episomal hiPSC line (Gibco; ThermoFisher Scientific) was used. HiPSCs were cultivated in mTeSR plus media (STEMCELL Technologies) on Matrigel-coated 6-well plates (Corning) under hypoxic conditions at 37 °C with 5 % CO_2 and 3 % O_2. For coating, frozen Matrigel aliquots were thawed on ice for 1 hour, followed by dilution in 12 ml cold DMEM/F-12 (ThermoFisher Scientific). The solution was dispensed in 6-well plates using 1 ml per well. The plates were incubated for 1 hour at 37°C and thereafter washed twice using the same volume of DMEM/F-12. Around 2 ml of media was left until usage (within one week) to prevent drying of Matrigel. Medium was exchanged every day, and hiPSCs were passaged every three to four days at a split ratio of 1:6 using StemPro Accutase (ThermoFisher Scientific) followed by cultivation in mTeSR plus media supplemented with 10 µM ROCK inhibitor (ROCKi) Y-27632 (STEMCELL Technologies). For cells harboring the *KMT2A::MLLT3* fusion gene, the medium was supplemented with 100 nM dTAG-13 (Tocris Bioscience) to specifically degrade the fusion protein and prevent interfering with the maintenance of the hiPSC state and differentiation.

The cell lines THP-1 and NOMO-1 were purchased from DSMZ and used as controls. Both express a KMT2A::MLLT3 fusion (*KMT2A* exon 8 or 9 fused to *MLLT3* exon 6) and harbor a NRAS G12D or KRAS G13D mutation (Drexler et al., 2004). Cell lines were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10 % fetal bovine serum (FBS) and Penicillin/Streptomycin (all from ThermoFisher Scientific) under normoxic conditions at 37 °C with 5 % CO_2. Cells were passaged every two to three days.

To exclude mycoplasma contamination, all cell lines were regularly tested using the MycoAlert Mycoplasma Detection kit (Lonza) following the manufacturer's instructions.

2.7 Generation of gene-edited hiPSCs and isolation of single cell clones

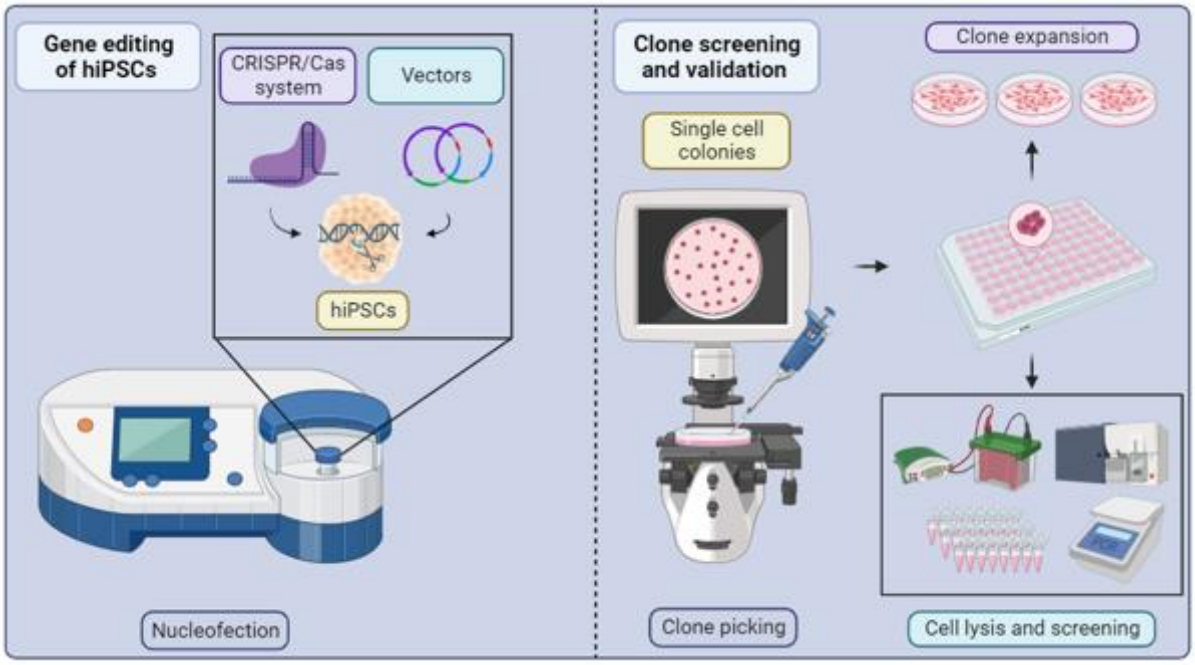


Figure 13: Generation of gene edited hiPSCs and screening of positive clones. Created in Biorender.com.

Gene-edited hiPSCs were generated by electroporation using a Nucleofector 2b (Lonza) (program A-023) and the human stem cell Nucleofector Kit 2 (Lonza). All steps are summarized in **Figure 13**. For nucleofection one to two million hiPSCs were used. CrRNA and tracrRNA were diluted to a final concentration of 200 µM in duplex buffer (IDT). To allow for ribonucleoprotein (RNP) formation, 1.5 µl of crRNA and 1.5 µl of tracrRNA were mixed in a PCR tube and incubated at 95 °C for 5 minutes with subsequent cooling to room temperature, followed by the addition of 4 µl 62 µM Cas9 D10 Nickase V3 (IDT). Next, cells were washed with Dulbecco's phosphate buffer saline (DPBS) (ThermoFisher Scientific) following detachment with Accutase, a DPBS wash and centrifugation at 450 g for 5 minutes at room temperature. The supernatant was aspirated, and the cell pellet was resuspended in 100 µl Amaxa human stem cell kit 2 solution (Lonza). Subsequently, 7 µl of the associated RNP (~ 250 pmol) and a total of 5 µg of donor vector were added. The respective RNPs were introduced along with 5 µg of the K-M_dTAG_Nick2_pUC57_donor for the *KMT2A::MLLT3* knock-in, 2 µg of pU6_pegRNA_GG_pegRNA1_NRAS G12D, and 3 µg of pCMV_PE2_P2A_hMLH1dn for the prime editing *NRAS* G12D approach. For the simple or

selection-based *in trans* paired nicking approach, 5 µg of NRAS_G12D_NarI_donor_trans or NRAS_PuroR_donor were employed, respectively.For induction of the *KMT2A::MLLT3* translocation, the *Sa*Cas9 and *Sp*Cas9 enzymes, as well as their respective gRNAs were expressed from plasmids. For transfection, 2 µg of plasmids Sa_KMT2A_gRNA1 and Sp_MLLT3_gRNA2 each, as well as 3 µg of K-M_PuroR_translocation donor, were used.

The cell suspension was then transferred into the provided cuvettes and nucleofected. After nucleofection the cells were resuspended in mTeSR plus media supplemented with 10 µM ROCKi and seeded into Matrigel-coated 6-well plates. Following 1-2 days of recovery, the cells were selected twice by a one-day 0.5 µg/ml puromycin (Merck) pulse followed by another nucleofection with 5 µg of a pCaGGS-Cre vector. After cell expansion, 1000-2000 cells were seeded as single cells in mTeSR plus media supplemented with 10 % CloneR (STEMCELL Technologies) on Synthemax II-SC (Corning) coated 10-cm dishes. On every other day, mTeSR plus medium was exchanged, and after nine to ten days, 96 of the emerged round colonies were picked manually with pipette tips under an EVOS XL Core Imaging System (ThermoFisher Scientific) and seeded into two replicate Matrigel coated 96-well plates in mTeSR plus media supplemented with 10 % CloneR (all from STEMCELL Technologies) and penicillin/streptomycin (ThermoFisher Scientific). After three to four days, replicate clones were lysed and genotyped via PCR. Positive clone replicates were expanded and further genotyped by sequencing.

2.8 FACS analysis and sorting

To assess the purity of *KMT2A::MLLT3* hiPSC knock-in clones, flow cytometric detection of mKO2 expression was performed. For this purpose, cells were detached with Accutase, singularized in mTeSR plus supplemented with 10 µM ROCKi, and transferred into FACS tubes. Intact cells were gated according to forward and sideward scatter and analyzed for mKO2 fluorescence in the phycoerythrin channel using the FCS Express 7 software. Purity was assessed by measuring the percentage of mKO2-positive knock-in cells. For impure clones containing ≥ 5% contaminating wild-type cells, sorting of mKO2-positive cells was performed by the CCRI FACS core facility.

2.9 Hematopoietic differentiation of hiPSCs and cultivation of myeloid cells

Hematopoietic differentiation was performed using the STEMdiff Hematopoietic Kit (STEMCELL Technologies) following the manufacturer's instructions with minor changes. In brief, 2000 hiPSCs per well were seeded at low density as cell aggregates into Matrigel-coated

12-well plates in mTeSR plus medium containing 10 µM ROCKi and 100 nM dTAG-13. Medium was changed to mTesR plus containing only 100 nM dTAG-13 after two days, and hematopoietic differentiation was started after three to four days, depending on colony density. For this purpose, the medium was exchanged to STEMdiff hematopoietic complete medium A (HEM A) containing 3 µM GSK-3 Inhibitor CHIR99021 (Sigma-Aldrich) and 100 nM dTAG-13 followed by cultivation under normoxic conditions. On day 1, the medium was exchanged to HEM A containing 100 nM dTAG-13 only, and on day 3, the medium was exchanged to STEMdiff hematopoietic complete medium B (HEM B) supplemented with 100 nM dTAG-13. On day 6, while one half of the cells was kept on HEM B with dTAG-13, the degrader was washed out and no longer added to the other half to induce KMT2A::MLLT3 protein expression. We delayed the induction of the fusion protein because previous experiments suggested that premature expression of KMT2A::MLLT3 interfered with early hematopoietic differentiation. Furthermore, as dTAG-13 from Sigma-Aldrich was not functional in a set of experiments and controls, we later switched to an effective dTAG-13 compound from Tocris Bioscience (both listed in **Supplementary Table 1**).

Subsequently, half HEM B media (with or without dTAG-13) exchanges were performed every two to three days, and on days 12 to 13, the supernatant was collected, and live cells were harvested using the MACS Dead Cell Removal Kit (Miltenyi) following the manufacturer's instructions. Briefly, up to 10^7 total cells were resuspended in 100 µl magnetic bead suspension and incubated for 15 minutes at RT. Meanwhile, MS columns were placed on a magnetic stand and equilibrated with 500 µl 1x binding buffer. After incubation, the cells were resuspended in 500 µl 1x binding buffer and transferred onto the magnetic columns, followed by washing them three times with 500 µl 1x binding buffer. The flow through was collected, and cell yield and viability were measured by mixing 10 µl cell suspension with 10 µl 0.2% Trypan Blue solution (Sigma-Aldrich), followed by counting dead (blue) and live (unstained) cells in a Bürker-Türk chamber (Brand). The adherent cell fraction was lysed and used for RNA or protein extraction. The hiPSC-derived HSCs were either kept in liquid culture, seeded for colony assays, or also lysed using TRIzol reagent for RNA extraction. For liquid culture, the cells were cultivated in StemPro34-SFM media (ThermoFischer Scientific) supplemented with SCF (50 ng/µl), EPO (10 ng/ml), GM-CSF, IL-3, IL-6, G-CSF, and TPO (all 20 ng/ml each; PeproTech), under normoxic conditions. Dead cell removal was performed every three to six weeks using the MACS Dead Cell Removal Kit to remove accumulated dead cell debris.

2.10 Cell counting and growth curve analysis

Growth curves of hiPSC-derived HSPCs in liquid culture were determined by Trypan Blue exclusion cell counting every two to five days. Cumulated total cell numbers were extrapolated to account for sampling and archiving. Fresh StemPro34 containing the seven cytokines, and 100 nM dTAG-13 or not was added regularly.

2.11 Cytokine dependency assay

To investigate whether the cells show cytokine independency, hiPSC-derived HSCs were cultivated in StemPro34 supplemented either with single SCF (50 ng/µl), EPO (10 ng/ml), GM-CSF, IL-3, IL-6, G-SCF or TPO (all 20 ng/ml), all 7 mentioned cytokines, or none. Cell growth was assessed by counting cells every two to five days, as described in **2.10 Cell counting and** growth curve **analysis**.

2.12 Oncogene dependency assay

To assess if cells depend on the expression of the KMT2A::MLLT3 fusion, long-term liquid cultured hematopoietic cells were cultivated in StemPro34 containing the seven cytokines and supplemented without or with 100 nM dTAG-13 or dTAG-v1 (both from Tocris Bioscience) to deplete the fusion protein. The medium was exchanged weekly, and the cells were counted every three to five days as described.

2.13 MethoCult and MegaCult colony formation assays

The colony formation assays were performed using cytokine enriched methylcellulose medium MethoCult H4435 Enriched or the MegaCult-C Complete Kit with cytokines (both STEMCELL Technologies). For all colony assays, either 100 nM dTAG-13 or vehicle (140 µM DMSO only) was added to the assay medium.

For MethoCult formation assays, 1000 cells were seeded after harvest from hematopoietic differentiation into 6-well plates with 300 µl Iscoves's Modified Dulbecco's Medium (IMDM) supplemented with 2 % FBS (both ThermoFisher Scientific) and 3 ml cytokine-enriched MethoCult H4435. After twelve to fifteen days, emerging colonies were counted under a 3D microscope (Leica M125) with dark field illumination using the STEMgrid-6 (STEMCELL Technologies) as an aid. Colony forming units (CFUs) were enumerated according to CFU types and colony size. Colonies comprising a minimum of 20 granulocytes were considered as CFU-G. CFU-Ms comprised a minimum of 20 macrophages/monocytes. A combination of both cell types found in single or multiple lobed colonies of at least 40 cells was considered as CFU-GM. The term CFU-GEMM describes a colony comprised of mostly more than 500 cells of

erythroid, granulocytic, monocytic and/or megakaryocytic cells. After colony counting, the cells were harvested by resuspension in 10 ml IMDM containing 2 % FBS, followed by centrifugation at 450 x g for 5 minutes at RT. Cell numbers were counted, and the cells were lysed using TRIzol reagent for RNA isolation and stored at -80 °C.

For MegaCult assays, 4000 cells were seeded after harvest and cultured for 12 days according to the manufacturer's instructions. Fixation and subsequent staining were also performed following the manufacturer's instructions. Stained colonies (non-megakaryocytic blue cells, megakaryocytic red CD41+ cells; mixed blue and red cells) were counted by two independent operators (an adhesive grid tape was applied and 20 % of the area was randomly chosen) under a light microscope (Nikon TMS). Non-megakaryocytic CFUs were composed of at least 20 cells, megakaryocytic colony forming units (CFU-Mks) were also classified by size: small 3-20, medium 21-49, and large ≥ 50 CD41+ cells.

2.14 Microscopy

Representative pictures of cells in culture as well as colonies of both MethoCult and MegaCult assays were taken using the EVOS XL Core Imaging System (ThermoFisher Scientific).

2.15 Cell lysis for PCR screening of single cell clones

To screen for successfully gene-edited hiPSC clones, cells were washed with DPBS and DNA was retrieved by crude cell lysis using a genotyping lysis buffer consisting of 50 mM KCl, 2 mM $MgCl_2$, 10 mM Tris-HCl pH 8.5, 0.45 % NP40, 0.45 % TWEEN-20, and 1 mg/ml Proteinase K (all Sigma-Aldrich/Merck). The lysates were then incubated at 65 °C for 3-4 hours, followed by incubation at 95 °C for 10 minutes.

2.16 Genotyping PCRs and sequencing of bulk cells and individual clones

To screen for positive clones as well as to genotype *KMT2A::MLLT3* knock-in bulk cells and expanded clones, PCRs were performed using 2 units of HotStarTaq DNA polymerase and buffer (QIAGEN), approximately 200 ng DNA or cDNA, 500 nM specific primers, and 200 nM dNTPs (Promega). The used primers for the *KMT2A::MLLT3* knock-in genotyping PCRs are listed in **Table 6**. PCR was conducted in a Biometra Trio thermocycler (Analytik Jena) using the cycling conditions provided in **Table 7**.

Table 6: Primers and respective PCRs for *KMT2A::MLLT3* knock-in genotyping/screening

Primer	PCR	Product size
PuroRmidF2	PuroR PCR	1315 bp

KMT2Ain10-R2	PuroR PCR	1315 bp
KMT2Aex9-F1	5'HA PCR	922 bp
MLLT3ex8-R1	5'HA PCR	922 bp
mKO2midF2	3'HA PCR	1276 bp
KMT2Ain10-R2	3'HA PCR	1276 bp
KMT2Aex9-F1	KMT2A WT PCR	1078 bp
KMT2Ain10-R2	KMT2A WT PCR	1078 bp

Table 7: PCR conditions using HotStarTaq DNA Polymerase

Temperature Settings			
Cycles	**Step**	**Temp. (°C)**	**Time**
1	Activation	95	15 min
40x	Denaturation	94	30 sec
	Annealing	60	30 sec
	Extension	72	1-2 min (~1 min/kbp)
1	Final Extension	72	5 min
1	Cool	4	∞

For genotyping of cell bulks generated by prime editing and *in trans* paired nicking, PCRs were performed using the OneTaq DNA Polymerase and buffer (NEB), approximately 200 ng DNA or cDNA, 500 nM specific primers, and 200 nM dNTPs (Promega). Additionally, 200 ng of the NRAS_G12D_NarI_donor_trans plasmid DNA were used for PCR as a positive digest control. The used primers are listed in **Table 8**. PCRs were conducted in a Biometra Trio thermocycler (Analytik Jena) with the conditions provided in **Table 9**. Additionally, these PCR products were directly digested by adding 1 µl NarI enzyme (NEB) to the reaction mix, followed by incubation at 37 °C for one hour.

Table 8: Primers and PCR for the *NRAS* G12D mutation

Primer	**PCR**	**Product size**
NRASex1-F3	NRAS PE/trans PCR	976 bp
NRASin2-R1	NRAS PE/trans PCR	976 bp
NRASex1-F1	NRAS digest PCR	976 bp
NRASin2-R1	NRAS digest PCR	976 bp

Table 9: PCR conditions using OneTaq DNA Polymerase

Temperature Settings			
Cycles	**Step**	**Temp. (°C)**	**Time**
1	Activation	94	30 sec

30x	Denaturation	94	30 sec
	Annealing	60	30 sec
	Extension	68	1 min
1	Final Extension	68	5 min
1	Store	4	∞

To genotype *NRAS* G12D knock-in cell bulks and to screen for positive clones generated by the selection-based gene editing approach, PCRs using the HotStarTaq DNA Polymerase (QIAGEN), approximately 200 ng DNA or cDNA, 500 nM specific primers, and 200 nM dNTPs (Promega) were performed. In addition, RNA was extracted from each of the positive *NRAS* G12D clones, and cDNA synthesis followed by PCR was performed to assess the expression of the *NRAS* G12D transcript variant. The used primers for the *NRAS* G12D knock-in genotyping PCRs are listed in **Table 10**. As usual, PCRs were conducted in a Biometra Trio thermocycler with the conditions provided in **Table 7**.

Table 10: Primers and PCRs for genotyping/screening of *NRAS* G12D knock-in

Primer	PCR	Product size
NRASex1-F3	NRAS 5'HA PCR	941 bp
NRASex3-4-R1	NRAS 5'HA PCR	941 bp
Donor_loxP2_for	NRAS 3'HA PCR	667 bp
NRASin2-R2	NRAS 3'HA PCR	667 bp
PuroRmidF2	PuroR PCR	1093 bp
NRASin2-R1	PuroR PCR	1093 bp
NRASex1-F1	NRAS WT PCR	1200 bp
NRASin2-R2	NRAS WT PCR	1200 bp
NRASex1-F3	NRAS G12D delta PCR	1537 bp
Donor_loxP2_rev	NRAS G12D delta PCR	1537 bp
Donor_loxP2_for	NRAS G12D KI PCR	1543 bp
NRASin2-R3	NRAS G12D KI PCR	1543 bp
NRASex1-F3	NRAS WT allele PCR	2200 bp
NRASin2-R3	NRAS WT allele PCR	2200 bp
NRASex1-F3	NRAS transcript PCR	460 bp
NRASex3-4-R1	NRAS transcript PCR	460 bp

For genotyping of *KMT2A::MLLT3* translocation cell bulks and screening for positive clones, PCRs were performed using the HotStarTaq DNA Polymerase (QIAGEN), approximately 200 ng DNA or cDNA, 500 nM specific primers, and 200 nM dNTPs (Promega). The used primers are summarized in **Table 11**, and the PCR conditions were the same as for other PCR reactions (**Table *7***).

Table 11: Primers and PCRs for genotyping/screening of the *KMT2A::MLLT3* translocation

Primer	PCR	Product size
KMT2Ain8-F2	KMT2A 5'HA PCR	1115 bp
mPgk1proR1	KMT2A 5'HA PCR	1115 bp
Donor_loxP2_for	MLLT3 3'HA PCR 1	912 bp
MLLT3in6-R2	MLLT3 3'HA PCR 1	912 bp
mPgk1_3UTR_F1	MLLT3 3'HA PCR 2	1180 bp
MLLT3in6-R2	MLLT3 3'HA PCR 2	1180 bp
KMT2Aex9-F1	KMT2A WT PCR 1	770 bp
KMT2Aex10in10-R1	KMT2A WT PCR 1	770 bp
KMT2Ain8-F2	KMT2A WT PCR 2	977 bp
KMT2Ain9-R1	KMT2A WT PCR 2	977 bp
MLLT3in5-F2	MLLT3 WT PCR	673 bp
MLLT3ex6-R1	MLLT3 WT PCR	673 bp
KMT2Ain8-F2	K::M PCR	1757 bp (+loxP site: 1837 bp)
MLLT3in6-R2	K::M PCR	1757 bp (+loxP site: 1837 bp)
KMT2Ain8-F2	KMT2A 5'HA delta PCR	949 bp
Donor_loxP2_rev	KMT2A 5'HA delta PCR	949 bp
KMT2Ain9-F2	MLLT3 3'HA delta PCR	1024 bp
MLLT3in6-R2	MLLT3 3'HA delta PCR	1024 bp
MLLT3in5-F4	M::K PCR	1200 bp
KMT2Ain9-R2	M::K PCR	1200 bp

For Sanger sequencing (Microsynth) PCR products were purified using the Monarch PCR & DNA Cleanup Kit (NEB) following the manufacturer's instructions and subsequently sent for sequencing using 3 µl (10 µM) of the respective primer. The amount of PCR product sent for sequencing was calculated via the following formula:

$$\frac{\frac{length\ of\ fragment\ [bp]}{100} \times 18\ ng}{DNA\ concentration\ [\frac{ng}{\mu l}]}$$

Sequences were aligned using the CLC Genomic Workbench 23 software (QIAGEN).

2.17 cDNA synthesis and gene expression analysis using RT-qPCR

cDNA synthesis was performed using 2 µg of total RNA, random and oligo-dT_{18} primers (500 ng each), 1 mM dNTPs, 200 U of M-MLV reverse transcriptase, and 20 U of RNasin ribonuclease inhibitor, all purchased from Promega.

To measure gene expression and allele frequencies, qPCR was performed in triplicates using the 7500-Fast cycler (Applied Biosystems) using the cycling conditions provided in **Table 12**. CDNA corresponding to 40 ng total RNA together with 200 nM forward and reverse primers as well as iTaq Universal SYBR-green Supermix (Bio-Rad) were used per 20 µl reaction.

Table 12: qPCR conditions

Temperature Settings				
Cycles	**Step**	**Temp. (°C)**	**Time**	**Ramp Rate**
1	Holding Stage	50	2 min	100 %
1	Holding Stage	95	10 min	100 %
40x	Cycling Stage	95	15 sec	100 %
	Cycling Stage	60	1 min	100 %
1	Melt curve Stage	95	15 sec	100 %
1	Melt curve Stage	60	1 min	100 %
1	Melt curve Stage	95	30 sec	1 %
1	Melt curve Stage	60	15 sec	100 %

PCR efficiencies were assessed by standard dilution series (1:2, 5 steps in duplicates) and specificity by melt curve analysis using the 7500 v2.3 software, which was also used for quantification along with the Data Assist v3.01 software from Applied Biosystems. Quality criteria (PCR efficiency of 90-100 %, negative no-template control, single peak of > 70°C in the melt curve analysis) were met for all conducted qPCRs. For relative quantification, gene expression was normalized to *GUSB* and *ABL1* expression and a control sample using the $2^{-\Delta\Delta Ct}$ method. Primers are listed in **Supplementary Table 2**.

Pluripotency of hiPSCs was verified via the expression of pluripotency factors *OCT4*, *SOX2* and *NANOG* using the primers OCT4_F1, OCT4_R1, SOX2_F1, SOX2_R1, NANOG_F1 and NANOG_R1 (**Supplementary Table 2**).

To validate the *KMT2A::MLLT3* knock-in and translocation hiPSC clones, qPCR to determine *KMT2A::MLLT3* knock-in/translocation allele frequencies and *KMT2A::MLLT3* mRNA levels was employed. Genomic *KMT2A* WT and knock-in alleles were analyzed using primers KMT2Ain9-F2 and KMT2Ain9-R1 for *KMT2A* and primers KMT2Ain9-F2 and MLLT3in5-R1 for *KMT2A::MLLT3* (**Supplementary Table 2**). Also, the expression of the reciprocal *MLLT3::KMT2A* translocation using primers MLLT3in5-F1 and KMT2Ain9-R1 was validated (**Supplementary Table 2**). For measuring mRNA levels of *KMT2A::MLLT3,* primers

KMT2Aex9-F1 and MLLT3ex6-R1, and for *KMT2A*, primers KMT2Aex9-F2 and KMT2Aex11-R1 were used (**Supplementary Table 2**).

After hematopoietic differentiation of both *KMT2A::MLLT3* knock-in and *NRAS* G12D clones, expression of the reported KMT2A::MLLT3 target *HOXA9* using primers HOXA9_ex1_F1 and HOXA9_ex2_R1 was assessed (**Supplementary Table 2**). In addition, the expression of the canonical *NRAS* targets *DUSP6* and *SPRY2* was determined. For this purpose, primers DUSP6_ex1_2_F1, DUSP6_ex2_R1, SPRY2_ex1_F2 and SPRY2_ex2_R1 were used (**Supplementary Table 2**).

2.18 Gel electrophoresis

Gel electrophoresis of PCR products and digests was conducted on 1 % agarose gels in 0.5x Tris-borate-EDTA (TBE)-buffer at 120 V and 50 mA for 30 minutes. Agarose gels were prepared by dissolving 1 g of agarose (VWR) in 100 ml 0.5x TBE-buffer in a microwave. After the cooling of the solution, 5 µl of ethidium bromide solution (10 mg/ml) (Sigma Aldrich) was added and cast into a gel stand. DNA fragments were loaded with a 6x DNA loading dye (6.2 Buffer recipes in Supplements) in a ratio of 1:6. Additionally, 5 µl of Peqlab DNA Ladder (VWR) was loaded to determine band sizes.

2.19 Protein isolation and Western blot analysis

Protein extraction was performed using a high salt lysis buffer [20 mM Tris-HCL pH 7.5, 400 mM NaCl, 0.5 % NP40 and 0.3 % Triton X-100 supplemented with PMSF (200 µM), Aprotinin, Leupeptin, Pepstatin (each 1 µg/ml) and benzonase nuclease (150 units/ml)], all from Sigma-Aldrich/Merck. Lysates were incubated at 4 °C for 90 minutes, followed by centrifugation at 13 000 x *g* for 30 minutes at 4 °C, and the supernatant was collected and stored at -20 °C until further usage.

Western blotting was performed using standard laboratory protocols. In brief, protein lysates were mixed with sodium dodecyl sulfate (SDS) loading buffer (**6.2** Buffer recipes in Supplements) and incubated at 95 °C for 5 minutes. Subsequently, the lysates and a PageRuler prestained ladder (ThermoFisher Scientific) were subjected to denaturing SDS-Poly Acrylamide Gel Electrophoresis (PAGE) using a 4-15 % acrylamide precast gel (BIO-RAD) and Tris/Glycine/SDS buffer (Bio-Rad). Separated proteins were then tank-blotted with Tris/Glycine blotting buffer onto nitrocellulose membranes (GE Amersham Protran 0.45 µm NC, Merck) and subsequently, equal loading was verified by staining the membranes with Ponceaus S (Sigma-Aldrich/Merck). Blocking was performed using 1x blocking reagent

(Roche) in tris buffered saline (TBS)-buffer, followed by incubation with primary antibody overnight. Membranes were washed and incubated with a secondary DyLight labeled antibody for two hours. Membranes were then scanned on a LI-COR Odyssey. Band signal intensities were quantified using Image Studio Lite 5.2.5 (LI-COR Bioscience).

For detection of KMT2A and HA-tagged KMT2A::MLLT3, the antibodies listed in **Table 13** were used. All antibodies were diluted in TBS-T supplemented with 0.5x blocking reagent (Roche). For detection of the N-terminal part of KMT2A and KMT2A::MLLT3 proteins, an MLL1 rabbit monoclonal antibody (Cell Signaling Technology) and, as secondary antibody, a goat anti-rabbit antibody (ThermoFisher Scientific) were used. To detect the tagged KMT2A::MLLT3 protein, a HA-tag mouse monoclonal antibody (Santa Cruz Biotech) and a goat anti-mouse antibody (ThermoFisher Scientific) as secondary antibody were used.

Table 13: Antibodies used for Western blot

Antibody	Manufacturer	Catalogue Number	(Clone/)lot Number	Dilution Radio
MLL1 Rabbit mAb (Amino-terminal antigen)	Cell Signaling Technology	14689	D2M7U/ 1	1:1000
HA-tag mouse antibody	Santa Cruz Biotech	sc-7392X	F-7/I1608	1:5000
Goat anti-Rabbit IgG (H+L) DyLight 650	ThermoFisher Scientific	84546	ND170165	1:12000
Goat anti-Mouse IgG (H+L) DyLight 800	ThermoFisher Scientific	35521	MG161209	1:12000

3. Results

3.1 *KMT2A::MLLT3* knock-in into hiPSCs and validation of *KMT2A::MLLT3* knock-in clones

Leveraging the advanced CRISPR/Cas9 technology, we created genetically modified hiPSC cell lines harboring a *KMT2A::MLLT3* fusion, with the aim of exploring its impact on *in vitro* hematopoietic differentiation. To introduce the *KMT2A::MLLT3* fusion into hiPSCs without off-targets, we utilized an established *in trans* paired nicking approach (Chen et al., 2017; Fortschegger et al., 2021). In brief, we knocked-in exons 6-11 of the *MLLT3* gene at the intron

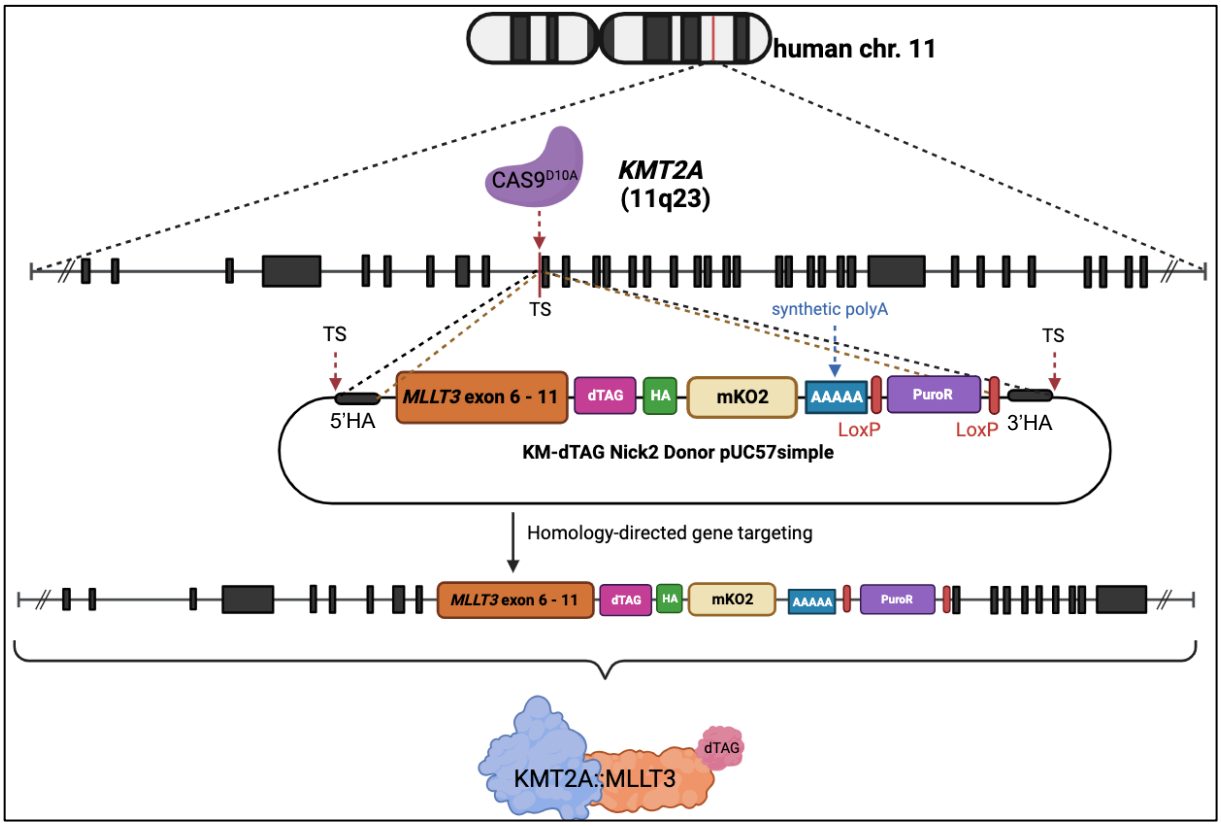


Figure 14: *KMT2A::MLLT3* knock-in strategy using *in trans* paired nicking. The Cas9^{D10A} enzyme, guided by a corresponding crRNA nicks the target site at the intron 9/exon 10 boundary of *KMT2A* and twice at the donor vector outside of the homology arms (HA). Homology-directed DNA repair in hiPSCs leads to the insertion of *MLLT3* exons 6-11, additional degradation (dTAG) and hemagglutinin (HA) tags, the fluorescent reporter mKusabira-Orange2 (mKO2), and a floxed puromycin resistance cassette (PuroR) into exon 10 of *KMT2A*. Target site (TS). Created with BioRender.com.

9 to exon 10 boundary of *KMT2A*, as depicted in **Figure 14**. By the addition of a C-terminal dTAG we aimed to conditionally modulate the expression of the fusion protein through targeted proteolytic degradation upon treatment with a heterobifunctional compound such as dTAG-13 (Nabet et al., 2020). In addition, we inserted an HA tag for specific detection, an IRES driven

fluorescent reporter, mKO2, to assess locus activity, and a synthetic poly-adenylation signal to induce transcriptional termination and avoid splice-out of the insert. Furthermore, we also introduced a PuroR for selection and enrichment of knock-in positive cells, which was flanked by two loxP sites (floxed) to allow subsequent removal of the potentially interfering resistance gene by transient Cre recombinase expression (Araki et al., 1997).

Utilizing this approach, we knocked in exons 6-11 of the *MLLT3* gene into the *KMT2A* locus of hiPSCs. Immediately after nucleofection we treated the hiPSCs with dTAG-13 to prevent any unintended effects of fusion protein expression at the pluripotent stage that might later interfere with hematopoietic differentiation. To genotype and analyze the generated bulk cells and subsequently derived single cell clones, we designed a series of PCRs depicted in **Figure 15 A**, which were performed on corresponding extracted cell DNA. Bulk cells underwent dual puromycin selection, followed by Cre transfection to excise the PuroR cassette, categorizing the former as KM KI floxed and the latter as KM KI delta bulk. Our PCR analysis confirmed the presence of the *KMT2A::MLLT3* knock-in in both examined cell bulks, with the *KMT2A* WT allele being intact, however, a fraction of the KM KI delta bulk retained the PuroR indicating inefficient Cre recombination (**Figure 15 B**). Nevertheless, we picked and seeded 192 single cell clones derived from the KM KI delta bulk. Next, we performed two PCR screens utilizing the 3'HA PCR and examined 192 clones in total (**Figure 15 C** and **D**, only second screen is shown).

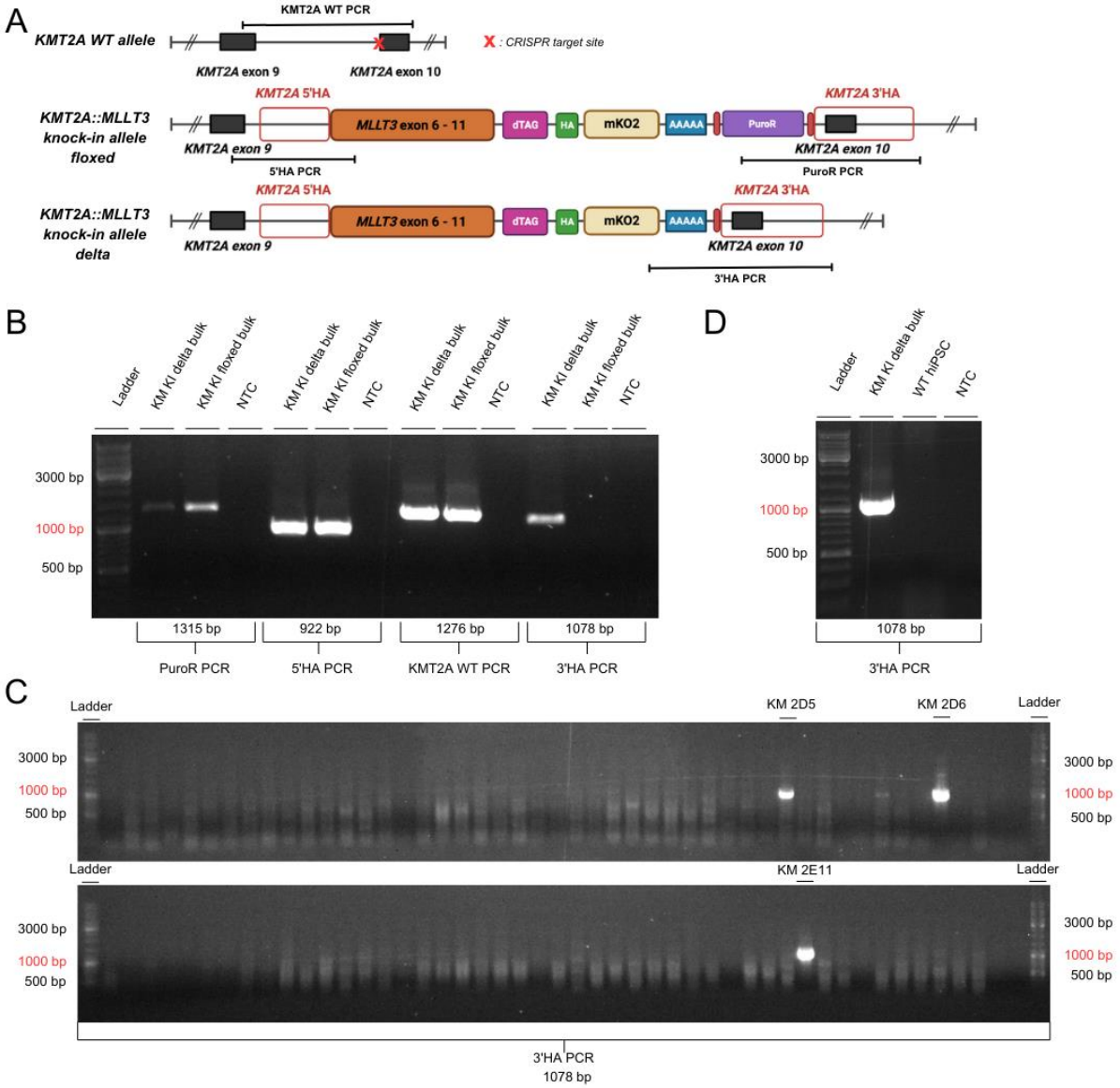


Figure 15: PCR genotyping and screening of KM knock-in bulk cells and KM single cell knock-in clones. A: Illustration of the *KMT2A* WT, *KMT2A::MLLT3* floxed and delta knock-in alleles with indicated PCR, 5' and 3' HAs, CRISPR target and LoxP sites (annotated by red bars). The PuroR PCR was conducted to detect the presence of the KM KI floxed allele. The 5'HA and 3'HA PCRs were performed to validate the correct insertion of the knock-in. Moreover, the 3'HA PCR was employed to screen for *KMT2A::MLLT3* knock-in delta positive clones and to confirm the excision of PuroR. The *KMT2A* WT PCR spans the intron 9 and exon 10 junction of *KMT2A*. **B:** Gel picture of indicated PCRs with product lengths of transfected and puromycin selected bulk cells (KM KI floxed bulk) as well as Cre transfected bulk cells (KM KI delta bulk). **C:** Gel pictures from the second PCR screen of 96 clones using the 3'HA PCR. Clones KM 2D5, 2D6 and 2E11 show presence of the delta allele. **D:** Gel picture of control conditions for the second PCR screen using the 3'HA PCR on KM KI delta bulk (positive control) and WT hiPSC (negative control) DNA. *KMT2A::MLLT3* (KM), degradation tag (dTAG), mKusabira-Orange2 (mKO2), poly A (AAAAA), puromycin resistance cassette (PuroR), homology arm (HA), clustered regularly interspaced short palindromic repeats (CRISPR), knock-in (KI), no template control (NTC), wild-type (WT).

Three clones termed KM 2D5, 2D6 and KM 2E11 were positive for the *KMT2A::MLLT3* knock-in allele delta (**Figure 15 C** and **D**). Therefore, we expanded the replicate cells from the second plate of these clones, extracted DNA and performed further genotyping PCRs (**Figure 16**).

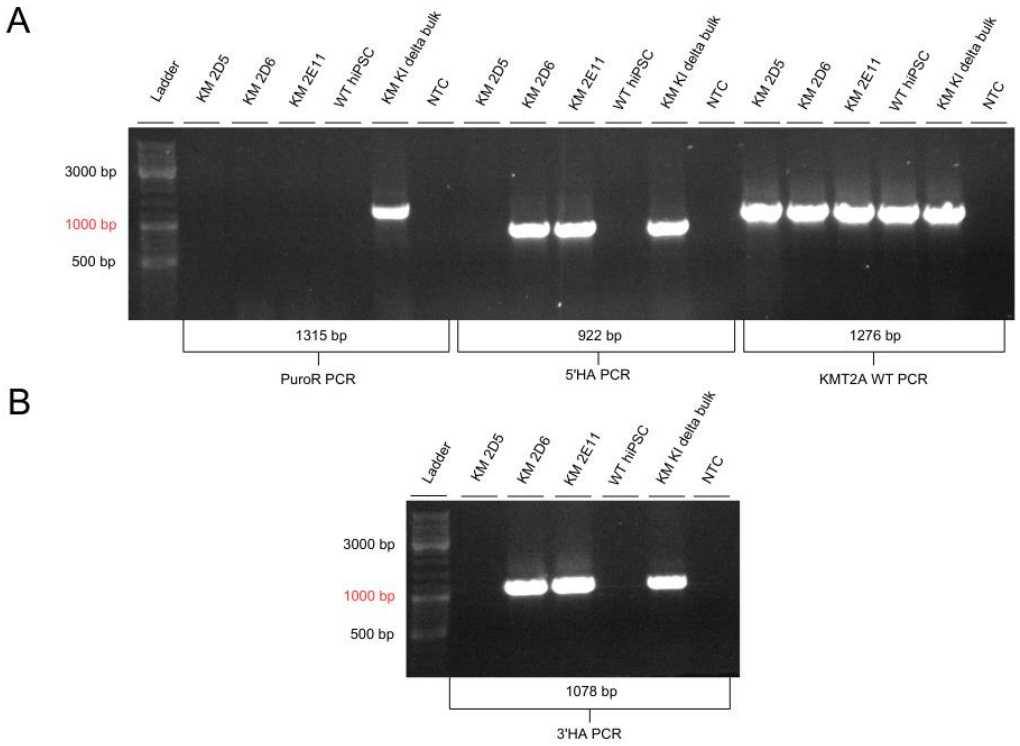


Figure 16: PCR genotyping of expanded KM single cell clones. WT hiPSC and KM KI delta bulk cell DNA were used as negative and positive controls, respectively. **A:** PuroR PCR negativity confirms successful PuroR excision in all clones. Clones KM 2D6 and 2E11 exhibit the expected 5'HA PCR product, while KM 2D5 lacks it. *KMT2A* WT PCR confirms the presence of the *KMT2A* WT allele in all clones. **B:** 3'HA PCR genotyping shows the *KMT2A::MLLT3* knock-in allele delta presence in clones KM 2D6 and KM 2E11. Conversely, for replicate clone KM 2D5, both 5'HA and 3'HA PCR analyses indicate the absence of the knock-in. *KMT2A::MLLT3* (KM), puromycin resistance cassette (PuroR), homology arm (HA), knock-in (KI), no template control (NTC), wild-type (WT).

Unfortunately, clone KM 2D5 did not retain the *KMT2A::MLLT3* knock-in allele delta after cell expansion, potentially attributable to an initially mixed composition and outgrowth of WT cells (**Figure 16 A** and **B**). However, clones KM 2D6 and 2E11 proved to be heterozygous, as they contained a *KMT2A::MLLT3* knock-in delta allele and retained an intact *KMT2A* WT allele (**Figure 16 A** and **B**). Furthermore, both clones tested negative for the PuroR, confirming successful Cre recombination (**Figure 16 A**). In addition, another knock-in clone termed KM E12 has already been generated, validated, and genotyped prior to the start of this thesis by my supervisor, Klaus Fortschegger.

We then sequenced the 5' and 3'HA as well as the *KMT2A* WT PCR product to validate the *KMT2A::MLLT3* knock-in and to control for mutations in the *KMT2A* WT allele (**Figure 17**). As seen by sequencing of the 5' and 3'HA of the *KMT2A::MLLT3* knock-in allele delta, we confirmed that clones KM 2D6 and KM 2E11 indeed harbor the desired knock-in (**Figure 17**

A, **B** and **C**). Moreover, no mutations were detected in the *KMT2A* WT allele, excluding CRISPR/Cas9-induced off-target insertions/deletions in this region (**Figure 17 D** and **E**).

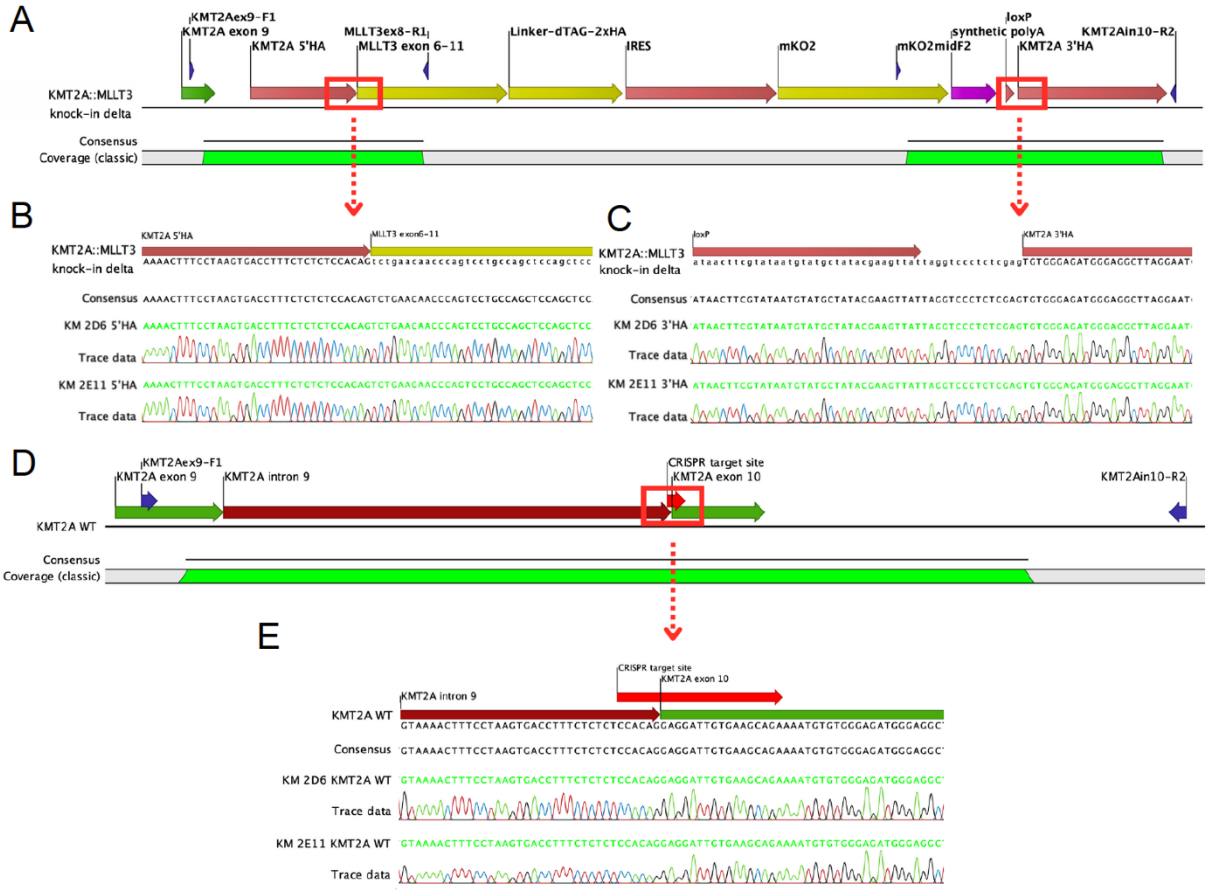


Figure 17: Sequencing results of the *KMT2A* WT and 5'HA and 3'HA PCRs of the *KMT2A::MLLT3* knock-in delta clones KM 2D6 and KM 2E11. Depicted sequences were obtained by Sanger sequencing of PCR products. Alignment to reference overviews show respective coverages, used primer pairs and CRISPR target sites. **A:** Overview of the intended *KMT2A::MLLT3* knock-in delta allele used as reference. Red boxes highlight regions magnified in B and C. **B:** Sequencing chromatogram showing the junction between the 5'HA and the *MLLT3* exon 6-11 insert. **C:** Sequencing chromatograms showing the region between the remaining loxP site and the 3'HA. **D:** Overview of the *KMT2A* WT allele used as alignment reference. Red square highlights region magnified in E. **E:** Sequencing chromatograms demonstrating lack of CRISPR/Cas9-induced off-target effects at the *KMT2A* intron 9 and exon 10 junction. *KMT2A::MLLT3* (KM), degradation tag (dTAG), internal ribosome entry site (IRES), mKusabira-Orange2 (mKO2), puromycin resistance cassette (PuroR), homology arm (HA), clustered regularly interspaced short palindromic repeats (CRISPR), wild-type (WT).

Next, the expression of *KMT2A::MLLT3* and *KMT2A* WT transcripts for each *KMT2A::MLLT3* knock-in clone was assessed by RT-qPCR. As shown in **Figure 18**, the fusion transcript was expressed in both *KMT2A::MLLT3* knock-in clonal hiPSC lines, with clone KM 2E11 expressing slightly higher levels of *KMT2A* WT than *KMT2A::MLLT3* transcripts.

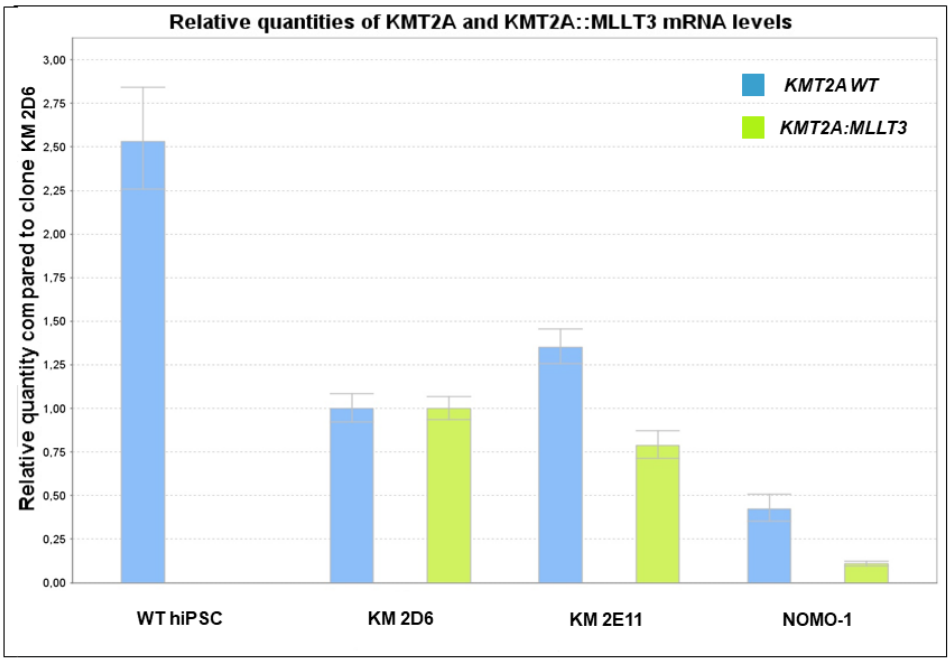


Figure 18: Relative quantification of *KMT2A* WT and *KMT2A::MLLT3* mRNA levels in single cell clones. CDNA was synthesized using RNA isolated from *KMT2A::MLLT3* knock-in clones KM 2E11, KM 2D6 and WT hiPSCs as well as NOMO-1 cells. Relative quantities were normalized to clone KM 2D6 and housekeeping genes *ABL1* and *GUSB* using the $2^{-\Delta\Delta Ct}$ method. NOMO-1 is a patient-derived cell line harboring a *KMT2A::MLLT3* fusion. *KMT2A::MLLT3* (KM), wild-type (WT).

To further assess clone purity, we measured the relative quantities of *KMT2A::MLLT3* knock-in and *KMT2A* WT alleles in genomic DNA by qPCR. Since we opted for a gene editing approach in which we inserted the *MLLT3* part into one of two *KMT2A* alleles, we considered clones that have equal quantities of WT and knock-in alleles as pure heterozygotes. Based on our qPCR analysis, we confirmed the purity of clones KM E12 and KM 2D6, as seen by the equal relative quantity of *KMT2A WT* and *KMT2A::MLLT3* genomic PCR products (**Figure 19**). For clone KM 2E11, higher amounts of *KMT2A* WT PCR product were detected, which is consistent with increased *KMT2A* WT transcript observed in RT-qPCR, and indicative of contamination with WT cells.

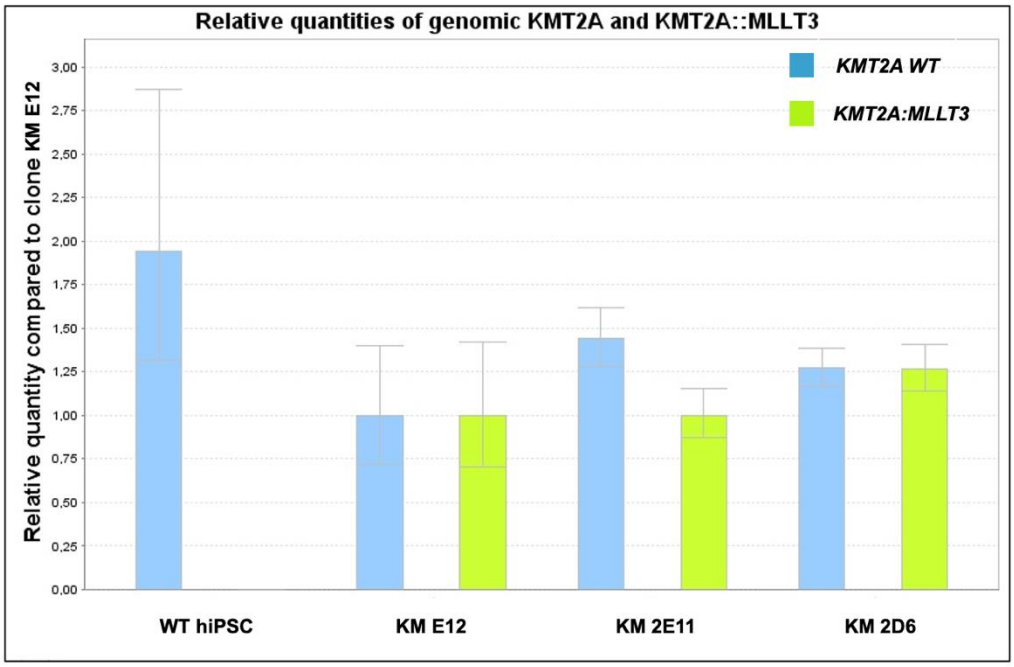


Figure 19: Relative quantity of genomic *KMT2A* WT and *KMT2A::MLLT3.* DNA was isolated from *KMT2A::MLLT3* knock-in clones KM E12, KM 2E11 and KM 2D6 and WT hiPSCs and subjected to qPCR using primers specific for the *KMT2*A WT and *KMT2A::MLLT3* alleles. Relative quantities were normalized to clone KM E12 and calculated via the $2^{-\Delta\Delta Ct}$ method. *KMT2A::MLLT3 (*KM*),* wild-type (WT).

To further validate the purity of the *KMT2A::MLLT3* knock-in clones, we measured mKO2-fluorescence of each clone using flow cytometry. We observed *mKO2* expression in over 95% of cells for clones KM E12 and KM 2D6 but in only 78 % for clone KM 2E11, confirming that this clone was indeed impure (**Figure 20**). Therefore, we flow sorted clone KM 2E11 for mKO2-positive cells, achieving a purity of ~98 % (**Figure 20**).

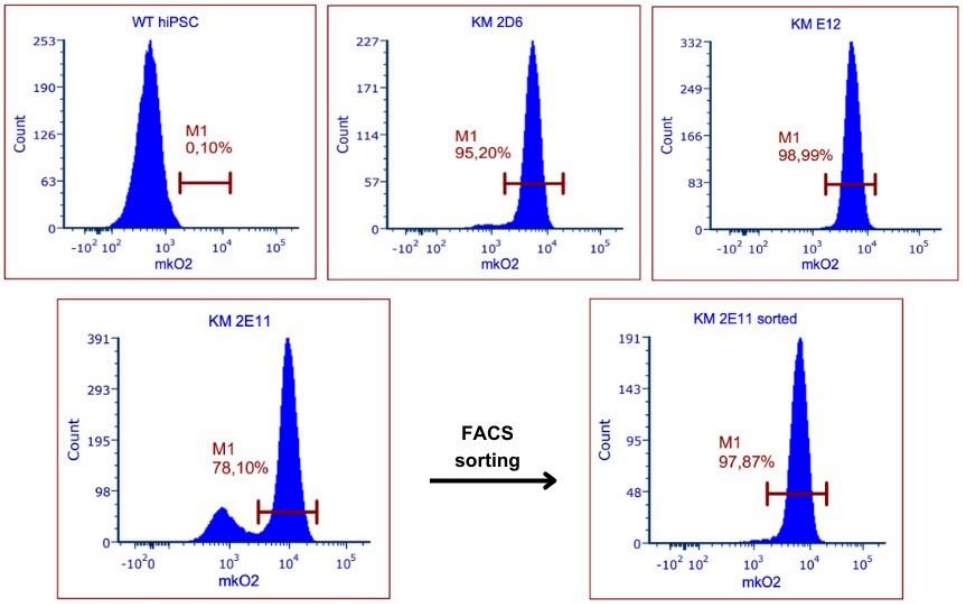


Figure 20: Flow cytometry analysis and FACS sorting of *KMT2A::MLLT3* knock-in clones based on mKO2 expression. The red bar termed M1 shows the expected orange fluorescence for mKO2-positive cells in histograms. Intact cells were analyzed for mKO2 fluorescence in the phycoerythrin (561 nm excitation, 586 nm emission) channel using FCS Express 7. FACS sorting was performed based on mKO2 expression. *KMT2A::MLLT3* (KM), mKusabira-Orange2 (mKO2). The flow cytometry analysis was performed by my supervisor, Klaus Fortschegger.

To detect the expression of the fusion protein in our *KMT2A::MLLT3* knock-in clones, we performed Western blot analysis. In the protein lysates of clone KM E12, we could indeed detect expression of the KMT2A::MLLT3 fusion protein, and in line with the genome-editing of one allele, the KMT2A WT protein levels in clone KM E12 showed approximately a fifty percent reduction compared to WT hiPSCs. (**Figure 21**). Moreover, dTAG-13 treatment caused almost complete depletion of the KMT2A::MLLT3 fusion protein. However, we could detect neither KMT2A WT nor KMT2A::MLLT3 proteins in further Western blot experiments.

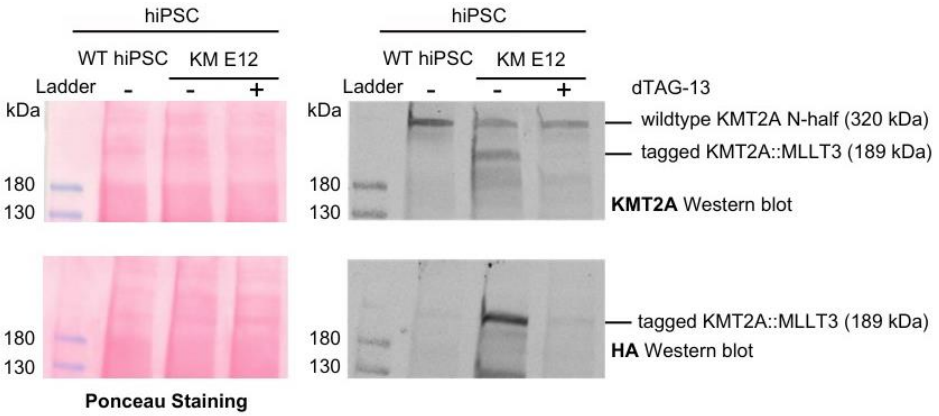


Figure 21: Western blot analysis of KMT2A::MLLT3 fusion protein expression. Left panel shows Ponceau staining images to verify equal loading. Right panel shows images of Western blot analysis performed with lysates of untreated (-) or dTAG-13-treated (+) wild-type hiPSC (WT) and KMT2A::MLLT3 knock-in clone KM E12. KMT2A Western blot analysis (upper right panel) was performed using an antibody specific for the N-terminal part of the KMT2A protein. HA Western blot analysis (lower right panel) was performed using an antibody specific for the C-terminal HA tag. Tagged KMT2A::MLLT3 refers to KMT2A::MLLT3 carrying dTAG and HA-tag. *KMT2A::MLLT3* (KM). Note that this experiment was conducted prior to the start of this thesis by my supervisor, Klaus Fortschegger.

3.2 Introduction of a *NRAS* G12D mutation into hiPSCs and clone validation

After the establishment of hiPSC cell lines containing the *KMT2A::MLLT3* fusion, we generated hiPSC cell lines harboring a secondary *NRAS* G12D mutation. For this purpose, we used the *KMT2A::MLLT3* knock-in clone KM E12, and WT hiPSCs to investigate the effects of this mutation in both genetic backgrounds. Since the *NRAS* G12D mutation is caused by a single point mutation, we first tried to utilize the optimized prime editor system 4 (PE4) with a pegRNA targeting exon 2 of *NRAS* encoding the G12D mutation along with a synonymous NarI cut site to facilitate screening of positive clones (**Figure 22**). This system further utilizes the dominant negative mutant of the DNA mismatch repair (MMR) protein MHL1 called MLH1dn, which reduces MMR and improves DNA editing efficiency (Chen et al., 2021).

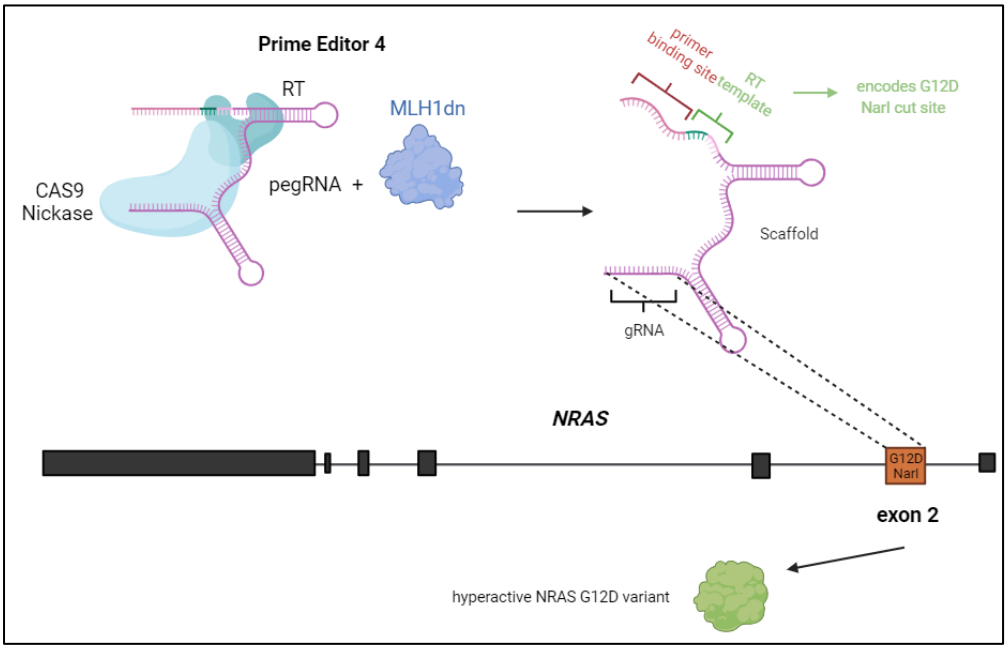


Figure 22: Prime editing strategy for the *NRAS* G12D mutation. Using the Prime Editor 4 system, the *NRAS* G12D mutation and a NarI cut site are encoded in the reverse transcription (RT) template of the pegRNA, which is used as a repair template after successful nicking leading to an incorporation into exon 2 of *NRAS*. MLH1dn enhances gene-editing efficiencies by inhibiting competing DNA mismatch repair.

Although we had prioritized this strategy, we neither detected any base editing in our hiPSC bulk cells by PCR followed by NarI digestion and gel electrophoresis nor by Sanger sequencing of respective PCR products (**Figure 23**).

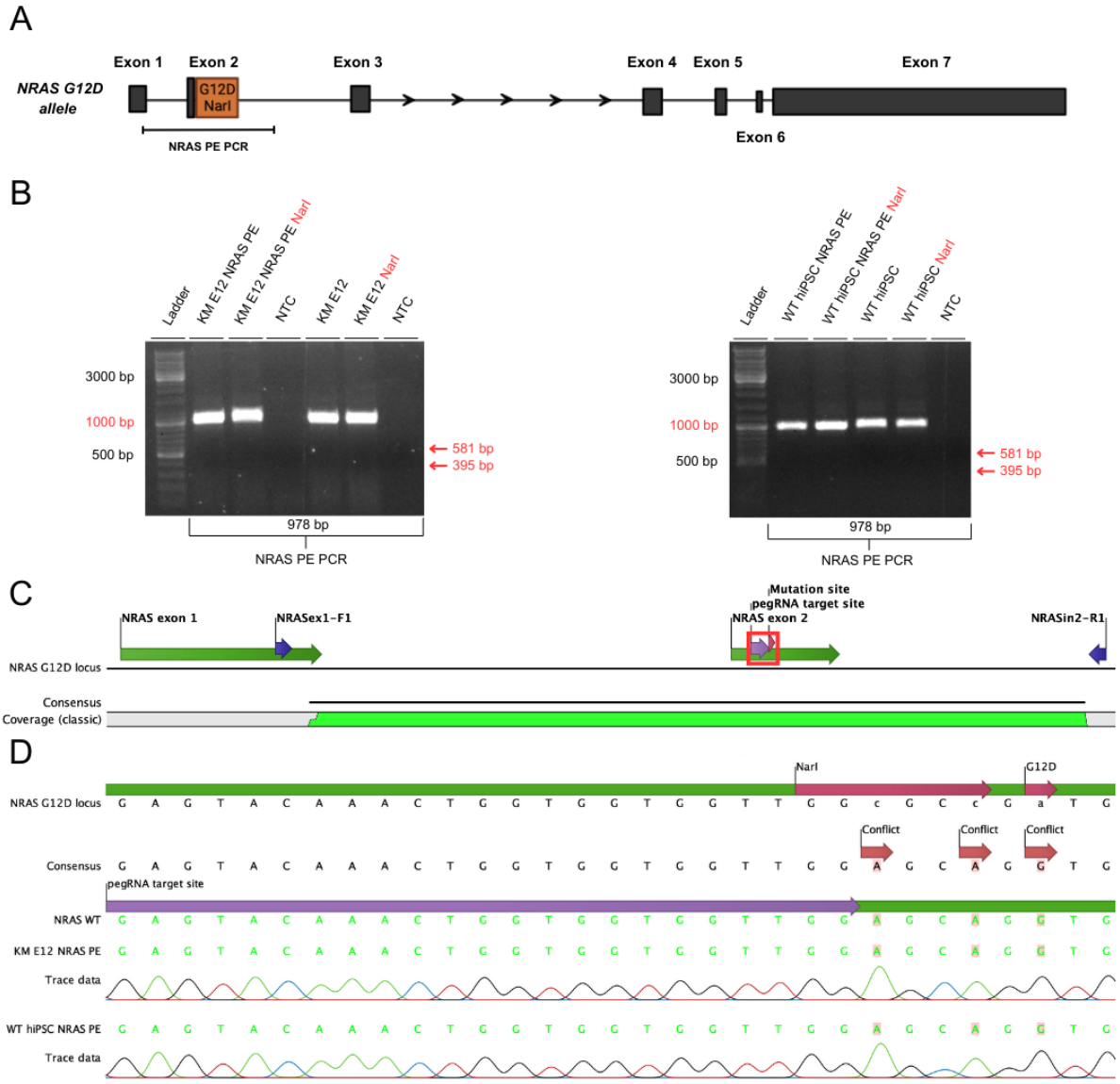


Figure 23: Genotyping of prime-edited bulk cells. Sequences were obtained by Sanger sequencing of PCR products from KM E12 and WT hiPSC *NRAS* PE bulk cell DNA. Sequences were aligned to a reference and the respective coverages, used primer pairs and pegRNA target site are indicated. **A:** Illustration of the intended G12D-mutated *NRAS* allele and the PCR-amplified region (*NRAS* PE PCR). **B:** Gel pictures of the respective PCR performed with bulk cell DNA isolated from transfected KM E12 *NRAS* PE (left panel) and WT hiPSC *NRAS* PE (right panel). Untransfected WT hiPSCs were used as negative control. PCR products were digested with NarI to assess successful prime editing (marked with NarI). Red arrows indicate the expected fragment sizes after NarI digestion. No digestion product was detected in KM E12 cells or WT hiPSCs, indicating inefficient prime editing. **C:** Overview of the intended *NRAS* G12D allele used as an alignment reference. Red box highlights the region magnified in D. **D:** Sequencing chromatograms of KM E12 and WT hiPSC *NRAS* PE bulk cells together with a *NRAS* WT sequence as reference. Successful prime editing would be recognizable by double peaks in the chromatogram. In our analysis, no such peaks were detected, suggesting unsuccessful prime editing. Prime editor (PE), no template control (NTC), *KMT2A::MLLT3* (KM), wild-type (WT).

Therefore, we resorted to an alternative approach, in which we tried to introduce the *NRAS* G12D mutation by *in trans* paired nicking knock-in (**Figure 24**).

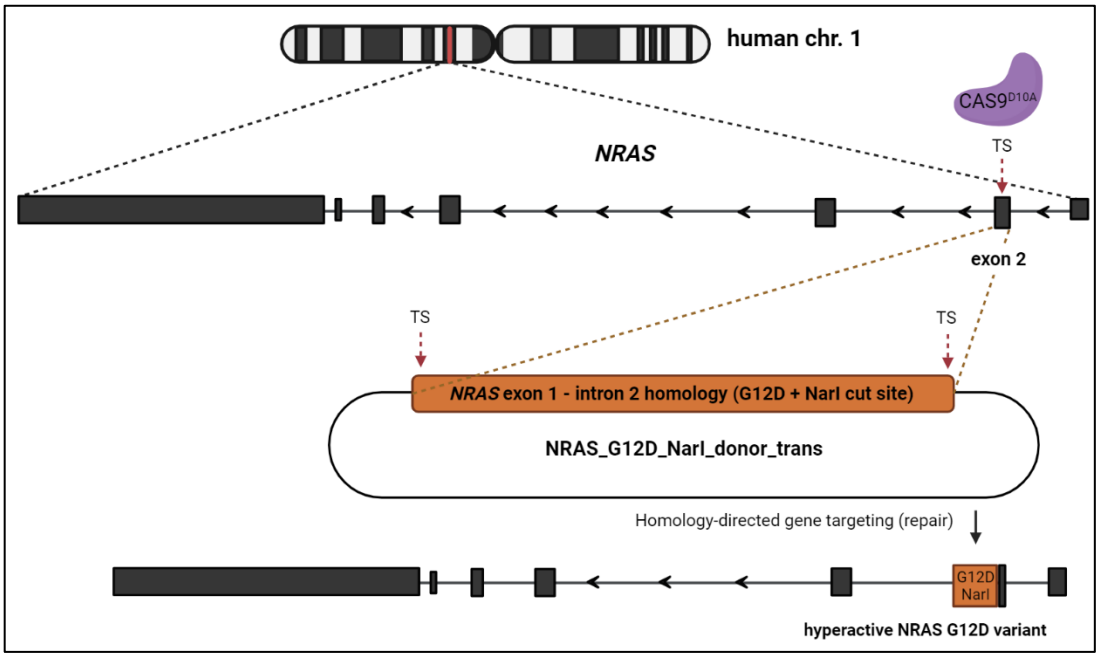


Figure 24: *In trans* paired nicking strategy for the introduction of an *NRAS* G12D mutation. The Cas9^{D10A} nickase induces nicks at the target sites (TS) inside *NRAS* exon 2 and the HDR donor construct. Through this approach, the *NRAS* homology containing both the *NRAS* G12D mutation and NarI cut site is used as an HDR repair template, thereby introducing mutations into the *NRAS* locus. The introduction of the NarI cut site allows for screening of gene edited cells. Homology directed gene repair (HDR).

However, our hiPSC bulk cell analysis of transfected clone KM E12 by PCR followed by NarI digestion and Sanger sequencing again showed inefficient gene editing (**Figure 25**) and this approach was not further followed-up.

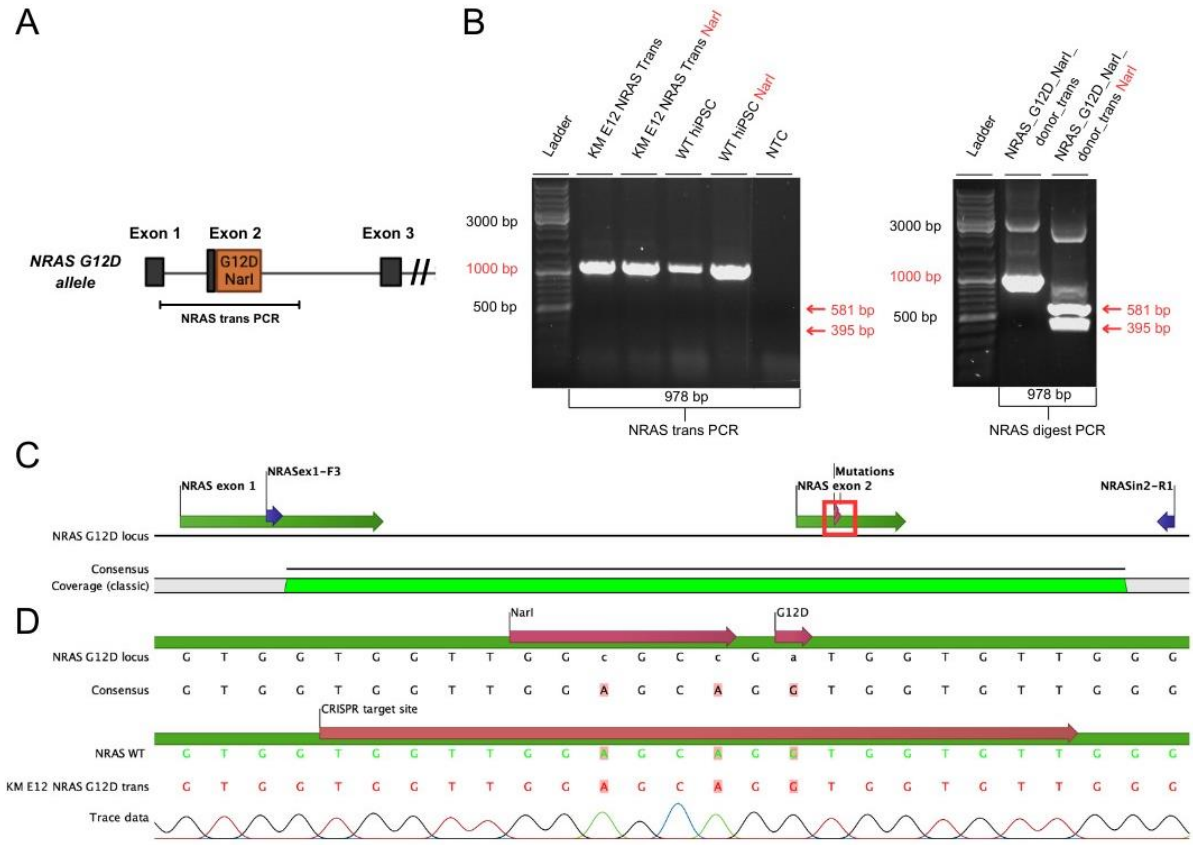


Figure 25: Genotyping of *in trans* paired nicking bulk cells. Sequences were obtained by Sanger sequencing of PCR products from KM E12 *NRAS* trans bulk cell DNA. Sequences were aligned to a respective reference and the coverages, used primer pairs and CRISPR target site are indicated. **A:** Illustration of the intended G12D-mutated *NRAS* allele and the PCR amplified region (*NRAS* trans PCR). The respective PCR was used to amplify *NRAS* exon 2 harboring the intended G12D mutation. **B:** Gel picture of the *NRAS* trans PCR performed using DNA from transfected KM E12 *NRAS* trans bulk cells (left). Additionally, the *NRAS* digest PCR was performed as a positive control on NRAS_ G12D_NarI_donor_trans plasmid DNA, which is specific for the *NRAS* homology inside the plasmid (right). Untransfected WT hiPSCs were used as negative control. Digested PCR products are marked with NarI. Red arrows indicate the expected fragment sizes after NarI digestion. **C:** Overview of the intended *NRAS* G12D allele used as a reference sequence. The red box highlights the region magnified in D. **D:** Sequencing chromatogram of KM E12 *NRAS* trans bulk cells and a *NRAS* WT sequence used as an additional reference. The analysis revealed the absence of double peaks, indicating inefficient gene editing. No template control (NTC), clustered regularly interspaced short palindromic repeats (CRISPR), *KMT2A::MLLT3* (KM), wild-type (WT).

Therefore, we designed a selection-based approach similar to the *KMT2A::MLLT3* knock-in strategy. In brief, we aimed to insert exons 2-7 of *NRAS* harboring the G12D mutation, a polyadenylation site and a floxed PuroR resistance cassette into exon 2 of *NRAS* (**Figure 26**).

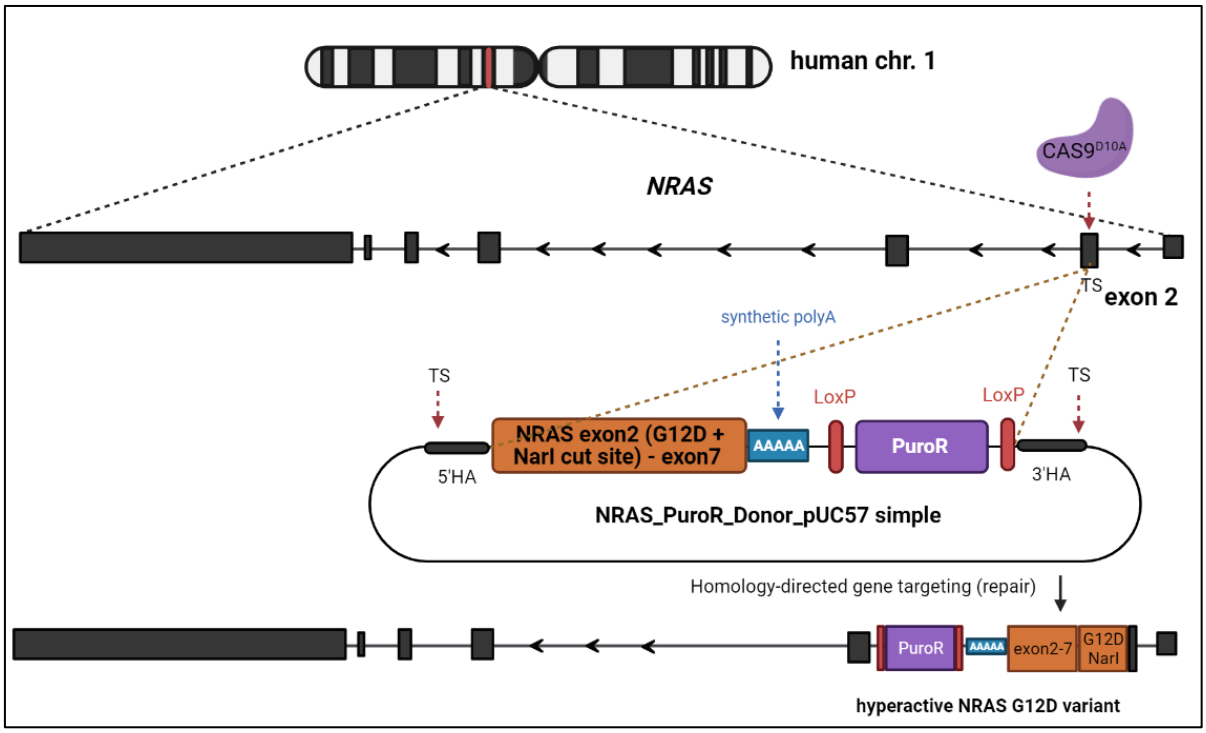


Figure 26: Gene editing strategy for *NRAS* G12D knock-in including a PuroR selection cassette. Similar to the *KMT2A::MLLT3 in trans* paired nicking approach, we utilized the Cas9^{D10A} nickase to induce nicks at target sites (TS) inside *NRAS* exon 2 and outside the HDR donor homology arms (HA) leading to the knock-in of the insert encompassing *NRAS* exons 2-7 with the *NRAS* G12D mutation, a polyadenylation signal (AAAAA) and a loxP-flanked puromycin resistance cassette (PuroR). Homology directed gene repair (HDR).

After cloning the respective constructs, clone KM E12 and WT hiPSCs were nucleofected followed by two pulses of puromycin selection and Cre transfection. Cell bulks were genotyped via a series of PCRs depicted in **Figure 27 A** before and after Cre transfection. The 5'HA and 3'HA were detected in both the KM E12 and WT hiPSC NRAS G12D floxed and delta bulk cells, while an *NRAS* WT allele was retained confirming the presence of the *NRAS* G12D knock-in (**Figure 27 B** and **C**). Some cells again retained the PuroR cassette in the KM E12 *NRAS* G12D delta cell bulk. The stronger PCR band intensity compared to the WT hiPSC *NRAS* G12D cell bulk indicated a higher percentage of gene-edited cells, also explaining the detection of cells with residual PuroR (**Figure 27 C**).

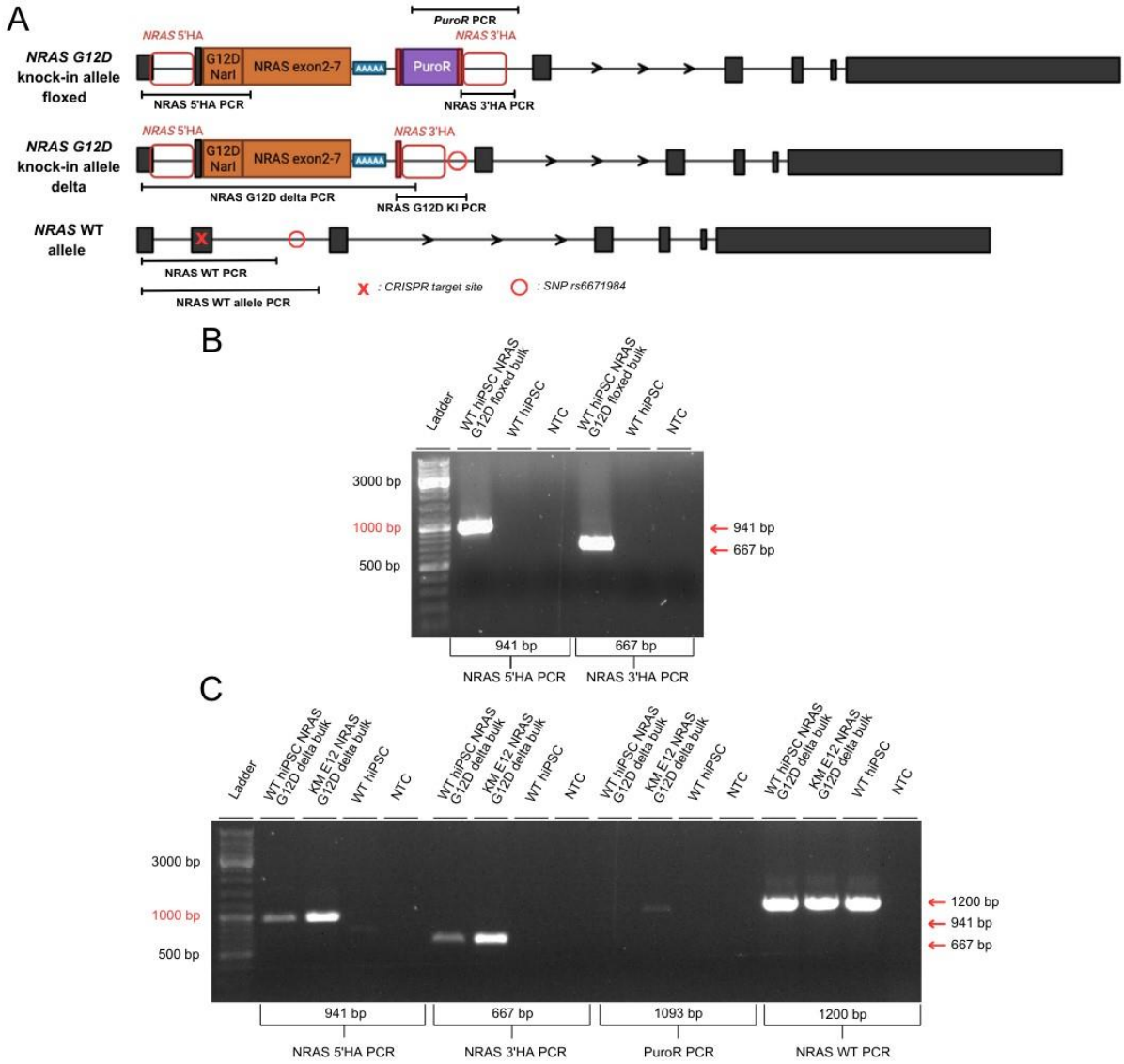


Figure 27: Genotyping of NRAS G12D knock-in bulk cells. A: Overview of the *NRAS* G12D knock-in floxed/delta and *NRAS* WT alleles with PCR amplified regions and indicated 5' and 3'HA. The CRISPR target site and a SNP are also shown. LoxP sites are denoted with red bars. The *NRAS* 5'HA and 3'HA PCRs amplified the HA regions of both the delta and floxed allele. Presence of PuroR was assessed by the PuroR PCR. The *NRAS* WT PCR amplifies *NRAS* exon 2 to intron 2 of the *NRAS* WT allele. To screen for gene edited clones, the *NRAS* G12D delta PCR was used, which amplifies the full knock-in insert on the *NRAS* G12D knock-in allele delta. The genomic region containing the SNP rs6671984, used for assessing clone purity, was amplified by the *NRAS* G12D KI PCR (*NRAS* G12D knock-in) and the NRAS WT PCR. **B:** Gel picture of nucleofected and puromycin-selected WT hiPSC *NRAS* G12D floxed bulk cells. **C:** Gel picture of Cre transfected delta bulk cells. Homology arm (HA), poly A (AAAA), puromycin resistance cassette (PuroR), knock-in (KI), single nucleotide polymorphism (SNP), *KMT2A::MLLT3* (KM). Red arrows indicate expected band sizes.

After introduction of the *NRAS* G12D mutation into WT hiPSC and *KMT2A::MLLT3* knock-in clone KM E12 using *in trans* paired nicking, selection and Cre recombination, we seeded and screened a total of 96 single cell clones using a PCR specific for the *NRAS* G12D knock-in

delta allele. We detected the *NRAS* G12D knock-in in ten KM E12-derived single cell clones termed B1, D1, B3, A4, E4, A5, B5, B6, B9 and F12 and two WT hiPSC clones termed H4 and E9 (**Figure 28 A, B** and **C**). We also detected weak bands in other KM E12 NRAS G12D knock-in clones but did not expand them since the weak band intensities suggested contamination with WT cells (**Figure 28 A**).

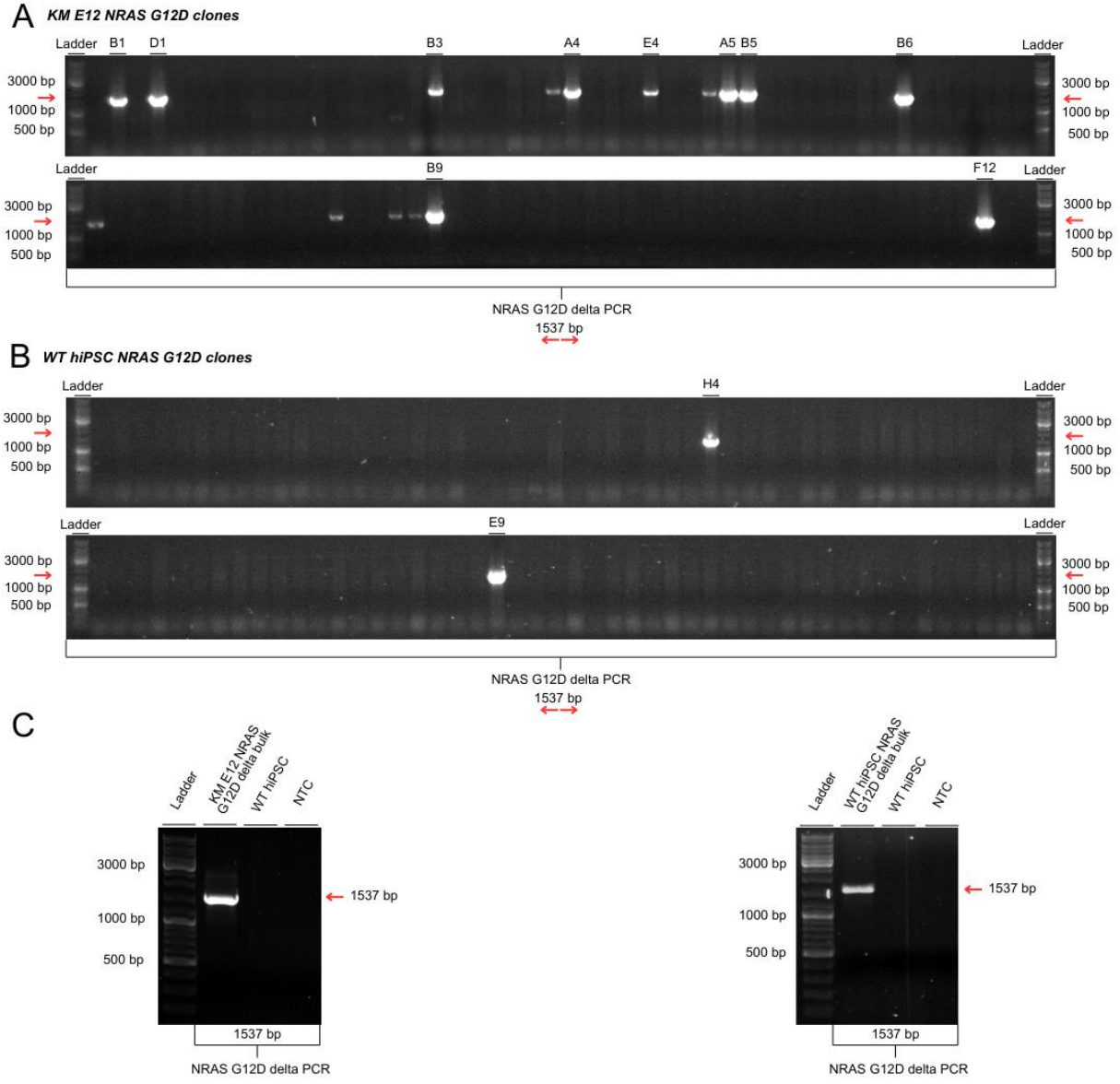


Figure 28: PCR screening of NRAS G12D knock-in clones. A: Gel pictures of the PCR screen of 96 KM E12 NRAS G12D clones using the *NRAS* G12D delta PCR. Clones B1, D1, B3, A4, E4, A5, B5, B6, B9 and F12 show the expected PCR product, suggesting the presence of the *NRAS* G12D knock-in delta allele. **B:** Gel pictures of the PCR screen of 96 WT hiPSC NRAS G12D clones. Clones H4 and E9 show the expected PCR product, suggesting the presence of the *NRAS* G12D knock-in delta allele. **C:** Gel picture of controls for the PCR screens using DNA of the respective delta cell bulk (positive control) and WT hiPSC (negative control). No template control (NTC), *KMT2A::MLLT3* (KM), wild-type (WT). Red arrows indicate expected band sizes.

We then expanded these positive clones and extracted DNA for further genotyping and clone selection. We performed *NRAS* G12D KI and *NRAS* WT allele PCRs to amplify a region containing the informative SNP rs6671984, which was used to assess clone purity (**Figure 29 A** and **B**). Unfortunately, KM E12 NRAS G12D knock-in replicate clones B3 and E4 were negative for the *NRAS* G12D knock-in allele, suggesting the presence of WT cells in these clones which outgrew the gene-edited cells. Therefore, we excluded these clones from further analyses.

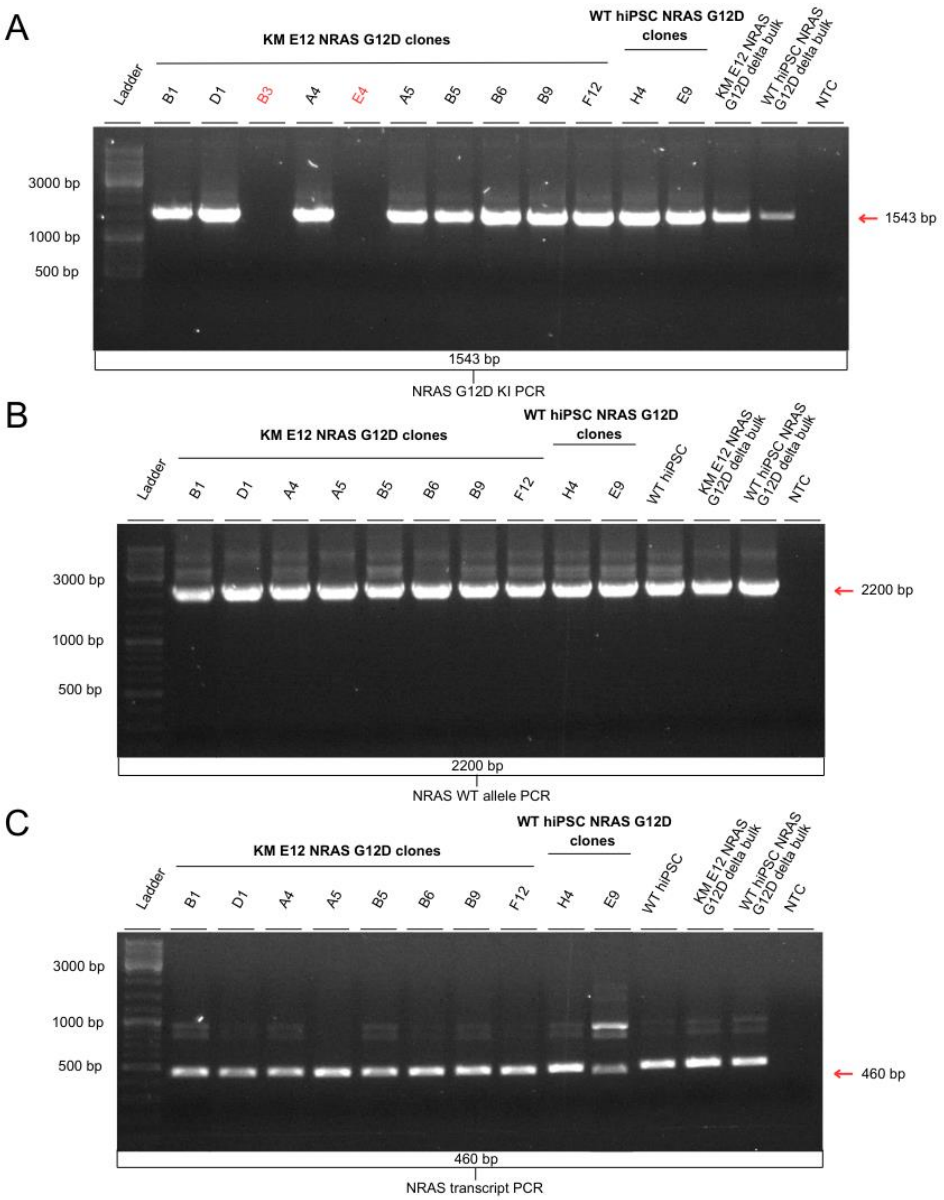


Figure 29: Genotyping of expanded NRAS G12D knock-in clones. A: Gel picture of the NRAS G12D KI PCR. All clones except for B3 and E4 show the expected PCR product. **B:** Gel picture of the *NRAS* WT allele PCR performed on clones and WT hiPSC with all samples showing the respective amplification product. **C:** Gel picture of the NRAS transcript PCR performed on clones and WT hiPSC using cDNA as template, amplifying both *NRAS* G12D-mutated and WT transcripts. The unspecific band (~900-1000 bp) stems from a product containing *NRAS* intron 2 due to genomic DNA contamination. No template control (NTC), *KMT2A::MLLT3* (KM), wild-type (WT). Red arrows indicate expected band sizes.

Next, we performed RNA extraction to synthesize cDNA via reverse transcription. Subsequently, the *NRAS* G12D mutation site was amplified using the *NRAS* transcript PCR (**Figure 29 C**). We Sanger sequenced the PCR product to validate expression of the *NRAS* G12D variant at the hiPSC stage. Sequencing of the PCR-amplified region revealed a total of

four positive KM E12 NRAS G12D knock-in clones (A5, B1, D1 and F12) and two positive WT hiPSC NRAS G12D knock-in clones (H4 and E9) harboring a pure heterozygous knock-in, as seen by the presence of double peaks in the sequencing chromatogram (**Figure 30 A** and **D**). Sanger sequencing of the genomic *NRAS* WT (**Figure 30 B** and **E**) and *NRAS* G12D KI allele (**Figure 30 C** and **F**) PCR products showed that clones E9, H4, A5, B1, D1 harbored the guanine SNP variant on the WT allele, while clone F12 harbored the adenine variant, with the corresponding variant being present on the knock-in allele. The absence of double peaks confirmed purity of the six generated clones (**Figure 30 B, C, E and F**).

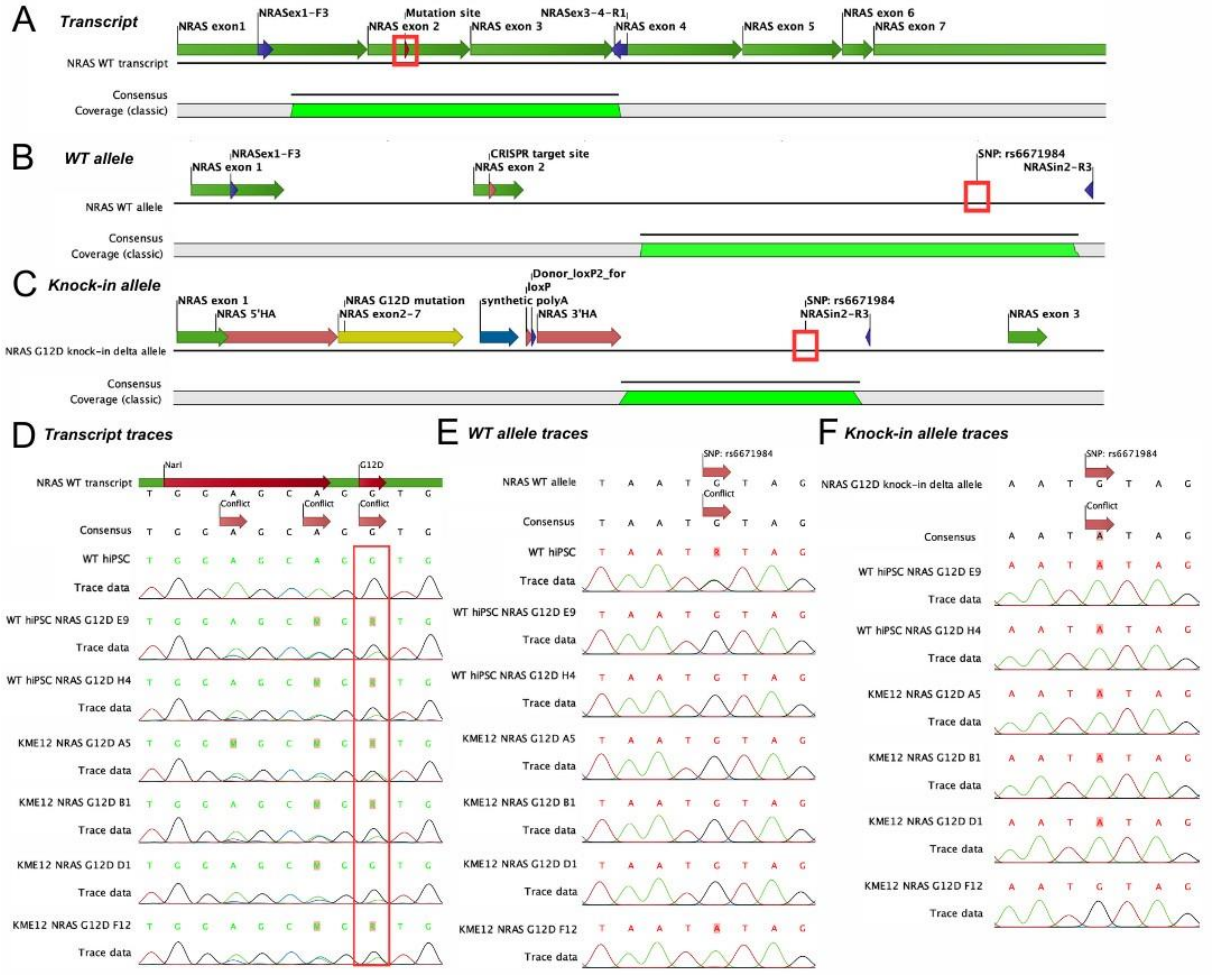


Figure 30: Sequencing results of NRAS G12D knock-in clones. Sequences were obtained by Sanger sequencing of PCR products from indicated NRAS G12D knock-in clones using either cDNA or genomic DNA as templates. Only chromatograms of positive NRAS G12D knock-in clones are shown. Sequences were aligned to respective references and their coverages, used primer pairs, as well as CRISPR target and SNP rs6681984 site are indicated. **A:** *NRAS* WT transcript used as a reference. Red box highlights the region magnified in D. **B:** Overview of the *NRAS* WT allele used as reference sequence. Red box highlights the region magnified in E. **C:** Overview of the *NRAS* G12D knock-in delta allele used as reference sequence. Red box highlights region magnified in F. **D:** Chromatograms of *NRAS* transcript traces of NRAS G12D knock-in clones and WT hiPSCs. Red box highlights the position of amino acid 12, the mutated site. **E:** Chromatograms of the *NRAS* WT allele, showing both SNP rs6671984 variants (G and A) in WT hiPSCs. **F:** Chromatograms of the *NRAS* G12D knock-in delta allele, showing knock-in positive clones. Clustered regularly interspaced short palindromic repeats (CRISPR), homology arm (HA), single nucleotide polymorphism (SNP), *KMT2A::MLLT3* (KM), wild-type (WT).

3.3 Induction of a t(9;11)(p21;q23)/*KMT2A::MLLT3* in hiPSCs and validation of clones

To induce a t(9;11)(p21;q23) translocation resulting in the expression of *KMT2A::MLLT3* fusion transcripts, we used two orthogonal Cas9 enzymes from different species, namely *Staphylococcus aureus* (*Sa*) and *Streptococcus pyogenes* (*Sp*), associating with the *Sa*Cas9 gRNA1 targeting intron 9 of *KMT2A* and the *Sp*Cas9 gRNA2 targeting intron 5 of *MLLT3* (**Figure 31**).

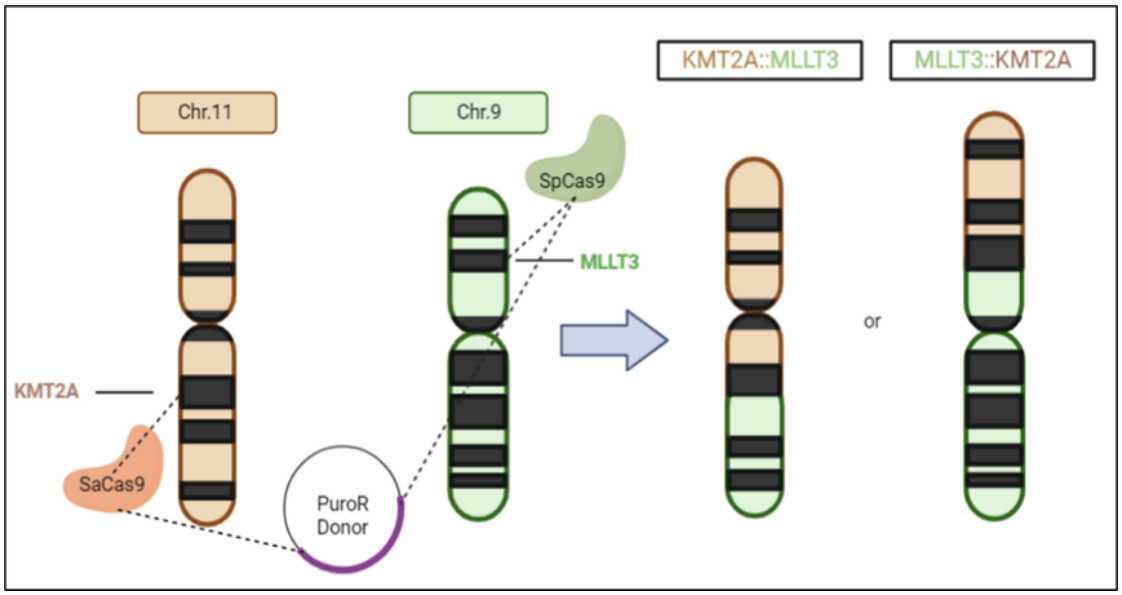


Figure 31: Gene editing strategy for t(9;11)(p21;q23)/*KMT2A::MLLT3* translocation. Using two orthogonal Cas9 enzymes generating double strand breaks in *KMT2A* on chromosome 11 and *MLLT3* on chromosome 9, we tried to induce a translocation that resulted in a *KMT2A::MLLT3* and/or a reciprocal *MLLT3::KMT2A* fusion. Due to low gene editing efficiencies, we later switched to a selection-based approach with a removable puromycin resistance cassette (PuroR). Chromosome (Chr.).

Because of low gene editing efficiency detected by qPCR, which indicated that only 0.05 % of cells harbored the intended translocation, we switched to a selection-based approach adapted from Vanoli et al., in which we knocked-in a removable, loxP-flanked puromycin resistance cassette at the translocation breakpoint (Vanoli et al., 2017) (**Figure 32**).

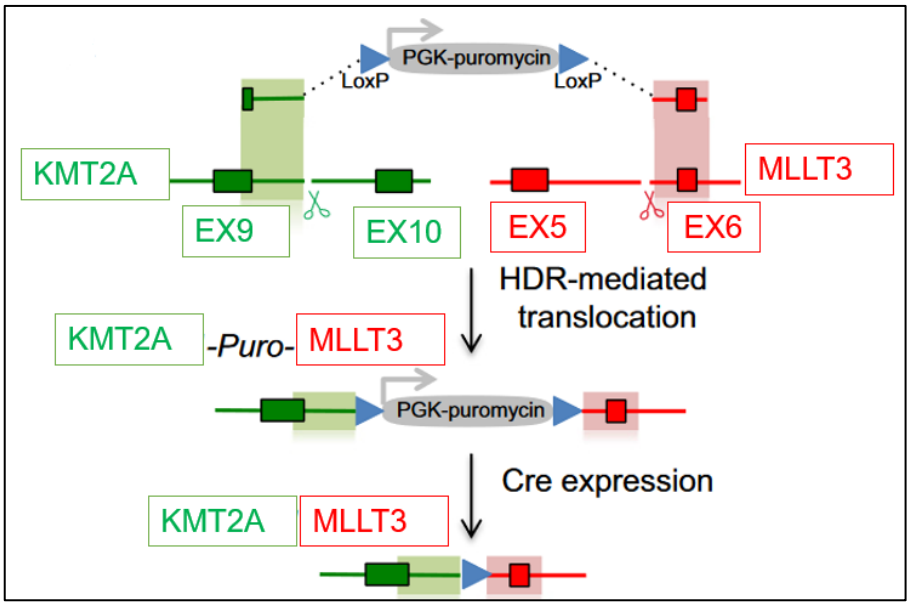


Figure 32: **Homology directed repair (HDR)-mediated translocation (modified from Vanoli et al., 2017).** Gene editing strategy using HDR to generate a *KMT2A::MLLT3* translocation. Double strand breaks are induced by CRISPR/Cas9 in intron 9 of *KMT2A* (red) and intron 5 of *MLLT3* (green) (marked with scissors). HDR uses the donor plasmid with two homology arms for *KMT2A* and *MLLT3* as a template and leads to the insertion of a loxP-flanked mPgk1-promoter driven puromycin resistance cassette. The selection marker can later be removed by transient Cre recombinase expression. Homology directed gene repair (HDR).

After nucleofection of WT hiPSCs, we performed two consecutive rounds of puromycin selection, generating the KM t(9;11) floxed bulk, followed by Cre transfection, resulting in the generation of the KM t(9;11) delta bulk. We designed and performed a series of PCRs to analyze and screen the bulks and, later, the generated single cell translocation clones (**Figure 33 A**). PCR analysis revealed the presence of the translocation in both cell bulks, with both the *KMT2A* and *MLLT3* WT alleles being intact (**Figure 33 B** and **C**). Note that the KMT2A 5'HA PCR and MLLT3 3'HA PCR 1 detect the *KMT2A::MLLT3* KM t(9;11) floxed allele. The successful amplification of the translocation breakpoint through the K::M PCR was achieved exclusively in the cell bulk derived from our initial gene editing strategy termed KM t(9;11) Cas9, wherein we employed only the two orthogonal Cas9 enzymes, although this approach showed low gene editing efficiencies that would hinder screening of translocation-positive clones. Since the applied primers bind outside the homology arms, the successful amplification of this region strongly suggests that the translocation is indeed present in the analyzed bulks.

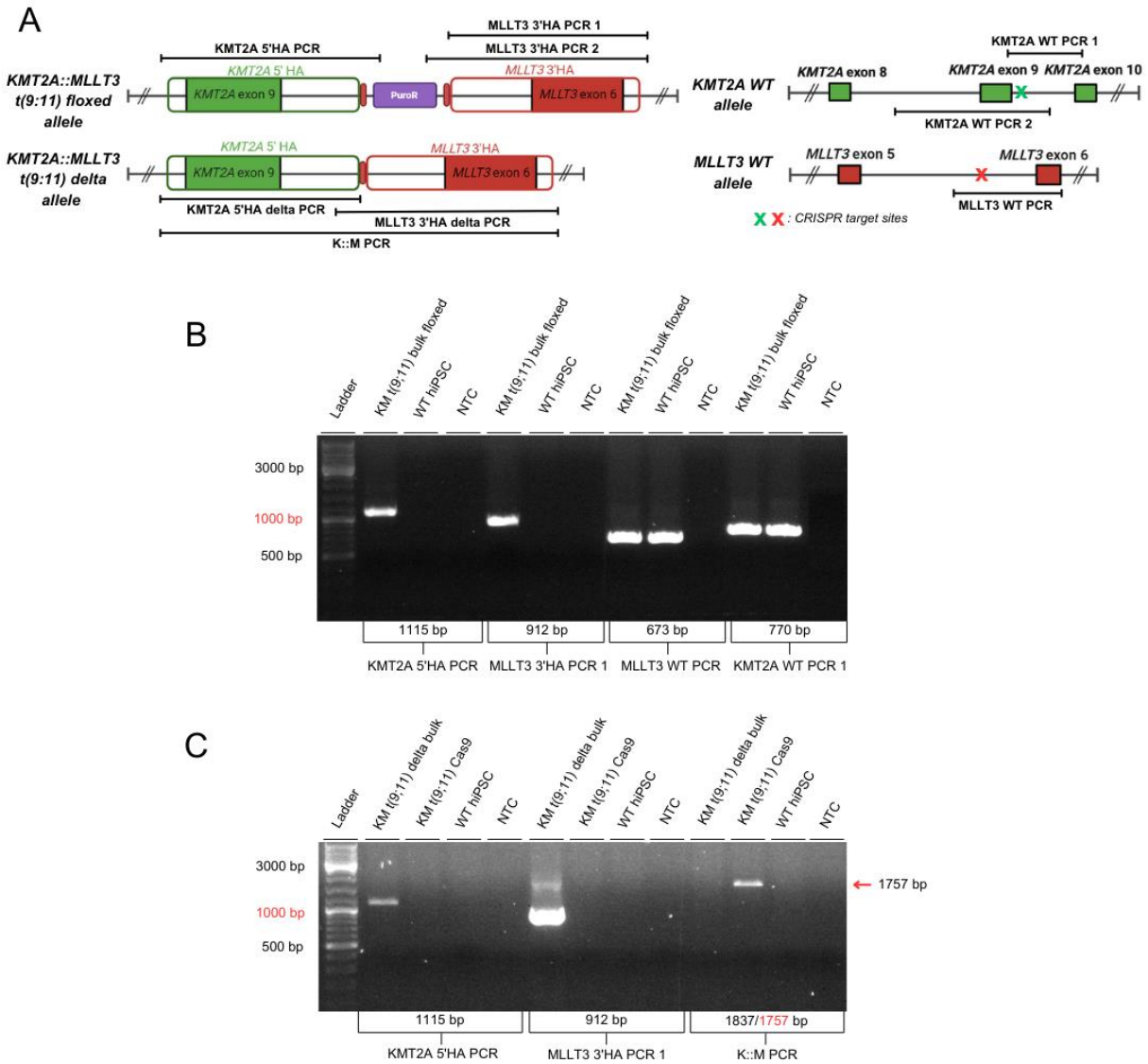


Figure 33: Genotyping of KM t(9;11) bulk cells. Bulk cell DNA from the first translocation approach, termed KM t(9;11) Cas9 bulk, was used as a positive control for the K::M PCR and WT hiPSC DNA as a negative control. **A:** Overview of the *KMT2A::MLLT3* t(9;11) floxed/delta and *KMT2A*/*MLLT3* WT alleles. The regions amplified by PCR are shown. The CRISPR target sites are indicated with an X, and LoxP sites are denoted as red bars. The *KMT2A* 5'HA PCR and *MLLT3* 3'HA PCR 1 and 2 were performed to detect the *KMT2A::MLLT3* t(9;11) floxed allele. Note that the *MLLT3* 3'HA PCR 1 also amplifies the 3'HA of the delta allele. The *KMT2A* WT and *MLLT3* WT PCRs were performed to amplify the respective exon and intron junctions containing the CRISPR target sites. To verify PuroR excision after Cre recombination and the presence of the translocation, the *KMT2A* 5'HA delta and *MLLT3* 3'HA delta PCRs were used. Lastly, the K::M PCR was conducted to amplify the whole translocation region between *KMT2A* intron 8 to *MLLT3* intron 6. **B:** Gel picture of PCR products of DNA from puromycin selected KM t(9;11) floxed bulk cells. **C:** Gel picture of PCR products from the Cre transfected delta and from the first translocation approach without donor (KM t(9;11) Cas9) cell bulks. The K::M PCR product was detected in the KM t(9;11) Cas9 bulk (1757 bp) but not in the delta bulk (1837 bp). No template control (NTC), homology arm (HA), puromycin resistance cassette (PuroR), clustered regularly interspaced short palindromic repeats (CRISPR), *KMT2A::MLLT3* (KM), wild-type (WT). Red arrows indicate expected band sizes.

Therefore, we seeded single cell clones from the KM t(9;11) bulk delta and performed a PCR screen with 96 picked clones using the MLLT3 3'HA PCR1 (**Figure 34 A** and **B**). Our initial PCR screen revealed three clones potentially harboring the translocation, termed KM t(9;11) A5, F5 and E11 (**Figure 34 A** and **B**). However, clone expansion from the replicate plate followed by PCR genotyping showed that only clone A5 might harbor both HAs specific for the *KMT2A::MLLT3* KM t(9;11) delta allele. However, the weak band intensity of the *KMT2A* 5'HA delta PCR product suggests that only a small fraction of cells harbors this HA (**Figure 34 C**). Residual PuroR could not be detected, although the *MLLT3* 3'HA PCR 2 produced unspecific bands and could not be used to assess successful PuroR excision (**Figure 34 C**). Both the *KMT2A* and *MLLT3* WT alleles remained intact in all conditions. The *MLLT3* WT PCR produced an unspecific band in clone A5, which might be a result of suboptimal primer specificity (**Figure 34 C**). Additionally, the presence of the reciprocal MLLT3::KMT2A translocation could not be detected, however no positive control was used (**Figure 34 C**).

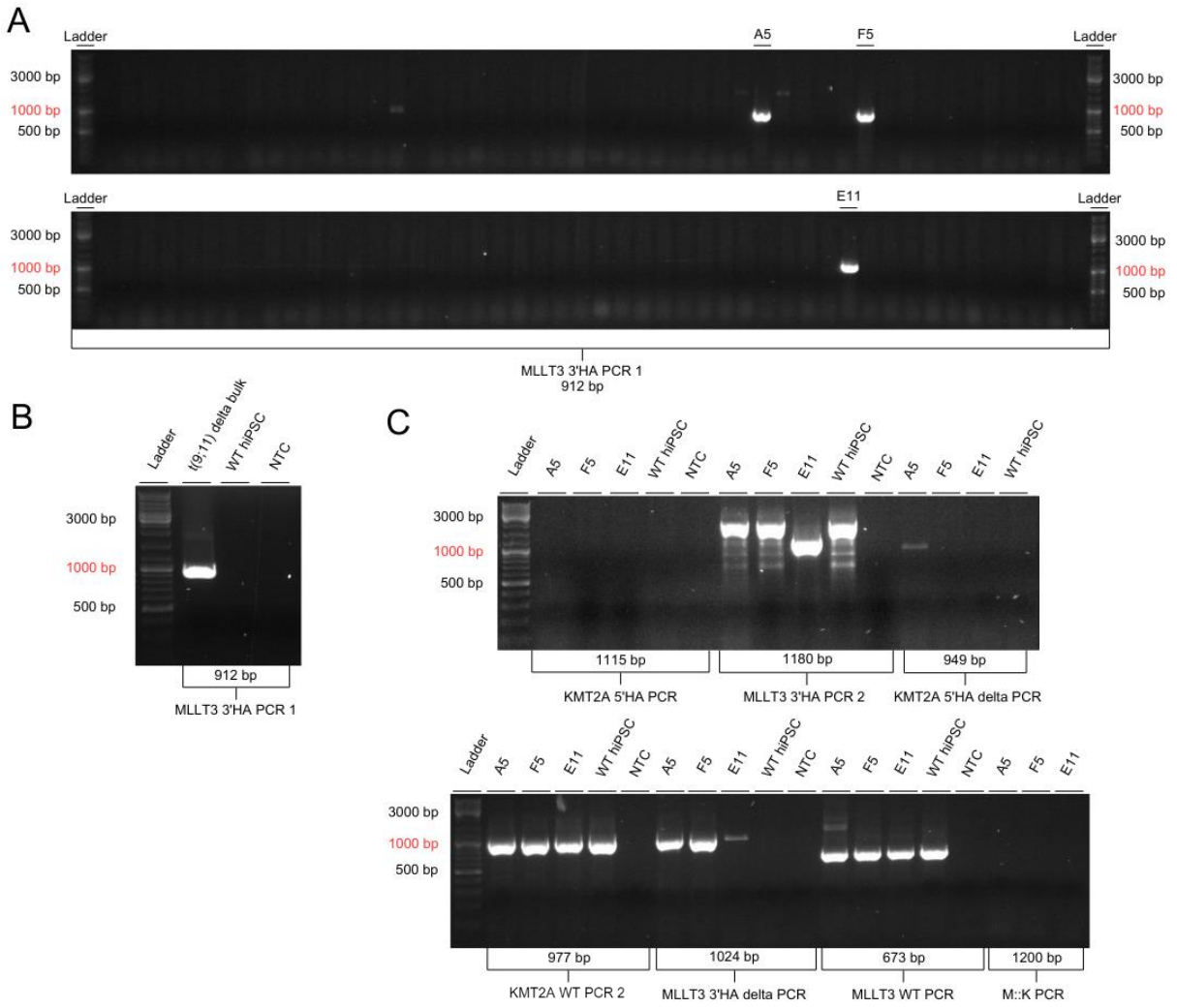


Figure 34: Genotyping of KM t(9;11) translocation clones. A: Gel pictures from the PCR screen of 96 clones using the *MLLT3* 3'HA PCR 1 with clone DNA. **B:** Gel picture of respective controls for the PCR screen using KM t(9;11) delta bulk (positive control) and WT hiPSC (negative control) DNA. **C:** Gel picture of indicated genotyping PCRs performed with DNA from expanded clones. WT hiPSC DNA was used as a negative control. The M::K PCR was performed to detect the reciprocal *MLLT3::KMT2A* translocation. Homology arm (HA), no template control (NTC), *KMT2A::MLLT3* (KM), *MLLT3::KMT2A* (MK), wild-type (WT).

Clone A5 was further analyzed by PCR amplification and sequencing of the *KMT2A* and *MLLT3* regions. Based on the results, KM t(9;11) clone A5 very likely indeed harbors the *KMT2A::MLLT3* translocation (**Figure 35 A, B** and **C**), although the weak band intensity of the *KMT2A* 5'HA delta PCR is indicative of impurity of the clone. Moreover, sequencing of the *KMT2A* and *MLLT3* WT alleles revealed deletions at the corresponding gRNA target sites, showcasing CRISPR/Cas9 induced off-target effects (**Figure 35 D-G**).

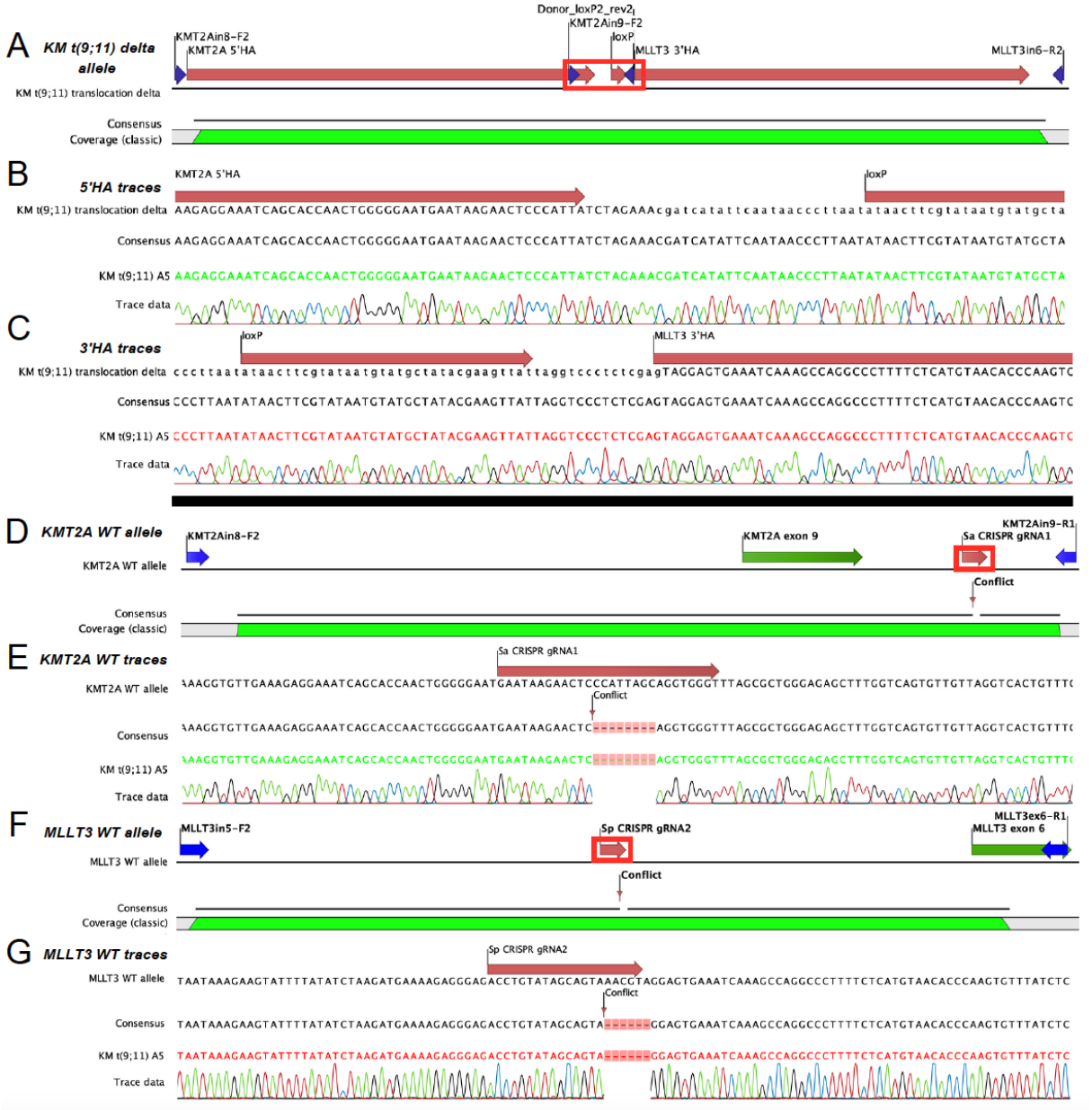


Figure 35: Sequencing results of clone KM t(9;11) A5. Sequences were aligned to their respective reference. Coverages, used primer pairs and CRISPR target sites are indicated. **A:** Overview of the KM t(9;11) delta allele sequence used as a reference. Red box highlights the region magnified in B and C. **B:** Sequencing chromatogram of the *KMT2A* 5'HA region, showing the junction between the *KMT2A* 5'HA and the remaining loxP site. **C:** Sequencing chromatogram of the *MLLT3* 3'HA region, showing the junction between the *MLLT3* 3'HA and the remaining loxP site. **D:** Overview of the *KMT2A* WT allele used as a reference. Red box highlights the region magnified in E. **E:** Chromatogram of the *KMT2A* WT allele showing an 8 bp deletion. **F:** Sequencing overview of the *MLLT3* WT allele used as a reference. Red square highlights the region magnified in G. **G:** Sequencing chromatogram of *MLLT3* WT, showing a 6 bp deletion. Clustered regularly interspaced short palindromic repeats (CRISPR), homology arm (HA), *KMT2A::MLLT3* (KM), wild-type (WT).

To assess the expression of *KMT2A::MLLT3* transcript in KM t(9;11) clone A5, we performed RT-qPCR. As shown in **Figure 36**, the fusion is expressed in KM t(9;11) clone A5. But we

detected significantly higher quantities of *KMT2A* WT transcripts in this clone, indicating the presence of contaminating WT cells (**Figure 36**). Nevertheless, this demonstrates that our gene editing approach indeed induces the t(9;11)(p21;q23)/*KMT2A::MLLT3* translocation in hiPSCs resulting in the expression of corresponding fusion transcripts. However, further in-depth analysis, which go beyond this work would be required before this clones may be used as an *in vitro* model to study the functional effects of KMT2A::MLLT3 expression. Therefore we refrained from conducting further experiments with this clone.

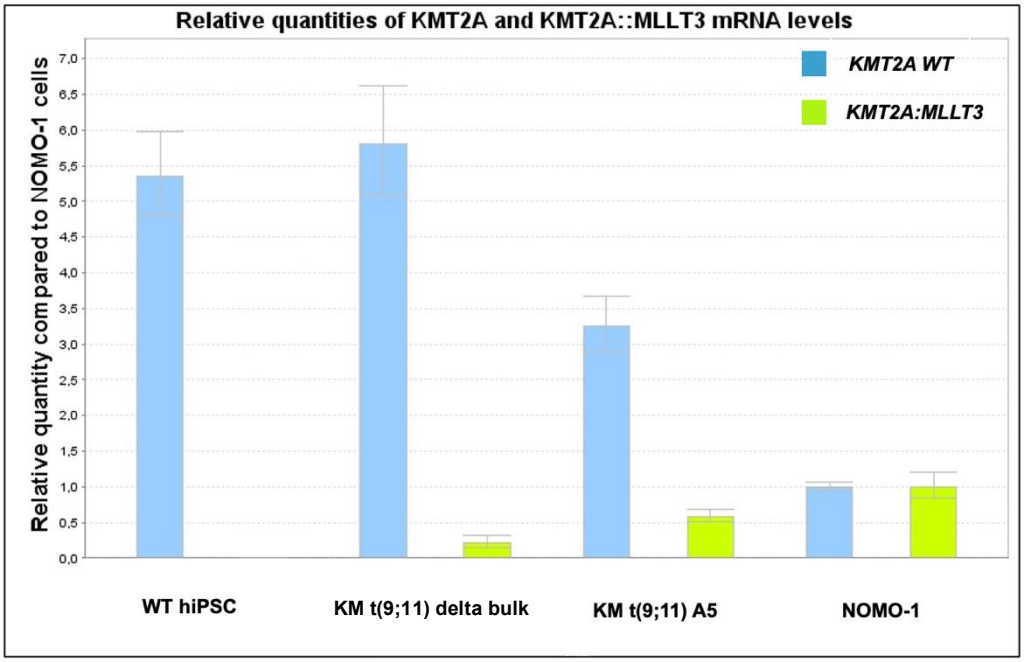


Figure 36: Relative quantities of *KMT2A* WT and *KMT2A::MLLT3* mRNA levels. Relative quantities were calculated by normalization to NOMO-1 cells and housekeeping genes *ABL1* and *GUSB* using the $2^{-\Delta\Delta Ct}$ method. NOMO-1 is a patient-derived cell line harboring a *KMT2A::MLLT3* translocation. *KMT2A::MLLT3* (KM), wild-type (WT).

3.4 Validation of pluripotency and hematopoietic differentiation

The pluripotency of the gene-edited hiPSC clones was assessed by RT-qPCR, analyzing the expression of pluripotency factors *POU5F1* (*OCT4*), *SOX2* and *NANOG*. Both *KMT2A::MLLT3* and NRAS G12D knock-in clones showed normal pluripotency gene expression, apart from the FACS sorted clone KM 2E11, which showed an about twofold reduced expression compared to the parental WT hiPSC cell line, indicating that this clone might have undergone some degree of differentiation and thereby shows reduced pluripotency (**Figure 37** and **Figure 38**).

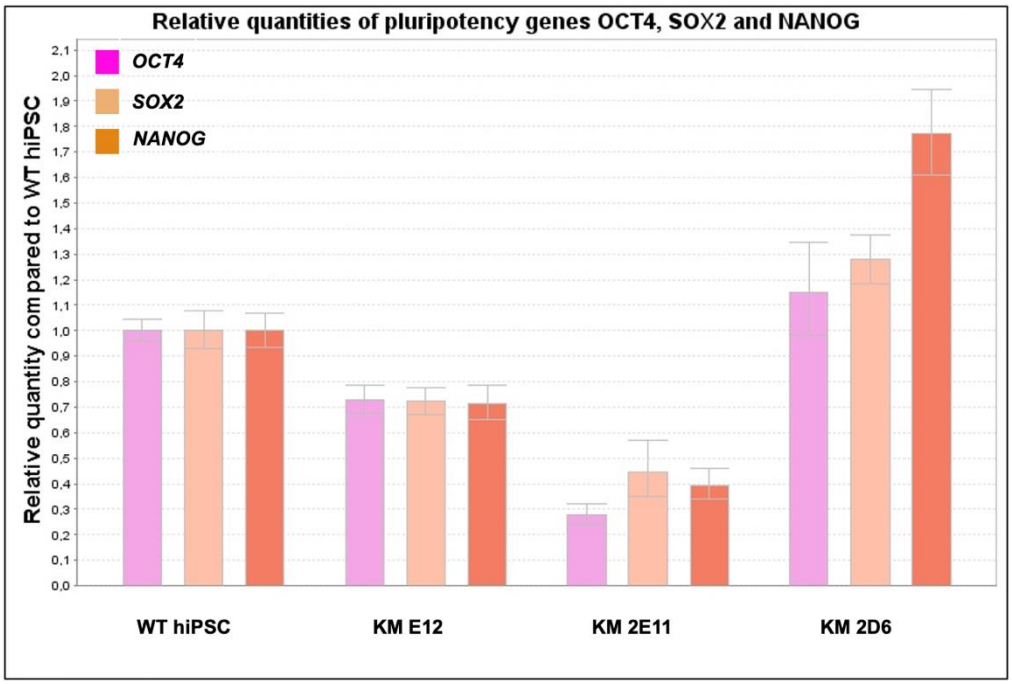


Figure 37: Expression of *OCT4*, *SOX2* and *NANOG* mRNA in *KMT2A::MLLT3* knock-in clones. RT-qPCR analysis of WT hiPSCs and *KMT2A::MLLT3* knock-in clones KM E12, KM 2E11 and KM 2D6. Relative quantities were calculated by normalization to WT hiPSC cells and the expression of the reference genes *GUSB* and *ABL1* using the $2^{-\Delta\Delta Ct}$ method. *KMT2A::MLLT3* (KM), wild-type (WT).

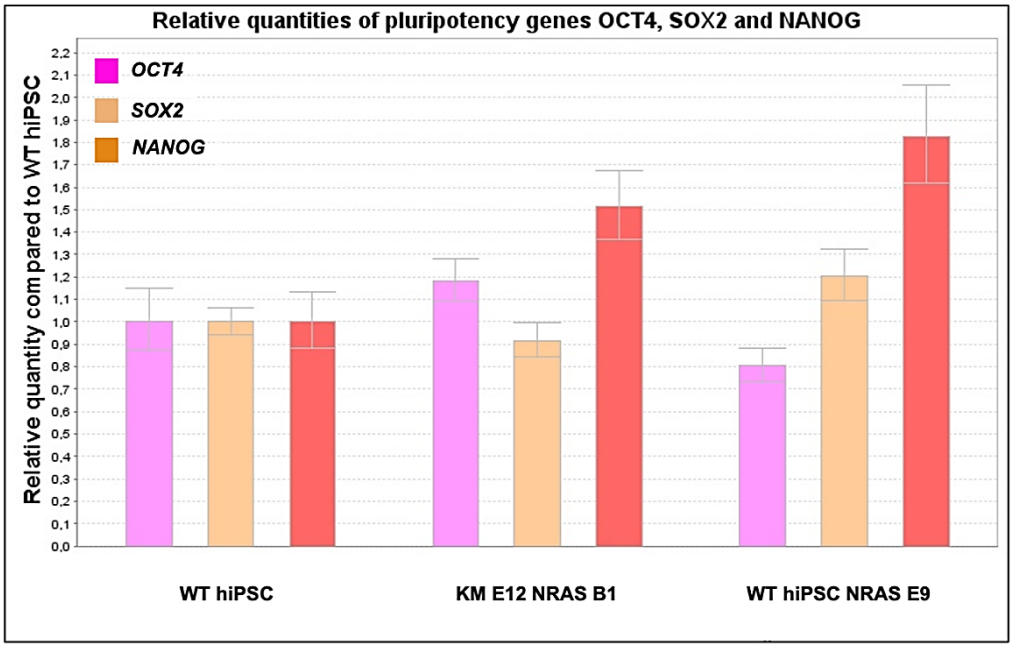


Figure 38: Expression of *OCT4*, *SOX2* and *NANOG* mRNA in *NRAS* G12D knock-in clones. RT-qPCR analysis of WT hiPSCs, KM E12 NRAS G12D clone B1 (KM E12 NRAS B1) and WT hiPSC NRAS G12D clone E9 (WT hiPSC NRAS E9). Relative quantities were calculated by normalization to WT hiPSC cells and the expression of the reference genes *GUSB* and *ABL1* using the $2^{-\Delta\Delta Ct}$ method. *KMT2A::MLLT3* (KM), wild-type (WT).

Due to limited capacity, we used only one each of the NRAS G12D knock-in clones, namely, KM E12 NRAS G12D clone B1 and WT hiPSC NRAS G12D clone E9, for hematopoietic differentiation. Within the scope of this work, we performed three hematopoietic differentiation experiments (HEM), termed HEM 33, 34, and 35. HEM 33 and 34 were performed using WT hiPSC, *KMT2A::MLLT3* knock-in clones KM E12, KM 2E11 and KM 2D6. Note that in HEM 33, the used dTAG-13 compound from Sigma-Aldrich was later found to be nonfunctional, thus the degrader-treated conditions had to be excluded. Therefore, we conducted HEM 34 as an additional differentiation with the same cell lines and a functional dTAG-13 compound from Tocris Bioscience. HEM 35 was performed using WT hiPSC, *KMT2A::MLLT3* knock-in clone KM E12, KM E12 NRAS G12D clone B1, and WT hiPSC NRAS G12D clone E9.

The gene-edited hiPSCs were *in vitro* differentiated into hematopoietic stem and progenitor cells (HSPCs) using the approach depicted in **Figure 39 A**, which resulted in the emergence of hematopoietic stem and progenitor cells (**Figure 39 B-E**). Hematopoietic differentiation was initiated by the cultivation of hiPSC clusters in HEM A media to induce mesoderm formation (**Figure 39 B**). During the first day, the GSK3 inhibitor CHIR99021 (GSKi) was added to enhance definitive hematopoiesis. After the third day of differentiation, HEM A was replaced with HEM B medium to direct differentiation toward the hemato-endothelial lineage (**Figure 39 C**). The medium was supplemneted with 100 nM dTAG-13 until the sixth day to prevent

interference of KMT2A::MLLT3 fusion protein expression with early differentiation. On the sixth day, dTAG-13 was washed out for half of the cells to express the KMT2A::MLLT3 fusion protein. On day 9, the first hiPSC-derived progenitors emerged, identifiable as clusters of cells with a small and rounded morphology, which were harvested after 12 to 13 days of hematopoietic differentiation. (**Figure 39 D** and **E**).

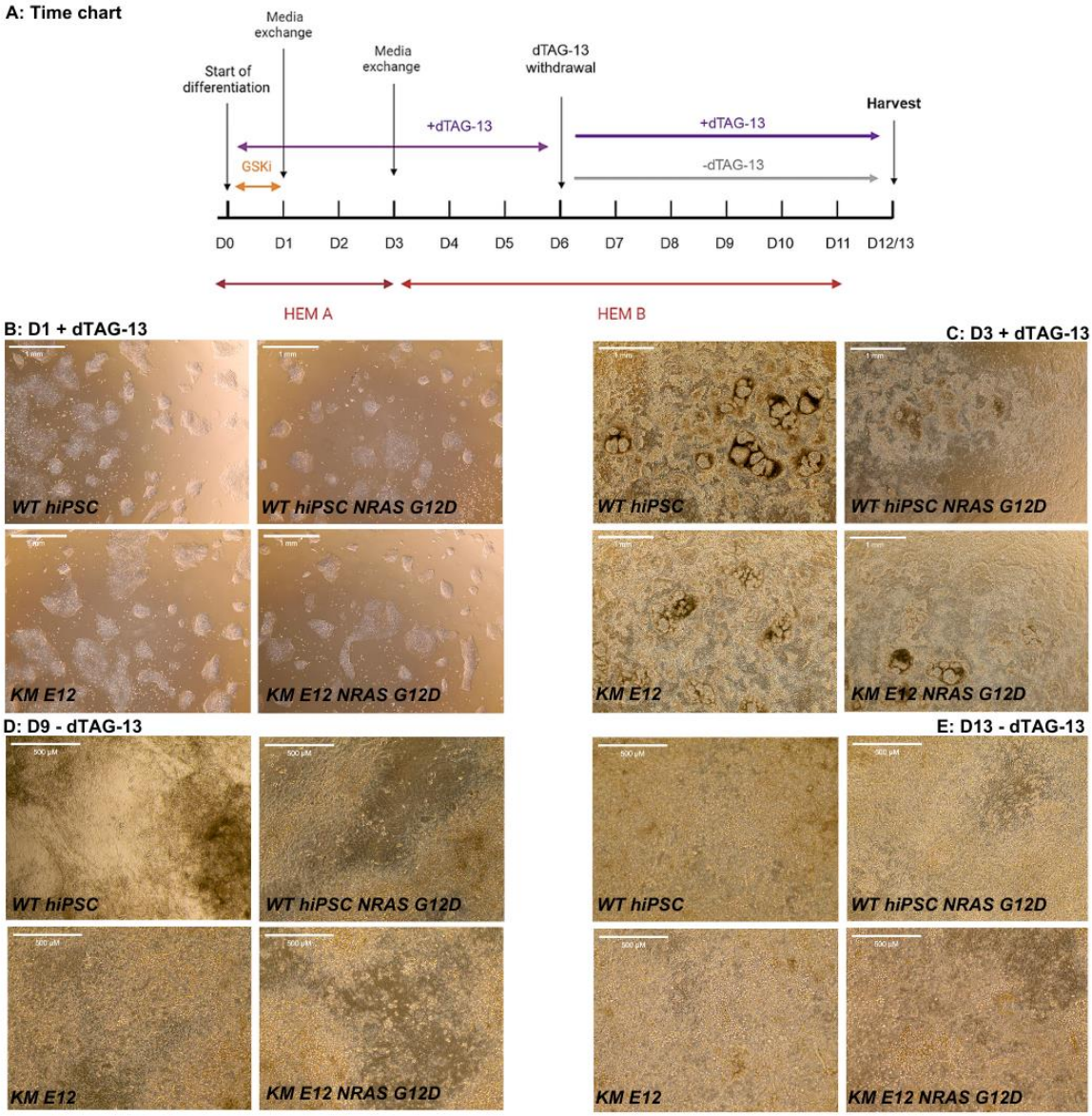


Figure 39: Hematopoietic differentiation using the STEMdiff hematopoietic kit. A: Schematic time chart of the hematopoietic differentiation protocol. **B:** Pictures of hiPSCs after one day (D1) and **C:** after three days (D3) of hematopoietic differentiation. **D:** Pictures of hiPSCs after nine days (D9) and **E:** after thirteen days (D13) of hematopoietic differentiation without dTAG-13 supplementation. *KMT2A::MLLT3* (KM), wild-type (WT). White bars correspond to 1 mm in B and C and 500 µm in D and E.

3.5 Growth curves of KMT2A::MLLT3-expressing and constitutive NRAS-mutated cells

After harvesting live cells from hematopoietic differentiation by dead cell removal, we further cultured the harvested HSPCs in medium supplemented with a combination of seven cytokines, and regularly counted the cell numbers to generate growth curves. As shown in **Figure 40** clone KM E12 derived from the HEM 33 experiment displayed outgrowth while clones KM 2E11, KM 2D6 and WT hiPSCs did not.

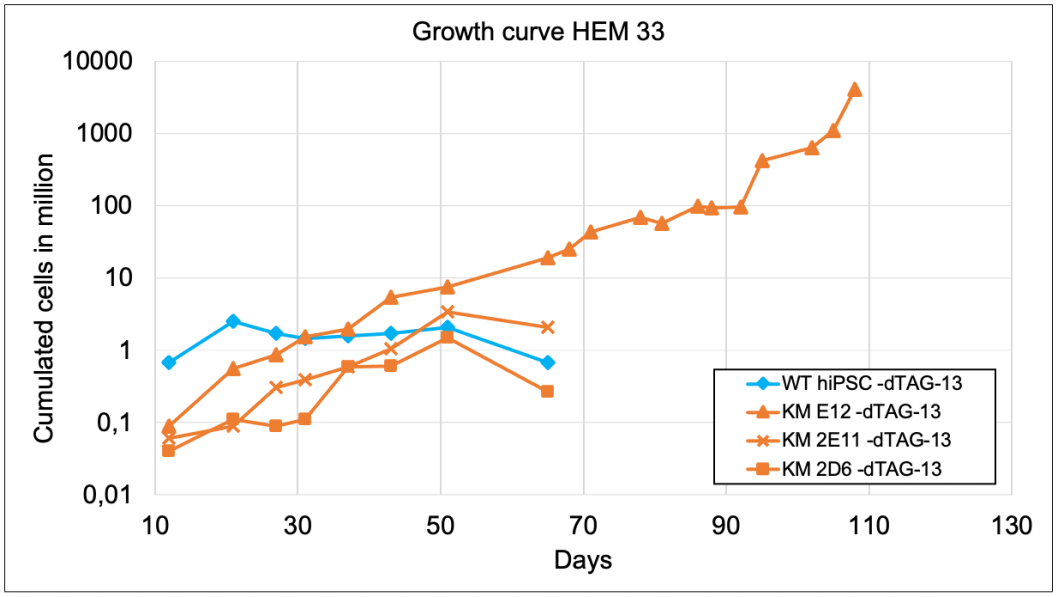


Figure 40: Growth curve of WT hiPSCs and KMT2A::MLLT3-expressing progenitors derived from the first hematopoietic differentiation (HEM 33) in liquid cultures. Supernatant cells were harvested and counted after hematopoietic differentiation and cultivated in medium supplemented with SCF (50 ng/µl), EPO (10 ng/ml), GM-CSF, IL-3, IL-6, G-CSF, and TPO (all 20 ng/ml each). Note that only conditions without supplementation of dTAG-13 are shown since, in HEM 33, the used dTAG-13 compound was not functional. X-axis depicts days of culture (Days). Wild-type (WT).

Assessing the growth of cells harvested from hematopoietic differentiation HEM 34 with a functional dTAG-13 compound, we observed no outgrowth for any of the used KMT2A::MLLT3 expressing knock-in clones or WT hiPSCs during two months of culture (**Figure 41**).

Growth curves of hiPSC-derived hematopoietic progenitors from the third hematopoietic differentiation, HEM 35, showed that the introduction of the NRAS G12D mutation into KMT2A::MLLT3-expressing knock-in clone KM E12 resulted in rapid outgrowth of hiPSC-derived hematopoietic progenitors (**Figure 42**), while the introduction of the NRAS G12D mutation into the WT background did not induce outgrowth.

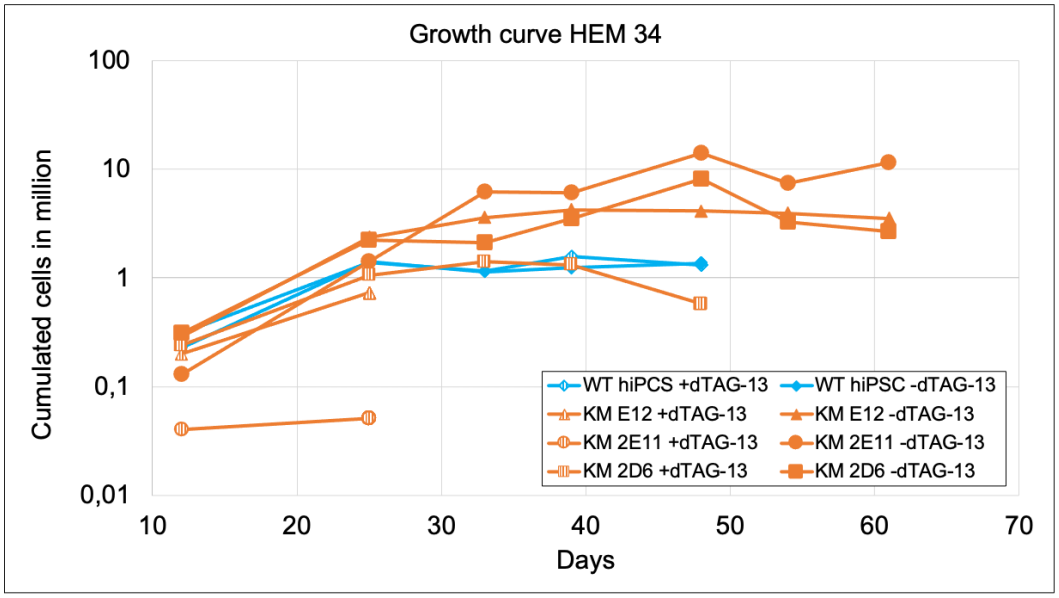


Figure 41: Growth curve of WT hiPSCs and KMT2A::MLLT3-expressing progenitors derived from the second hematopoietic differentiation (HEM 34) in liquid cultures. Supernatant cells were harvested and counted after hematopoietic differentiation and cultivated in medium supplemented with SCF (50 ng/µl), EPO (10 ng/ml), GM-CSF, IL-3, IL-6, G-CSF, and TPO (all 20 ng/ml each). Clones were either cultivated with or without 100 nM dTAG-13. X-axis depicts days of culture (Days). Wild-type (WT).

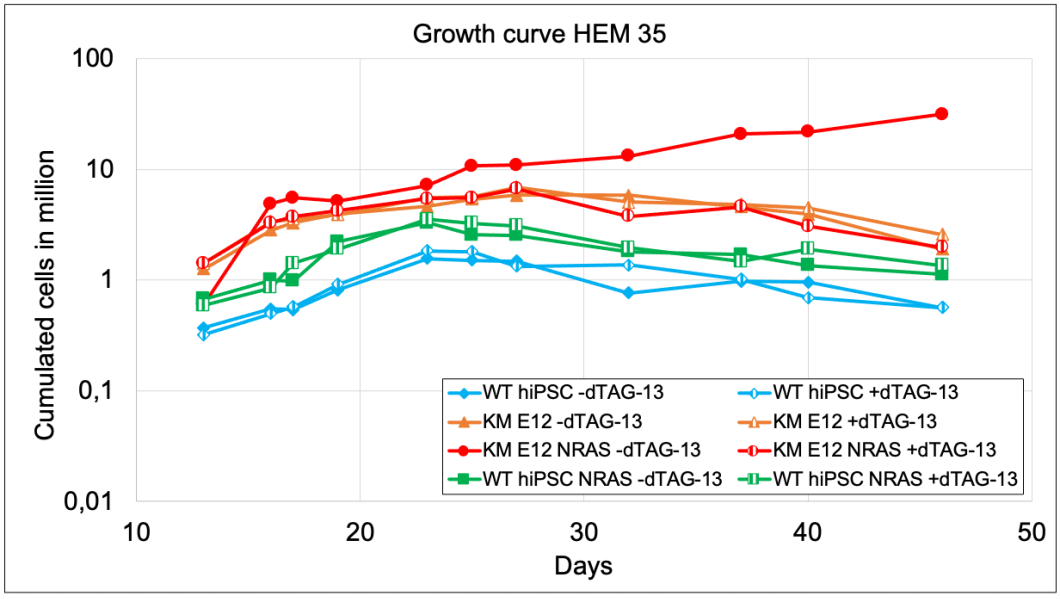


Figure 42: Growth curve of WT hiPSCs, KMT2A::MLLT3 and NRAS G12D expressing progenitors derived from the third hematopoietic differentiation experiment (HEM 35) in liquid cultures. Supernatant cells were harvested and counted after hematopoietic differentiation and cultivated in medium supplemented with SCF (50 ng/µl), EPO (10 ng/ml), GM-CSF, IL-3, IL-6, G-CSF, and TPO (all 20 ng/ml each). Clones were either cultivated with or without dTAG-13. X-axis depicts days of culture (Days). KM E12 NRAS G12D knock-in clone B1 (KM E12 NRAS), WT hiPSC NRAS G12D knock-in clone E9 (WT hiPSC NRAS), wild-type (WT).

Moreover, when comparing the growth curves of dTAG-13 compound untreated clone KM E12 with and without the NRAS G12D mutation in experiments HEM 33, 34 and 35, it became evident that constitutive expression of G12D-mutated NRAS in combination with the KMT2A::MLLT3 fusion protein leads to accelerated outgrowth as compared to cells that only carry the fusion (**Figure 43**).

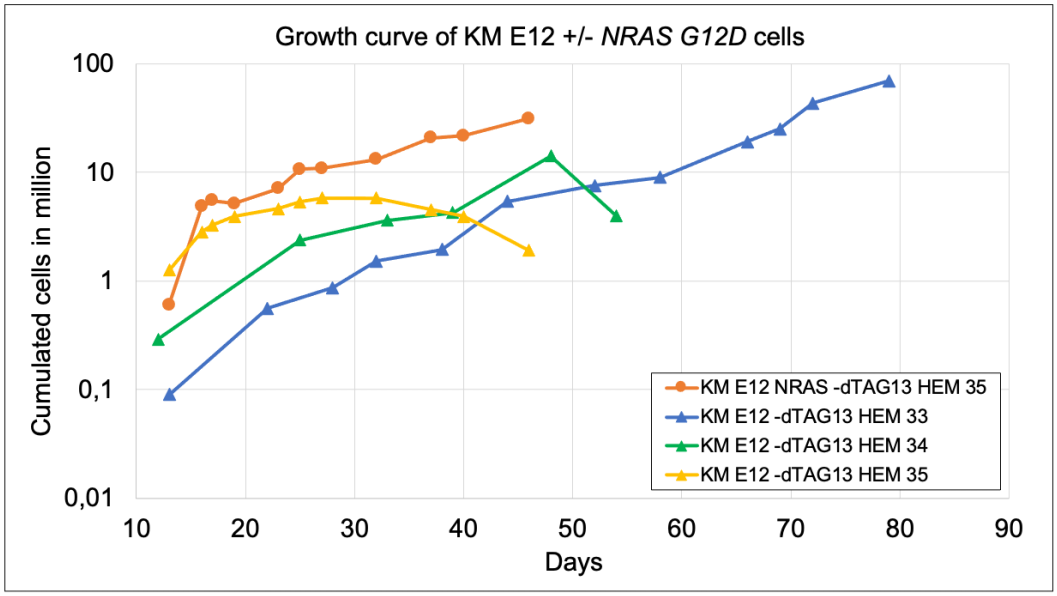


Figure 43: Growth curve of untreated KMT2A::MLLT*3*-exressing progenitors derived from KM E12 from HEM 33, 34 and 35 with or without the NRAS G12D mutation in liquid culture. Supernatant cells were harvested, after hematopoietic differentiation, cultured in medium supplemented with SCF (50 ng/µl), EPO (10 ng/ml), GM-CSF, IL-3, IL-6, G-CSF, and TPO (all 20 ng/ml each), and counted regularly. Only conditions without dTAG-13 are shown. X-axis depicts days of culture (Days). KM E12 NRAS G12D knock-in clone B1 (KM E12 NRAS).

Monitoring the cells under the microscope, we noted morphological differences between WT cells and KMT2A::MLLT3-expressing clones. WT cells showed features associated with myeloid differentiation, such as the presence of large adherent cells typical for macrophages, while the KMT2A::MLLT3-expressing cells showed features associated with undifferentiated hematopoietic progenitors, such as a non-adherent, small and regularly round shaped appearance (**Figure 44**).

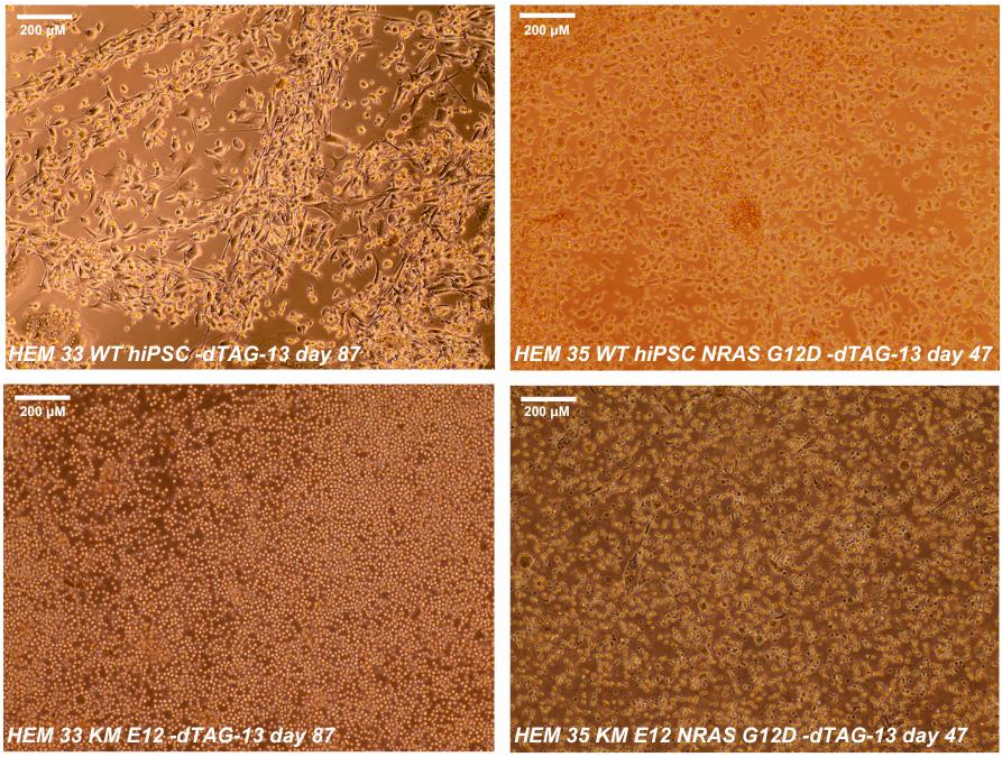


Figure 44: Images of cells of WT hiPSCs, KMT2A::MLLT3-expressing KM E12 with and without the NRAS G12D mutation, derived from hematopoietic differentiation, and grown in liquid culture. Pictures were taken on the indicated days of long-term cultivation following hematopoietic differentiation (HEM) experiments 33 or 35. White bars correspond to 200 µm. KM E12 NRAS G12D B1 (KM E12 NRAS), WT hiPSC NRAS G12D E9 (WT hiPSC NRAS), wild-type (WT).

3.6 Cytokine and oncogene dependency assays

To investigate which supplemented cytokines are required for growth and/or survival of KMT2A::MLLT3-expressing cells, we performed a cytokine dependency assay using cells of clone KM E12 after 91 days cultivation in liquid culture. As shown in **Figure 45,** cells only survived in medium supplemented with either all cytokines, GM-CSF or IL-3 or a combination of both, suggesting that KMT2A::MLLT3 expressing cells are still dependent on the presence of cytokines essential for granulocytes and macrophages.

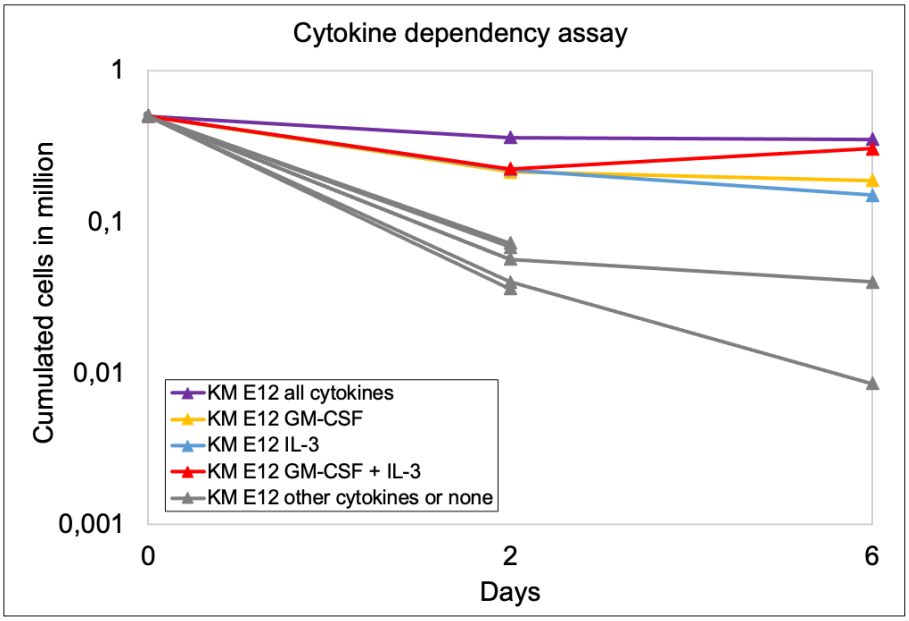


Figure 45*:* Cytokine dependency assay of KMT2A::MLLT3-expressing knock-in clone KM E12 cells derived from HEM 33. KM E12 cells were cultured in media supplemented with the indicated cytokine(s) and counted on days 2 and 6. All cytokines means that the medium was supplemented with the cytokines SCF (50 ng/µl), EPO (10 ng/ml), GM-CSF, IL-3, IL-6, G-SCF and TPO (all 20 ng/ml). Other cytokines or none (grey lines) correspond to conditions where the medium was supplemented with only one (SCF, G-CSF, IL-6, EPO, or TPO) or no cytokine. Hematopoietic differentiation (HEM).

Additionally, we also investigated if cells are addicted to the expression of the KMT2A::MLLT3 fusion protein for cell growth by performing an oncogene dependency assay in which we cultured KM E12 derived cells after 101 days in liquid culture with either dTAG-13, dTAGv-1 or no compound. Cultivation of clone KM E12 without dTAG promoted cell growth, while treatment with either 100 nM dTAG-13 or dTAGv-1 halted cell proliferation but did not completely impair their survival (**Figure 46**).

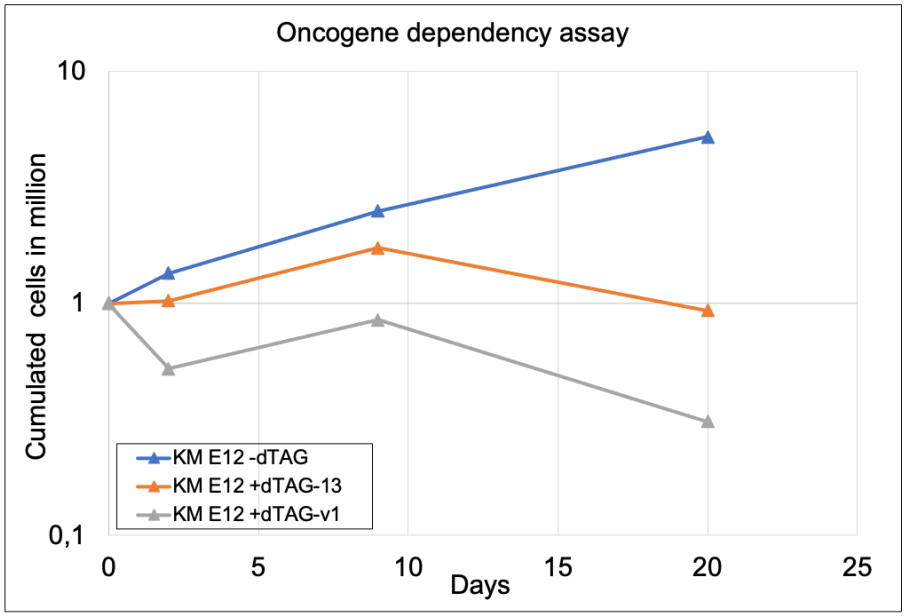


Figure 46: Oncogene dependency assay of KMT2A::MLLT3 clone KM E12 cells derived from HEM 33. The cell counts of clone KM E12 cultured in media supplemented with either 100 nM dTAG-13, 100 nM dTAG-v1 or no dTAG compound are shown. Medium was always supplemented with SCF (50 ng/µl), EPO (10 ng/ml), GM-CSF, IL-3, IL-6, G-CSF, and TPO (all 20 ng/ml each). Hematopoietic differentiation (HEM).

3.7 Colony formation assays

To investigate clonogenicity, we performed a series of colony formation assays with the hiPSC-derived HSPCs harvested at days 12-13 of differentiation. To assess the effects of the KMT2A::MLLT3 fusion protein on myeloid differentiation, we performed MethoCult assays using the hiPSC-derived HSPCs of experiments HEM 33-35. As shown in **Figure 47**, both dTAG-13 treated and untreated KMT2A::MLLT3-expressing cells exhibit a significant increase in undifferentiated mixed granulocyte/macrophage colony forming units (CFU-GM) and an overall greater number of total CFUs indicating that the fusion is still, at least weakly, expressed despite dTAG-13 treatment. This increase in CFUs is further enhanced in KMT2A::MLLT3-expressing cells, which in addition harbor a constitutive NRAS G12D mutation (**Figure 47**).

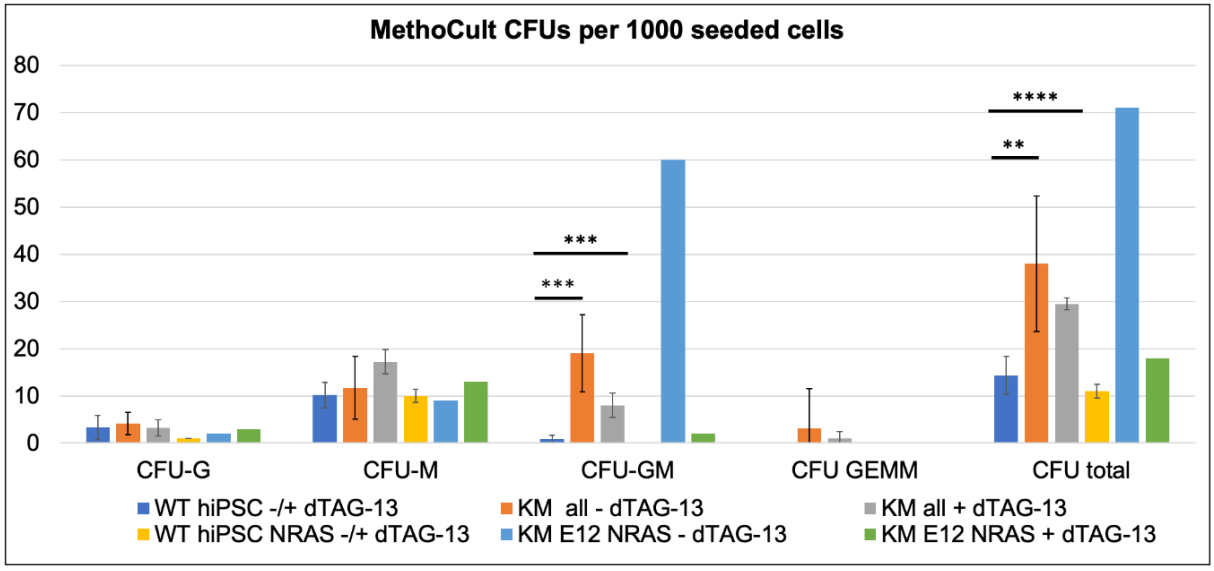


Figure 47: Results of MethoCult assays. Depicted are mean CFU counts ± standard deviations from data of hiPSC-derived hematopoietic stem and progenitor cells from HEM 33, 34, and 35 experiments. 1000 cells were seeded into 3 ml semi-solid MethoCult medium in 6-well plates, and CFUs were enumerated after 14 days of cultivation. Statistical significance was calculated by a one-sided ANOVA and post hoc analysis to investigate group differences using a two tailed t-test. To account for multiple testing, p-values were adjusted using the Bonferroni correction. Significance values : $p \leq 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$. WT hiPSC -/+ dTAG13 (n=6), KM E12 - all (n=7), KM all + dTAG13 (n=4), WT hiPSC NRAS -/+ dTAG13 (n=2), KM E12 NRAS - dTAG13 (n=1), KM E12 NRAS + dTAG13 (n=1). Colony forming unit (CFU), granulocyte (G), macrophage (M), granulocyte/macrophage (GM), granulocyte-erythrocyte-macrophage-megakaryocyte (GEMM), KM 2D6/2E11/E12 (KM all), KM E12 NRAS G12D B1 (KM E12 NRAS), WT hiPSC NRAS G12D E9 (WT hiPSC NRAS), wild-type (WT).

The increase in CFU numbers upon introduction of the NRAS G12D mutation in KMT2A::MLLT3-expressing cells is also reflected by the overall live cell numbers harvested from MethoCult assays, however, the differences were not statistically significant (**Figure 48**).

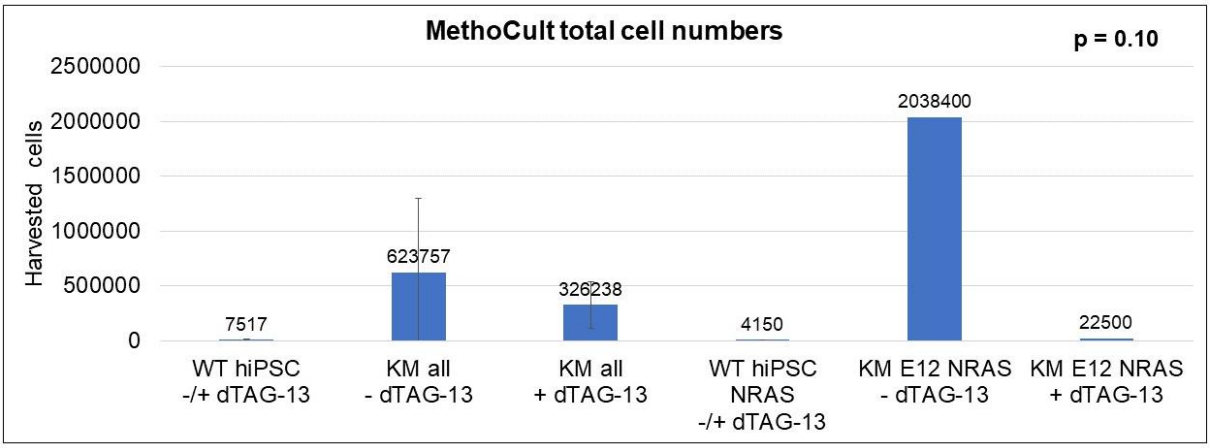


Figure 48: Cell counts of cells harvested from MethoCult semi-solid medium. Depicted are mean cell counts ± standard deviations. Cells were harvested and counted immediately after enumerating CFU numbers. To assess statistical significance, a one-sided ANOVA was performed, with $p \leq 0.05$ considered significant. WT hiPSC -/+ dTAG13 (n=6), KM all - dTAG13 (n=7), KM all + dTAG13 (n=4), WT hiPSC NRAS -/+ dTAG13 (n=2), KM E12 NRAS - dTAG13 (n=1), KM E12 NRAS + dTAG13 (n=1). KM 2D6/2E11/E12 (KM all), KM E12 NRAS G12D B1 (KM E12 NRAS), WT hiPSC NRAS G12D E9 (WT hiPSC NRAS), wild-type (WT).

Inspection of MethoCult assays under the microscope showed that expression of the KMT2A::MLLT3 fusion resulted in highly enlarged CFU-GMs consisting mainly of atypical, undifferentiated, blast-like cells (**Figure 49**).

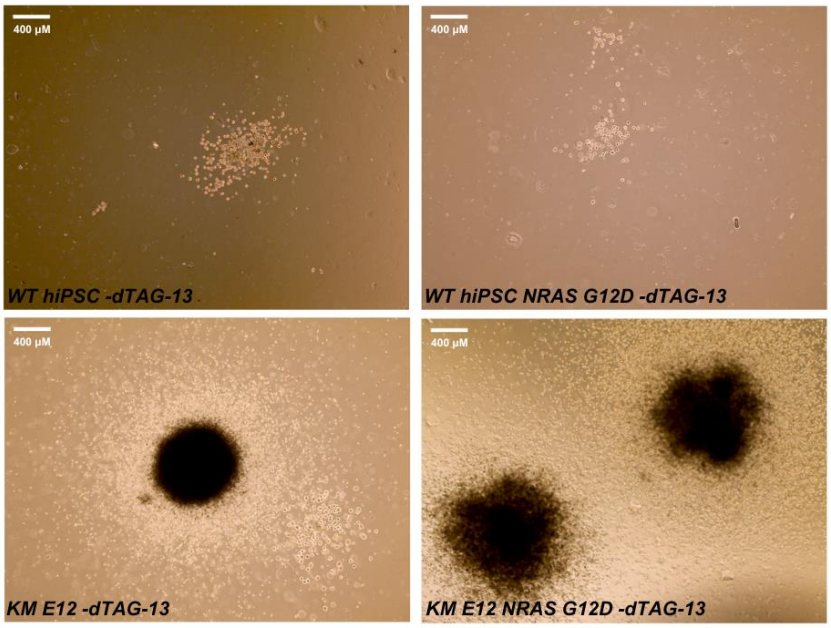


Figure 49: Pictures of colonies from MethoCult assays. Large, atypical CFU-GM are present in KMT2A::MLLT3-expressing conditions. White bars correspond to 400 µm. Granulocyte/Macrophage colony forming units (CFU-GM), KM E12 NRAS G12D B1 (KM E12 NRAS), WT hiPSC NRAS G12D E9 (WT hiPSC NRAS), wild-type (WT):

Using MegaCult assays, we assessed the effects of the KMT2A::MLLT3 fusion with and without the NRAS G12D mutation on megakaryocytic differentiation. As shown in **Figure 50**, we observed a significant increase in non-megakaryocytic CFUs for KMT2A::MLLT3-expressing cells, and an about two-fold increase in small megakaryocytic colonies/CFUs of KM NRAS G12D cells, resulting in an increase in total colony/CFU numbers. Moreover, also dTAG-13 treated KM cells showed an increase in non-megakaryocytic CFUs, again indicating incomplete degradation of the fusion protein.

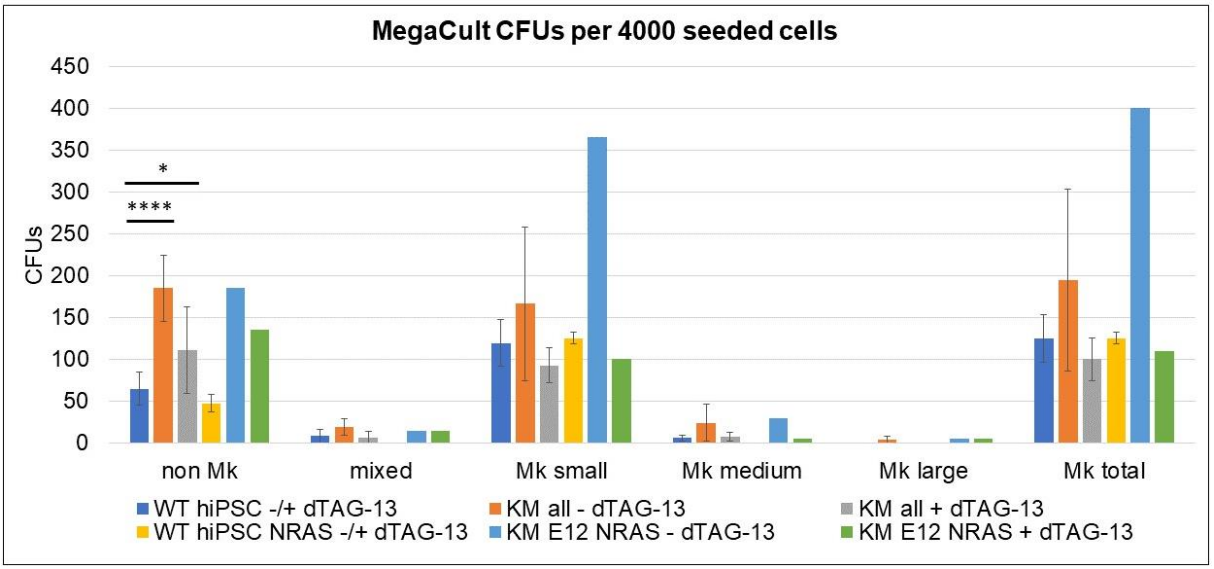


Figure 50: Results of MegaCult assays. Mean CFU counts ± standard deviations of three MegaCult experiments with hiPSC-derived hematopoietic stem and progenitor cells from HEM 33, 34, and 35. 4000 cells were seeded in MegaCult-C medium containing cytokines and after 12-14 days of culture, emerging colonies were fixed, stained, and enumerated. Statistical significance was calculated by a one-sided ANOVA and post hoc analysis to investigate group differences using a two tailed t-test. To account for multiple testing, p-values were adjusted using the Bonferroni correction. Significance values : $p \leq 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$. WT hiPSC -/+ dTAG13 (n=6), KM all - dTAG13 (n=7), KM all + dTAG13 (n=4), WT hiPSC NRAS -/+ dTAG13 (n=2), KM E12 NRAS - dTAG13 (n=1), KM E12 NRAS + dTAG13 (n=1). Megakaryocytic CFUs (Mk), Non-megakaryocytic CFUs (non-Mk), mixed non-Mk and Mk CFUs (mixed), CFU-Mk small: 3 to 20 cells, CFU-Mk medium: 21 to 49 cells and CFU-Mk large: ≥ 50 cells. KM 2D6/2E11/E12 (KM all), KM E12 NRAS G12D B1 (KM E12 NRAS), WT hiPSC NRAS G12D E9 (WT hiPSC NRAS), wild-type (WT).

Inspecting fixed and stained MegaCult CFUs by microscopy, we observed that most non-Mk CFUs of the KMT2A::MLLT3-expressing conditions were substantially larger than WT controls (**Figure 51**).

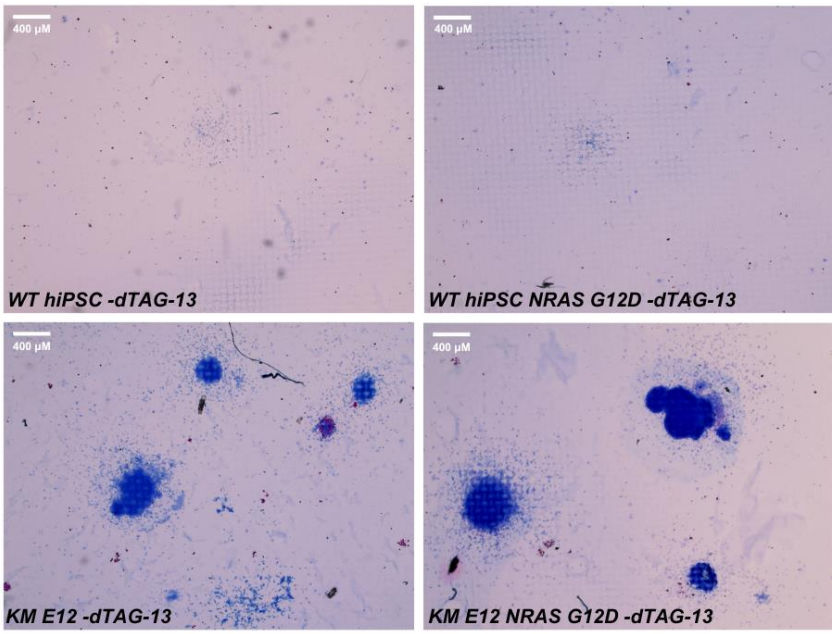


Figure 51: Images of stained CFUs from MegaCult assays. Non-megakaryocytic cells appear blue. Megakaryocytic cells are CD41-positive (red). After 14 days of culture in serum-free MegaCult-C medium, cells were fixed and stained, and pictures were taken. KM E12 NRAS G12D B1 (KM E12 NRAS), WT hiPSC NRAS G12D E9 (WT hiPSC NRAS), wild-type (WT).

3.8 RT-qPCR analysis of hiPSC-derived hematopoietic progenitors in liquid culture and MethoCult assays

To investigate whether our KMT2A::MLLT3- and/or constitutive NRAS G12D mutation expressing hiPSC-derived hematopoietic progenitors recapitulate gene expression of AML blasts, we analyzed the expression of some of their target genes. For this purpose, we performed RT-qPCR analysis to quantify the mRNA expression of two negative regulators associated with prolonged RAS signaling, namely, *DUSP6* and *SPRY2* (Muhammad et al., 2018; Shojaee et al., 2015; Tsavachidou et al., 2004). Additionally, we assessed the expression of a primary gene target of KMT2A::MLLT3, *HOXA9*. In our RT-qPCR analysis, we didn't observe significant alterations in the expression levels of *DUSP6* and *SPRY2* within our hiPSC-derived hematopoietic progenitors harboring the NRAS G12D mutation (**Figure 52**). Furthermore, *HOXA9* mRNA expression was not increased in our hiPSC-derived progenitors expressing KMT2A::MLLT3 (**Figure 52**).

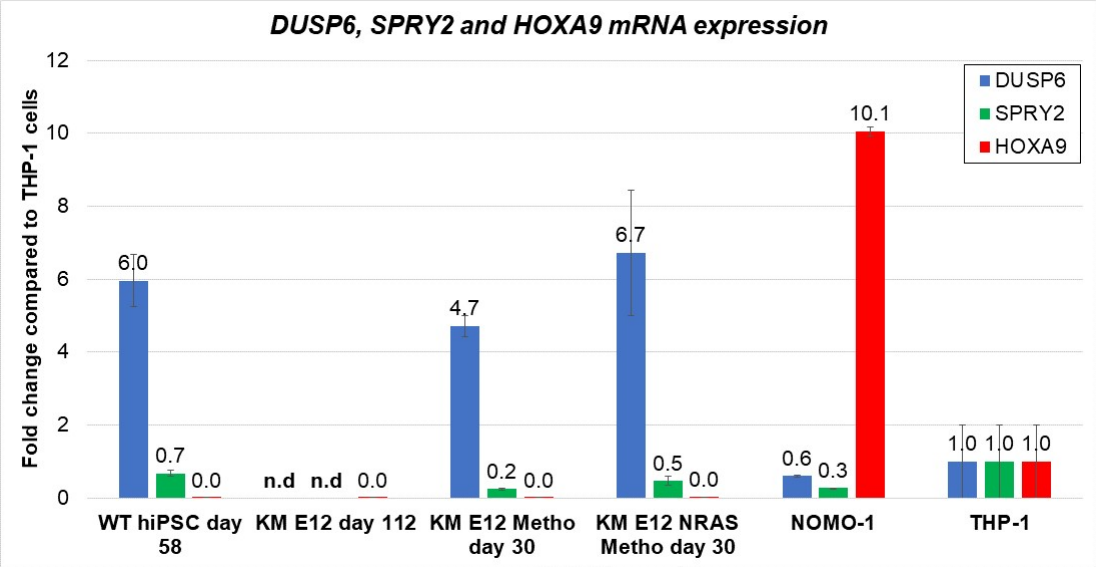


Figure 52: *DUSP6*, *SPRY2* and *HOXA9* RT-qPCR results. All conditions shown are dTAG untreated. WT hiPSCs were harvested on day 58 of liquid culture and are derived from HEM34. KM E12 cells were harvested on day 112 of liquid culture and are derived from HEM33. KM E12 and KM E12 NRAS G12D Metho cells were harvested on day 30 after HEM35 from MethoCult semi-solid media. Relative quantities were calculated by normalization to THP-1 cells and housekeeping genes *ABL1* and *GUSB* by the $2^{-\Delta\Delta Ct}$ method. NOMO-1 and THP-1 are patient-derived cell lines harboring the *KMT2A::MLLT3* translocations and activating RAS mutations. Not determined (n.d.), MethoCult (Metho), KM E12 NRAS G12D B1 (KM E12 NRAS), KMT2A::MLLT3 (KM), wild-type (WT), hematopoietic differentiation (HEM).

4. Discussion

In this study, we aimed to establish leukemia model systems to study the oncogenic mechanisms of the KMT2A::MLLT3 fusion protein with and without an additional NRAS G12D mutation. Taking advantage of the unique differentiation potential of hiPSCs and the CRISPR/Cas9 gene editing technology, we could generate cell lines harboring a conditional *KMT2A::MLLT3* fusion allele, a constitutively active NRAS mutation, or both. Subsequent *in vitro* differentiation of the respective hiPSCs into hematopoietic progenitors has been shown to recapitulate important aspects of leukemogenesis (Chao et al., 2017; Papapetrou, 2019). Therefore, we hypothesized that expression of the KMT2A::MLLT3 oncoprotein would transform the affected cells and thereby phenocopy pediatric AML.

Using the dTAG system allowed for the conditional degradation of the fusion protein to study the immediate and direct effects of KMT2A::MLLT3 on hematopoietic differentiation without the confounding effects of knock-down or knock-out approaches (Nabet et al., 2018). Although hiPSC models for KMT2A::MLLT3-driven AML have been established before, none harbored an additional NRAS G12D mutation (Heuts et al., 2023; Liu et al., 2014). After hematopoietic differentiation, we subjected the hiPSC-derived hematopoietic progenitors to colony formation assays to assess their clonogenic potential. Additionally, we established long-term cultures to investigate the effects of the KMT2A::MLLT3 fusion with or without the NRAS G12D mutation on cellular proliferation. In our study, we observed the outgrowth of hiPSC-derived hematopoietic progenitors expressing KMT2A::MLLT3, which was further accelerated by the constitutive NRAS G12D mutation. Colony formation assays showed a significant increase in the number and size of atypical colonies consisting mainly of aberrant myeloid progenitors, further enhanced by the introduction of the NRAS G12D mutation. These findings suggest that the KMT2A::MLLT3 fusion and the NRAS G12D mutation are synergistic and together result in increased proliferation and blocked differentiation of hiPSC-derived progenitors, thereby recapitulating two major hallmarks of AML.

We used recent gene editing strategies to introduce both the *KMT2A::MLLT3* fusion and *NRAS* G12D mutation into hiPSCs. Employing *in trans* paired nicking, we successfully knocked in the 3’ part of *MLLT3* together with a fluorescent reporter and a removable selection marker into the endogenous *KMT2A* locus, which resulted in expression of *KMT2A::MLLT3* transcripts. The main advantage of this gene editing approach is that the fusion is expressed under the control of endogenous regulatory elements, avoiding potentially unphysiological ectopic

expression of the fusion protein. In addition, the Cas9^{D10A} nickase utilized for *in trans* paired nicking produces SSBs instead of DSBs, thereby promoting HDR rather than error-prone NHEJ and reducing the likelihood of undesirable lesions at the target site (Chen et al., 2017). Furthermore, it has been demonstrated that Cas9-induced DSBs have a detrimental impact by triggering TP53-mediated apoptosis in hiPSCs, thus constraining the effectiveness of CRISPR/Cas9-mediated gene editing (Ihry et al., 2018). Although we selected the nucleofected hiPSC pool twice, we still had to screen almost 200 clones to retrieve only two *KMT2A::MLLT3* knock-in clones, with one clone requiring additional FACS sorting because of its admixture with wild-type cells. Although DNA nicks promote HDR over NHEJ, the majority of SSBs are repaired through the base excision repair pathway, thereby limiting editing efficiencies, which might explain the low number of gene-edited clones (Abbotts & Wilson, 2017). According to Western blot analysis, we could detect the fusion protein only in one of the knock-in clones in only one experiment. Since fusion gene expression is under the control of the endogenous gene regulatory elements of *KMT2A*, it is conceivable that the fusion protein is expressed at very low levels, rendering its detection by Western blot difficult. This would also explain why we were unable to detect neither the fusion nor the WT protein in the other experiments. Another explanation may involve suboptimal experimental conditions such as incomplete transfer of these large proteins or poor antibody performance, which would require further optimization.

So far, no hiPSC models that express both the KMT2A::MLLT3 fusion and a constitutive NRAS G12D mutation have been described. Therefore, using a novel PE system, we intended to introduce the *NRAS* G12D mutation into one of the *KMT2A::MLLT*3 knock-in clones and parental WT hiPSC cells to investigate the effects of the NRAS G12D variant on cellular transformation. The main advantage of the PE system is that it can be used to induce a broad range of highly precise edits, such as single base changes with reduced off-target effects, since prime editing also relies on DNA nicking and avoids DSBs (Anzalone et al., 2019). Unfortunately, we were unable to detect the intended editing in our hiPSCs. As the PE system is relatively new, it requires optimization of experimental parameters such as the length and sequence of the RT template and primer binding site of the pegRNA to achieve high editing efficiencies (Chen et al., 2021; Doman et al., 2022; Zhao et al., 2023). Since time and resources to optimize prime editing were limited, we resorted to knock-in of a *NRAS* homology harboring the G12D mutation into the *NRAS* locus via *in trans* paired nicking. However, using this approach did not result in successful gene editing in our generated cell bulks. It is important

to emphasize that bulk cell analyses using PCR, Sanger sequencing or restriction digestion cannot effectively detect gene-edited cells if they constitute only a minor fraction of the population. Time- and resource-intensive next-generation sequencing as well as digital PCR offer a viable and sensitive alternative. But successful detection would still require screening a large number of clones. With the aim to increase the fraction of gene-edited cells, we next employed a selection-based approach. For this purpose, we again used *in trans* paired nicking to knock-in *NRAS* exons 2-7 containing the G12D mutation into *NRAS*, together with a removable selection cassette. Since knock-in editing often faces low efficiencies as compared to knock-out approaches, enrichment by introduction of selection or reporter cassettes is highly recommended (Reuven & Shaul, 2022).

Previous leukemia models often relied on vector-based ectopic overexpression of the fusion protein, which may cause random integration artefacts, bypass regulatory mechanisms and often does not accurately reproduce the physiological expression levels in cancer cells (Alonso & Dow, 2021; Gaussmann et al., 2007; Heisterkamp et al., 1991). The CRISPR/Cas9 gene editing technology not only facilitates knock-in approaches but also the establishment of cell lines harboring chromosomal translocations (Brunet & Jasin, 2018; Torres et al., 2015). Therefore, we also tried to utilize this system to introduce a t(9;11)(p22;q23) giving rise to the *KMT2A::MLLT3* rearrangement. Recreating the actual translocation would fully recapitulate the situation found in human leukemia, including the location of all gene regulatory regions up and downstream of the breakpoints. Using two different Cas9 orthologs with gRNAs targeting *KMT2A* and *MLLT3*, we intended to induce the translocation in our hiPSCs by interchromosomal DSB repair via NHEJ. Because such an event is expected to be highly unlikely, gene editing efficiencies are very low and would require screening of a large number of clones. Thus, we adopted the approach from Vanoli et al. and inserted a removable selection cassette at the translocation breakpoint (Vanoli et al., 2017). We succeeded in generating at least one translocation-positive clone that expressed *KMT2A::MLLT3* transcripts. However, we detected intronic deletions in both remaining *KMT2A* and *MLLT3* WT alleles of this clone, representing NHEJ-induced off-target effects of the conventional Cas9 enzymes, which might impair gene function. We refrained from conducting further experiments with the positive t(9;11)(p22;q23) clone because it appeared to be impure. Moreover, this clone constitutively expressed the fusion protein which might interfere with early differentiation. While the oncogenic potential of cells expressing KMT2A::MLLT3 from the actual t(9;11)(p22;q23)

compared to knock-in cells was not assessed, we demonstrated that the CRISPR/Cas9 system is capable to introduce one of the most prevalent translocations in pediatric AML into hiPSCs.

The establishment of long-term liquid cultures of hematopoietic progenitors obtained by *in vitro* differentiation of gene edited hiPSCs showed outgrowth of KMT2A::MLLT3-expressing cells only in one of several attempts. As this fusion alone has been reported to be sufficient to transform HSCs, we expected that its expression would always result in increased proliferation (Stavropoulou et al., 2016). One plausible explanation why this was not the case could be that HSCs require a synergistic combination of biomechanical cues to adequately proliferate and grow (Li et al., 2021). Such cues are provided by the bone marrow environment and are therefore absent in liquid culture. This would also explain why we consistently observed increased clonogenicity of all KMT2A::MLLT3-expressing clones in our colony formation assays in semi-solid media, which might have provided sufficient biomechanical cues. Another reason might be that only a small fraction of hiPSC-derived progenitors expresses the fusion protein at an optimal level to initiate transformation and escape oncogene induced apoptosis. As we repeatedly obtained low cell yields after hematopoietic differentiation, these might have been too low to suffice for outgrowth. Furthermore, KMT2A::MLLT3 mainly drives the upregulation of genes involved in self-renewal such as *HOXA9* and *MEIS1* (Collins & Hess, 2016; Kumar et al., 2009), which we were unable to detect in our hiPSC-derived progenitors. Another explanation for the inconsistency of outgrowth in our liquid culture experiments could be the acquisition of secondary mutations. While the genomic stability of hiPSCs is similar to that of other cell lines, prolonged culture of pluripotent stem cells has been shown to increase mutation rates, especially under suboptimal culture conditions (Busuttil et al., 2003; Jacobs et al., 2016). Furthermore, oncogenic stress induced by KMT2A::MLLT3 may increase the mutation rate by which this clone acquired a secondary hit, promoting its outgrowth (Kuijk et al., 2020; Papapetrou, 2019). Whole exome or genome sequencing could shed light on the spectrum of mutations and genetic lesions acquired during prolonged culture.

Full-blown AML often develops as a result of a combination of type I and type II mutations altering both differentiation and proliferation (de Rooij et al., 2015). Indeed, introduction of the constitutive NRAS G12D mutation into KMT2A::MLLT3 expressing hiPSCs resulted in enhanced cell outgrowth upon hematopoietic differentiation, an effect that was not observed if the NRAS G12D mutation was present alone. These results are coherent with murine *in vivo* experiments in which secondary activating mutations accelerated KMT2A::MLLT3-driven leukemia onset (Hyrenius-Wittsten et al., 2018). Moreover, analysis of KMT2A::MLLT3-

expressing cells showed that these had frequently acquired secondary mutations, although mutation rates in mice are higher than in humans (Hyrenius-Wittsten et al., 2018; Milholland et al., 2017; Perlman, 2016). Our data suggests that KMT2A::MLLT3 is required to initiate AML, and secondary NRAS mutations lead to disease progression.

One of the major hallmarks of AML is aberrant and arrested differentiation, which in most cases gives rise to a heterogeneous population of leukemic blasts containing highly clonogenic LSCs (Marchand & Pinho, 2021). To investigate whether KMT2A::MLLT3-expressing hiPSC-derived hematopoietic progenitors show an increased clonogenicity and/or blocked differentiation, we performed colony formation assays. Our results show that the introduction of the fusion gene resulted in a substantial increase in number and size of undifferentiated granulocytic/monocytic-like colonies compared to other myeloid colony types. This suggests that KMT2A::MLLT3 skews differentiation toward the monocytic lineage mimicking a FAB M5 phenotype. This observation is consistent with *KMT2A::MLLT3* being more frequently present in FAB M5 than M7 AML (Quessada et al., 2021). Furthermore, expression of KMT2A::MLLT3 in hematopoietic progenitors derived from hiPSCs has been shown to interfere with myelomonocytic development (Heuts et al., 2023). Flow cytometric analysis of myeloid surface markers may help to confirm whether our hiPSC-derived hematopoietic progenitors resemble monocytic blasts.

Cytokines play an important role in normal hematopoiesis as well as AML leukemogenesis. It has been shown that AML blasts, similar to healthy HSCs, require an adequate cytokine milieu to develop and proliferate both *in vivo* and *in vitro* (Luciano et al., 2022). We aimed to investigate cytokine dependency of our gene edited hiPSC-derived hematopoietic progenitors by cultivating them with different cytokine combinations. Our results show that KMT2A::MLLT3-expressing cells survive in media supplemented with GM-CSF and IL-3, which are both essential for the maintenance and differentiation of granulocytic and monocytic progenitors (Metcalf, 2008). Both GM-CSF and IL-3 serum levels are elevated in pediatric AML (Elbaz & Shaltout, 2000). Moreover, *in vitro*, GM-CSF promotes the outgrowth of AML blasts while inducing differentiation of healthy HSCs (Schmetzer et al., 1999). The same has been shown for IL-3, which additionally promoted the establishment of autocrine loops that further enhanced proliferation (Nowak et al., 1999). Collectively, our hiPSC-derived KMT2A::MLLT3-expressing progenitors exhibit dependency for AML-related cytokines. We assessed cytokine dependency only using KMT2A::MLLT3 clones with wild-type NRAS, and further analysis of

cells harboring both lesions may yield novel insights into the modulatory effects of RAS pathway activation on cytokine-dependent growth.

As expected, clonogenicity was further exacerbated by constitutive expression of G12D-mutated NRAS. As previously shown, rapidly proliferating myeloid progenitors are more susceptible to KMT2A::MLLT3-driven transformation compared to resting or quiescent progenitors (Chen et al., 2019). Since NRAS G12D results in accelerated proliferation of myeloid progenitors, it might promote the transformation of the affected cells, further supporting the argument that the synergy of fusion protein expression and RAS pathway activation promotes transformation (Wang et al., 2013). These findings suggest that our hiPSC model recapitulates the increased clonogenicity of leukemic stem cells, indicating that it might be suitable to screen for drugs, which target this crucial cell population. Inhibition of menin function has been shown to induce differentiation of KMT2A fusions expressing leukemic blasts, offering a viable treatment option for AML patients (Issa et al., 2023). By utilizing menin inhibitors in functional assays, it will be possible to investigate the applicability of our KMT2A::MLLT3 model for drug testing. However, in our hematopoietic differentiation experiments, we frequently observed low yields of hiPSC-derived hematopoietic progenitors despite using a standardized serum and feeder free monolayer-based differentiation kit. This poses a potential challenge for downstream analyses, which require high cell numbers, such as drug screens. This emphasizes the importance of optimizing culture conditions and differentiation protocols to enhance cell yields.

Oncogene addiction is defined as the dependency of cancer cells on oncoproteins or deregulated pathways to survive and grow (Weinstein & Joe, 2008). Many anti-cancer therapies rely on this principle to effectively target cells, for example, by inhibition of the given oncoprotein. We assessed oncogene addiction of our hiPSC-derived hematopoietic progenitors by dTAG-13-induced degradation of the KMT2A::MLLT3 fusion protein. Rather unexpectedly, dTAG-13 treatment did not completely reverse the KMT2A::MLLT3-induced increase in clonogenicity. DTAG-13 treated cells still produced higher colony and cell numbers than wild-type controls, but at least less than untreated cells. Although Western blot analysis clearly showed degradation of the tagged KMT2A::MLLT3 fusion protein upon dTAG-13 treatment, its sensitivity might not be sufficient to detect remnant low protein amounts and verify complete protein degradation. “Leakiness” of PROTAC systems has been reported, in which high concentrations of linker compound led to saturation of the system (hook effect), hampering the degradation of the target protein and resulting in protein activity despite

treatment (Pettersson & Crews, 2019). Optimization of protein degradation may be achieved by applying various concentrations of dTAG-13 or alternative compounds, such as dTAG-v1. However, as we did not ectopically overexpress the fusion protein and its basal expression level appears to be rather low, our findings suggest that already minute quantities suffice to initiate transformative effects, indicating a hit-and-run mechanism.

In addition, we investigated changes in gene expression by assessing *HOXA9* expression in hiPSC-derived progenitors via RT-qPCR. Intriguingly, we were unable to detect KMT2A::MLLT3-mediated induction of *HOXA9* in the hematopoietic progenitors. Although *HOXA9* upregulation is a key feature of *KMT2A*-rearranged AML, a *KMT2A::AFF1*-positive subgroup exists in infant ALL, which is characterized by low *HOXA9* but high *IRX1* expression (Symeonidou & Ottersbach, 2021). Furthermore, we assessed expression levels of *DUSP6* and *SPRY2,* which are targets and attenuators of RAS signaling known to constitute a negative feedback loop (Muhammad et al., 2018; Shojaee et al., 2015; Tsavachidou et al., 2004). However, we detected no substantial change in the expression levels of *DUSP6* and *SPRY2* in our hiPSC-derived NRAS G12D hematopoietic progenitors. Importantly, our prelimary gene expression analysis was limited to a few predefined target genes, and only whole transcriptome (WTS) sequencing will reveal the global gene expression profile of other targets, such as the entire HOX cluster genes or *MEIS1*. At least so far, we could neither confirm that our hiPSC-derived progenitors show gene expression profiles similar to those of AML patient samples, nor could we confirm the expression of *KMT2A::MLLT3* protein after hematopoietic differentiation. However, we have stored hiPSC-derived cells at the progenitor stage and after prolonged cell culture for global gene expression analysis, which in the future will be used to track dynamic changes in gene expression upon KMT2A::MLLT3-induced transformation.

In summary, our study revealed that the introduction of *KMT2A::MLLT3* into hiPSCs, followed by their *in vitro* differentiation, led to elevated clonogenicity and self-renewal of the derived hematopoietic progenitors. This effect was further enhanced by the introduction of a constitutively active NRAS G12D mutation. Our findings suggest that the observed AML-like phenotype aligns more closely with FAB type M5 than M7. Although gene expression of the hiPSC-derived hematopoietic progenitors was not closely examined, our study demonstrates that our hiPSC model system effectively captures two key characteristics of AML, namely impaired differentiation and uncontrolled proliferation.

5. Abstract

Pediatric acute myeloid leukemia (AML) is a rare subtype of leukemia, which is characterized by uncontrolled proliferation and blocked differentiation of myeloid progenitors. This malignancy is mainly caused by chromosomal rearrangements, which lead to the expression of oncogenic fusion proteins driving leukemogenesis. *KMT2A* rearrangements are found in about 20% of AML patients and are generally associated with an inferior outcome. One of the most frequent fusion genes, *KMT2A::MLLT3*, results from a t(9;11)(p22;q23) translocation and is often accompanied by secondary mutations in the RAS signaling pathway. Although several downstream targets of the KMT2A::MLLT3 fusion protein are already well-known, it is still poorly understood at what stage of hematopoietic development and how this fusion drives malignant transformation. As patient material only captures the late stage of full-blown leukemia and mouse models do not fully reflect human disease, innovative model systems faithfully recapitulating leukemia development are required. Human induced pluripotent stem cells (hiPSCs) offer a promising alternative to other model systems since they can be used to address both the early stages of leukemia development and disease progression. Based on previous studies showing that expression of leukemia-associated fusion oncoproteins in hiPSCs upon their *in vitro* differentiation into hematopoietic progenitors gives rise to leukemia-like cells resembling the human disease, we hypothesized that this system may also be utilized to accurately model KMT2A::MLLT3-driven leukemia.

The aim of this study was to apply genome editing technologies to generate hiPSC lines harboring the *KMT2A::MLLT3* fusion and/or the RAS pathway-activating NRAS G12D mutation. We have successfully generated the respective hiPSC clones and subsequently *in vitro* differentiated them toward hematopoietic stem and progenitor cells. Assessment of their clonogenic potential by colony formation assays revealed strongly increased self-renewal of KMT2A::MLLT3-expressing progenitors compared to hiPSC carrying only the NRAS G12D mutation and wild-type hiPSCs. Notably, through the expression of KMT2A::MLLT3 fusion protein, hematopoietic progenitors were maintainable in long-term liquid cultures, and outgrowth was further exacerbated by the additional NRAS-activating mutation indicating cellular transformation. Taken together, we have established an hiPSC-based KMT2A::MLLT3-driven leukemia model, which has the potential to recapitulate leukemia development.

Die pädiatrische akute myeloische Leukämie (AML) ist eine seltene Unterform der Leukämie, die durch unkontrollierte Proliferation und blockierte Differenzierung von myeloischen Vorläuferzellen charakterisiert ist. Diese bösartige Erkrankung wird hauptsächlich durch chromosomale Alterationen verursacht, die zur Expression von onkogenen Fusionsproteinen führen, welche für die Entstehung der Leukämie verantwortlich sind. Chromosomale Umstrukturierungen des *KMT2A* Gens werden bei etwa 20 % der AML-Patienten gefunden und sind im Allgemeinen mit einem eher ungünstigen Krankheitsverlauf verbunden. Eines der häufigsten Fusionsgene, *KMT2A::MLLT3*, resultiert aus einer t(9;11)(p22;q23)-Translokation und wird oft von sekundären Mutationen im RAS-Signalweg begleitet. Obwohl einige durch KMT2A::MLLT3 regulierte Signalwege bereits bekannt sind, ist immer noch unzureichend geklärt, in welchem Stadium der hämatopoetischen Entwicklung und wie dieses Fusionprotein die maligne Transformation der Zellen antreibt. Da Patientenmaterial nur das Spätstadium der bereits vollständig entwickelten Leukämie abbildet und Mausmodelle die Krebserkrankung des Menschen nicht vollständig widerspiegeln, werden innovative Modellsysteme benötigt, welche die Leukämieentwicklung präzise reproduzieren.

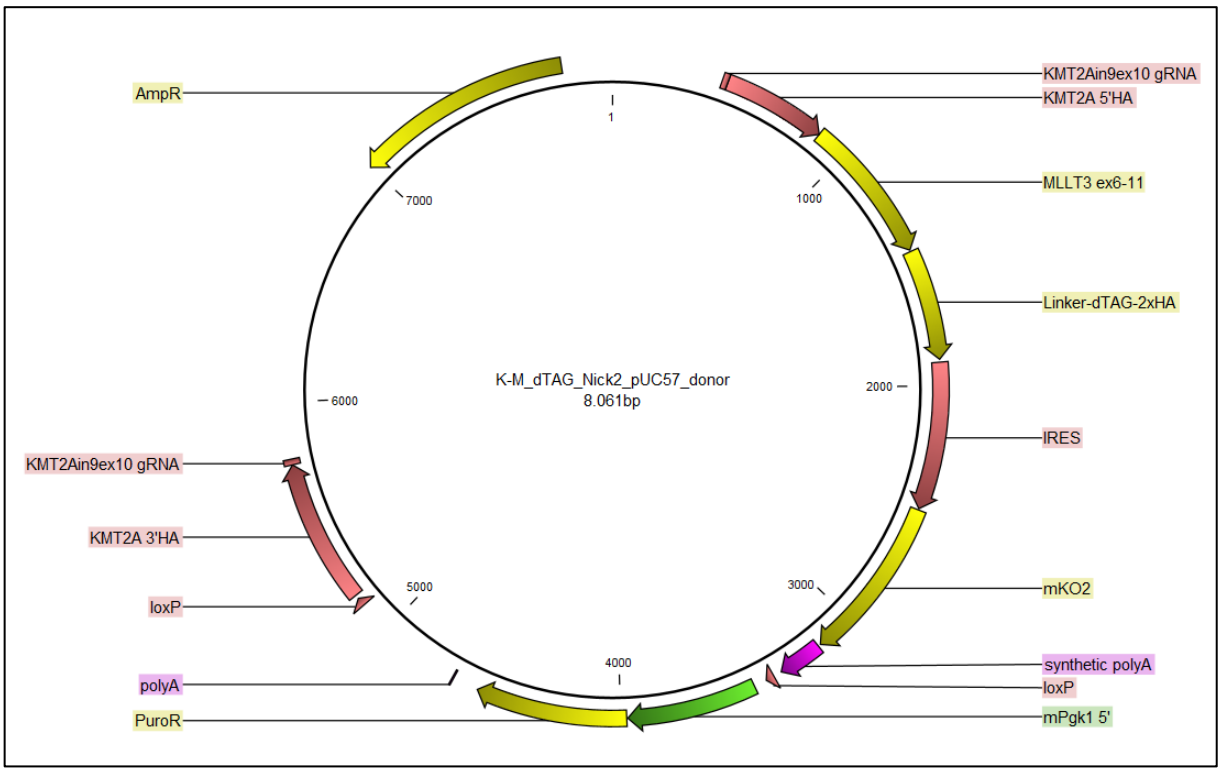
Humane induzierte pluripotente Stammzellen (hiPSZ) bieten eine vielversprechende Alternative zu anderen Modellsystemen, da sie dazu verwendet werden können, sowohl die frühen Stadien der Leukämieentwicklung als auch das Fortschreiten der Erkrankung zu untersuchen. Aufgrund früherer Studien, in denen gezeigt wurde, dass die Expression von Leukämie-assoziierten Fusionsonkoproteinen in hiPSZ, nach ihrer *in vitro* Differenzierung in hämatopoetische Vorläuferzellen zur Bildung von Leukämie-ähnlichen Zellen führt, die der menschlichen Krankheit ähneln, haben wir die Hypothese aufgestellt, dass dieses System auch für die Modellierung der von KMT2A::MLLT3-getriebenen Leukämie verwendet werden kann.

Ziel dieser Studie war, mittels Genomeditierung hiPSZ-Linien herzustellen, welche die KMT2A::MLLT3-Fusion und/oder eine den RAS-Signalweg aktivierende Mutation tragen. Wir haben die entsprechenden hiPSZ-Klone erfolgreich generiert und sie anschließend *in vitro* in hämatopoetische Stamm- und Vorläuferzellen differenziert. Die Bewertung ihres klonogenen Potenzials durch Koloniebildungstests zeigte eine deutlich erhöhte Selbsterneuerung von KMT2A::MLLT3-exprimierenden Vorläuferzellen im Vergleich zu entsprechenden Wildtyp-Zellen oder solchen, die nur eine NRAS G12D-Mutation tragen. Bemerkenswert ist, dass durch die Expression des KMT2A::MLLT3-Fusionsproteins das Wachstum der hämatopoetischen Vorläuferzellen in Flüssigkulturen langfristige aufrechterhalten werden konnte und durch die

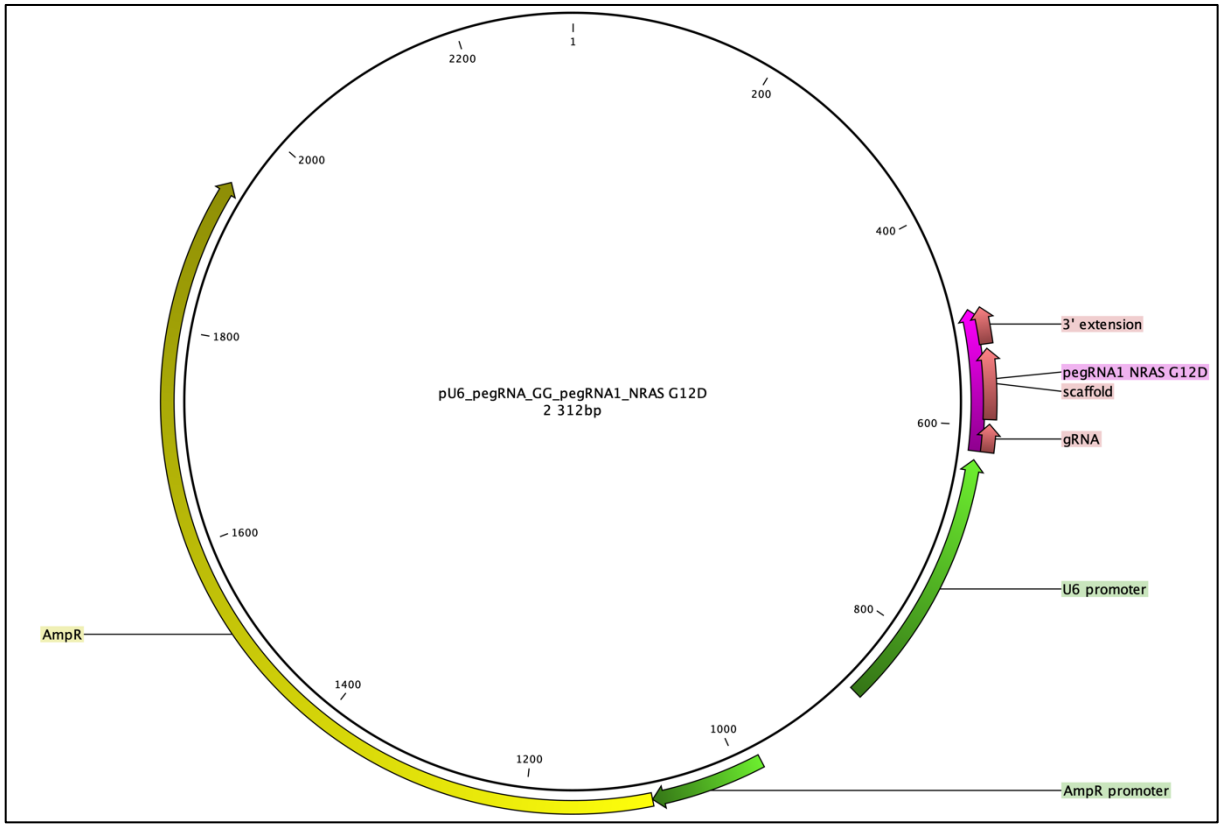
zusätzliche NRAS-aktivierende Mutation noch beschleunigt wurde, was auf eine zelluläre Transformation hinweist. Zusammenfassend haben wir ein hiPSZ-basiertes KMT2A::MLLT3-getriebenes Leukämiemodell etabliert, welches das Potenzial hat, die Leukämieentwicklung im Reagenzglas zu rekapitulieren.

6. Supplements

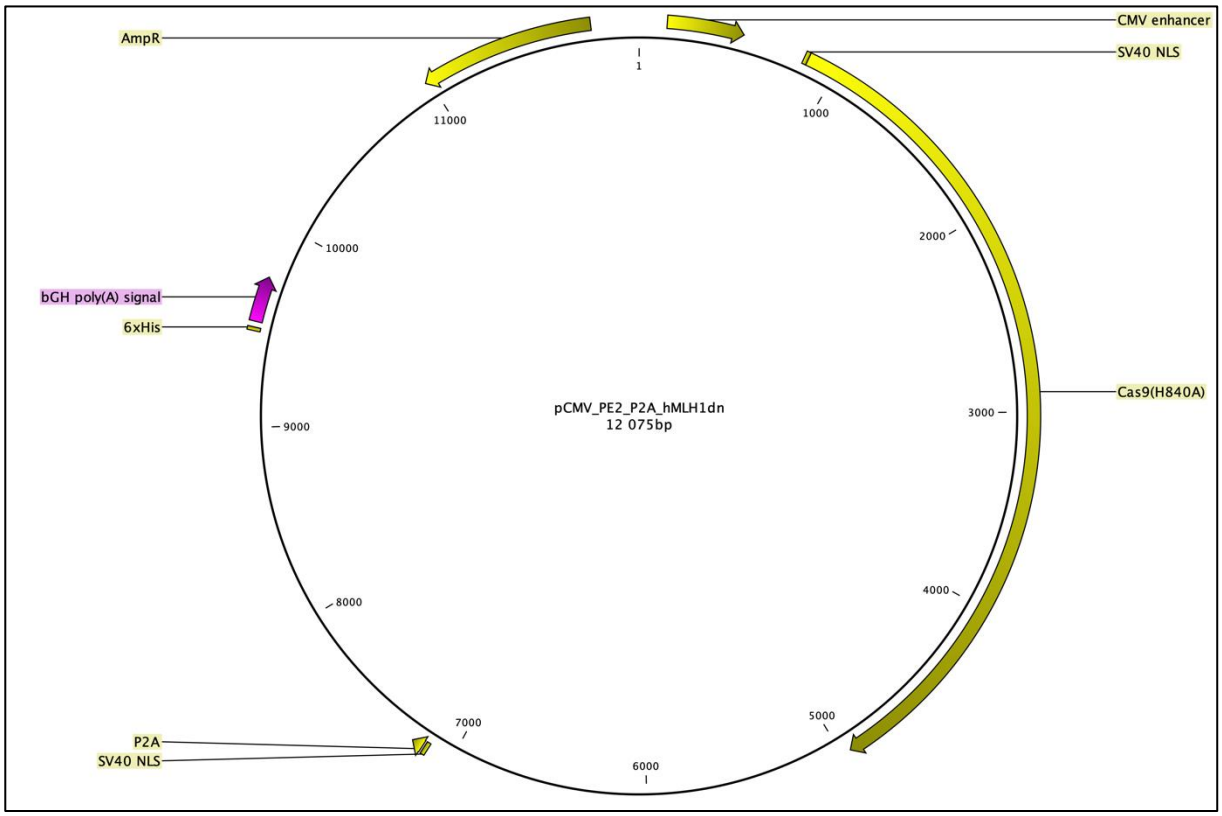
6.1 Gene editing constructs



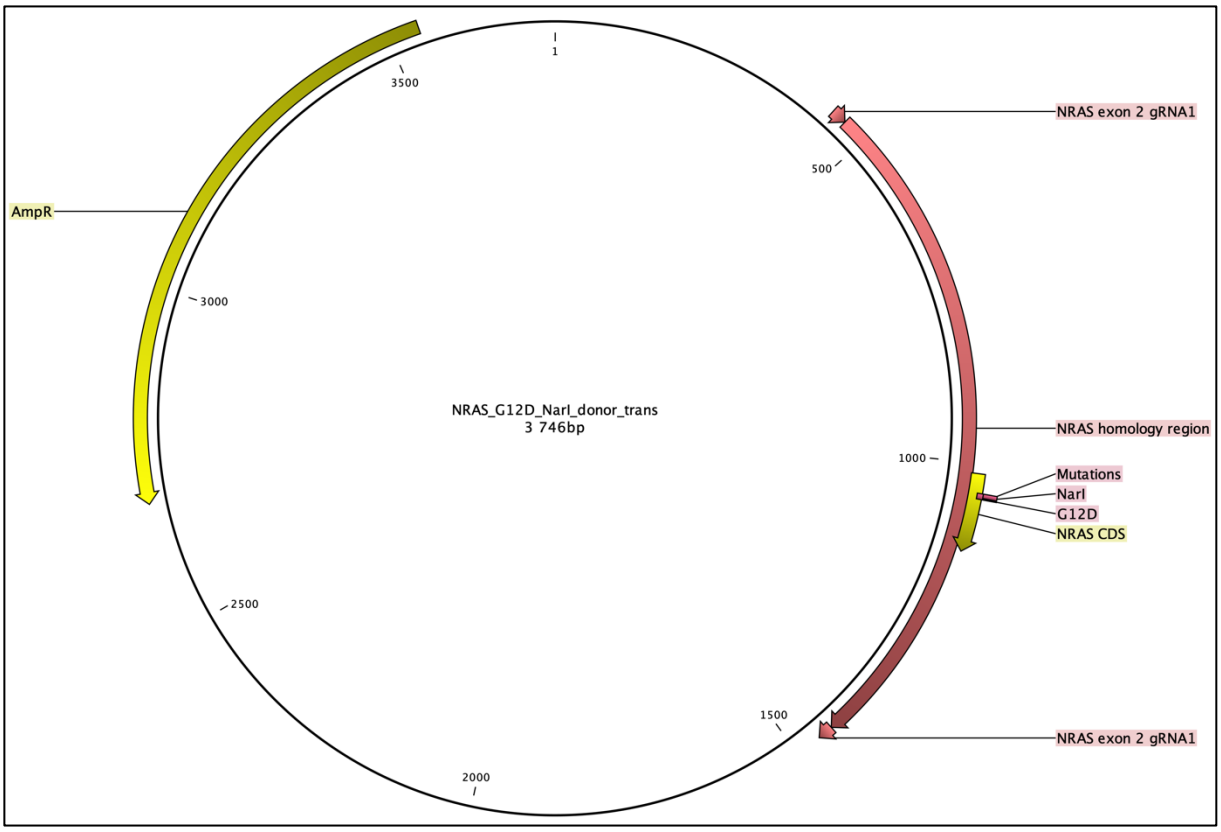
Supplementary Figure 1: Knock-in donor for the KMT2A::MLLT3 fusion (K-M_dTAG_Nick2_pUC57_donor). Homology arm (HA), exon (ex), intron (in), degrader Tag (dTAG), internal ribosome entry site (IRES), mKusabira-Orange2 (mKO2), phosphoglycerate kinase 1 5' promotor region (mPgk1 5'), puromycin resistance cassette (PuroR), ampicillin resistance (AmpR).



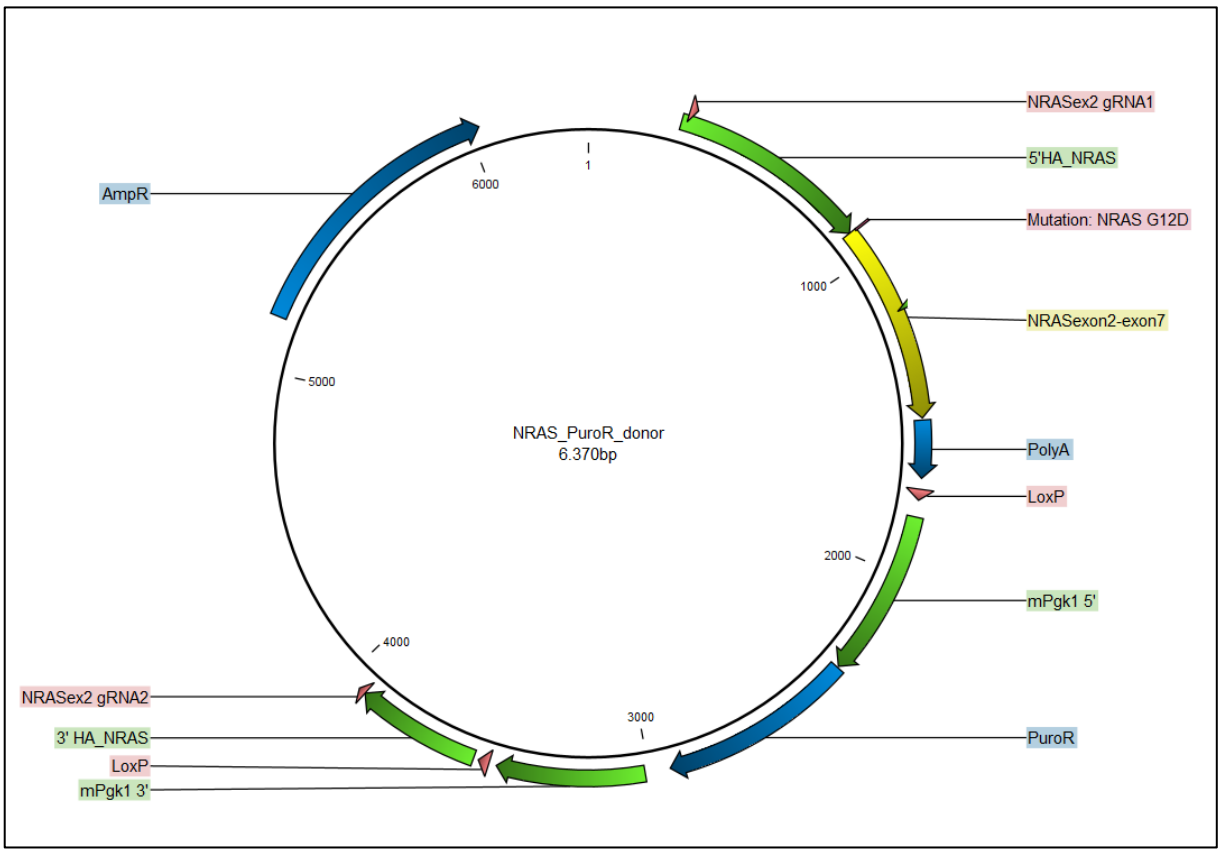
Supplementary Figure 2: PegRNA vector used for prime editing of the *NRAS* G12D mutation (pU6_pegRNA_GG_pegRNA1_NRAS G12D). Ampicillin resistance (AmpR).



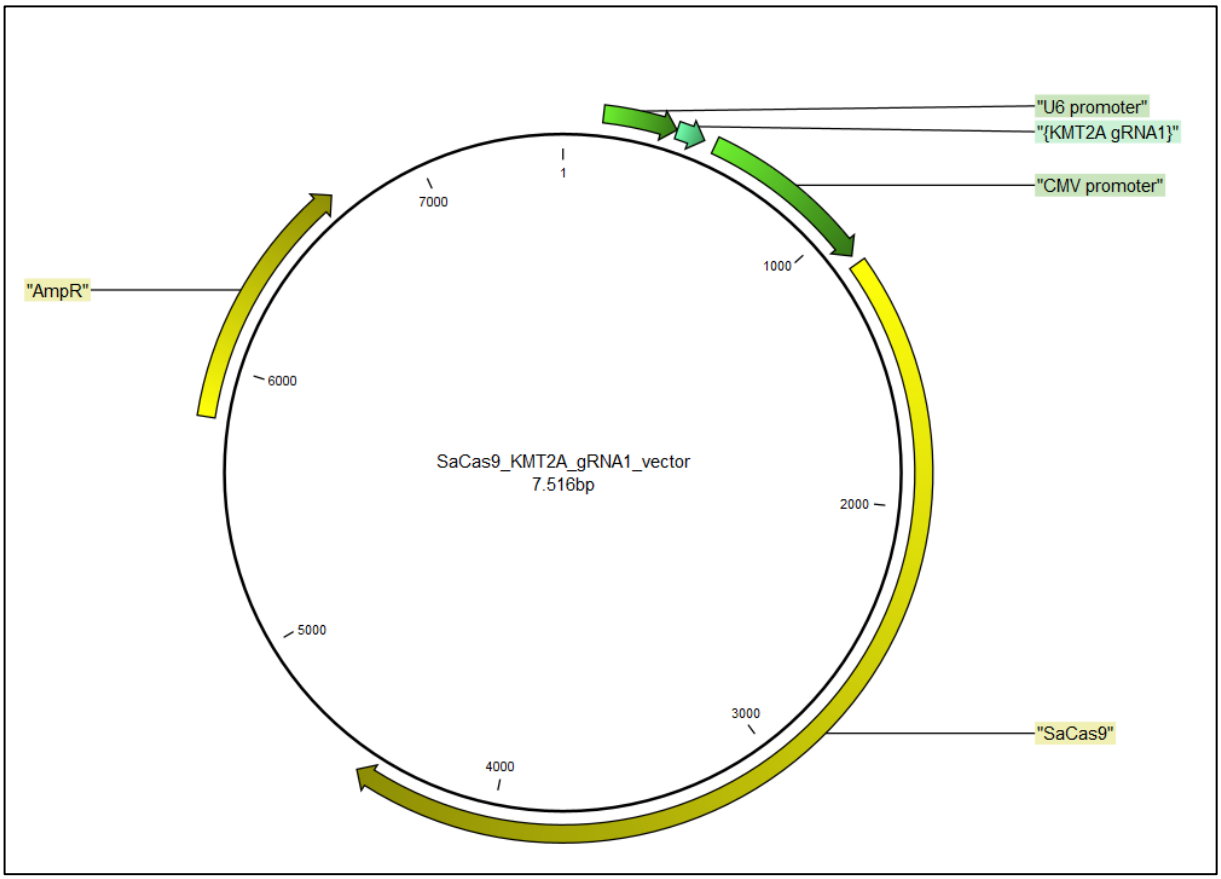
Supplementary Figure 3: Prime editing construct containing the nickase Cas9 H840A and MLH1dn genes. Cytomegalovirus (CMV), simian virus 40 Nuclear localization signal (SV40 NLS), self-cleaving peptide sequence P2A (P2A), 6x histidine (6xHis/His tag), bovine growth hormone polyadenylation signal (bGH poly(A) signal), ampicillin resistance (AmpR).



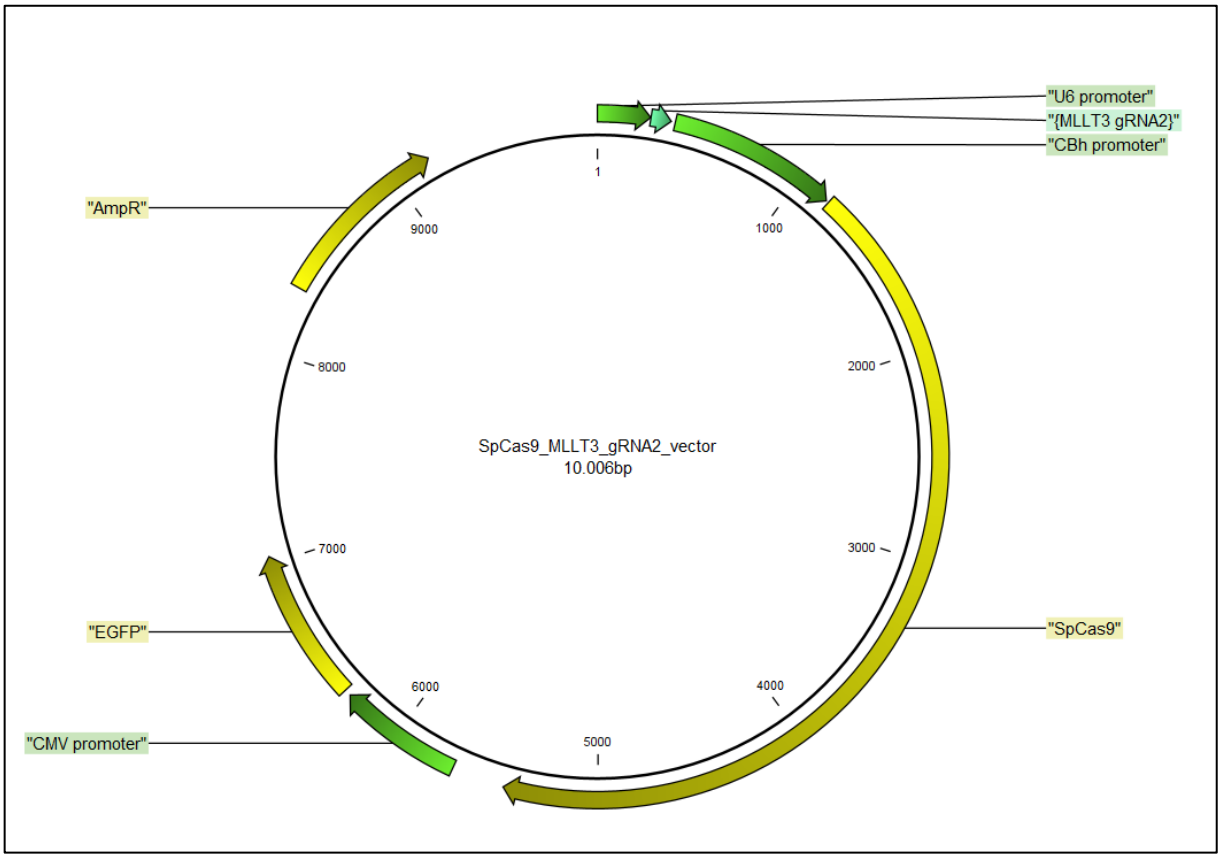
Supplementary Figure 4: *In trans* paired nicking *NRAS* G12D donor containing the *NRAS* homology (NRAS_G12D_NarI_donor_trans). Coding sequence (CDS), ampicillin resistance (AmpR).



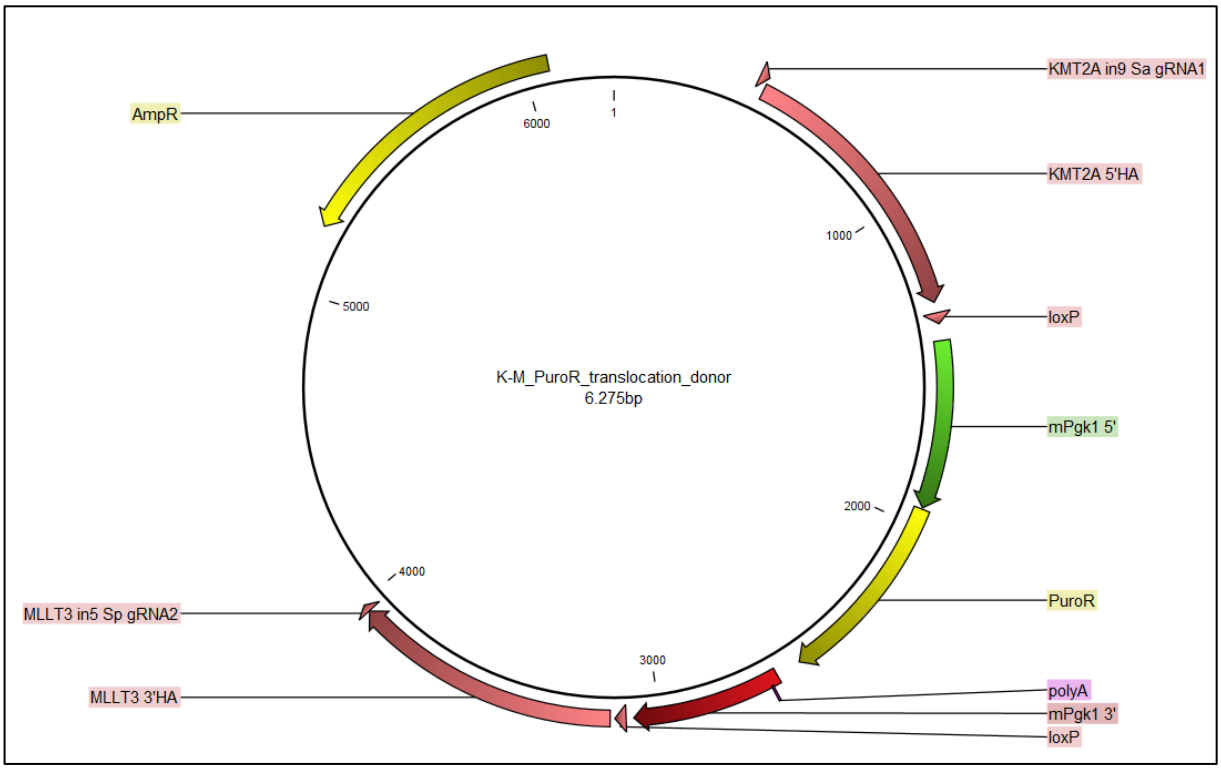
Supplementary Figure 5: Knock in donor for the selection based *NRAS* G12D knock-in strategy (NRAS_PuroR_donor). Exon (ex), homology arm (HA), phosphoglycerate kinase 1 5' promotor region (mPgk1 5'), puromycin resistance cassette (PuroR), phosphoglycerate kinase 1 3' promotor region (mPgk1 3'), ampicillin resistance (AmpR).



Supplementary Figure 6: Vector for the SaCas9 enzyme and *KMT2A* gRNA (SaCas9_KMT2A_gRNA1_vector). Cytomegalovirus (CMV), *Staphylococcus aureus* Cas9 (SaCas9), ampicillin resistance (AmpR).



Supplementary Figure 7: Vector for the SpCas9 enzyme and *MLLT3* gRNA (SpCas9_MLLT3_gRNA2_vector). Chicken β-actin (CBA), CBA promoter with cytomegalovirus enhancer and hybrid intron (CBh promotor), *Streptococcus pyogenes* Cas9 (SpCas9), Cytomegalovirus (CMV), enhanced green fluorescent protein (EGFP), ampicillin resistance (AmpR).



Supplementary Figure 8: Puromycin resistance cassette donor for the *KMT2A::MLLT3* translocation (K-M_PuroR_translocation_donor). Intron (in), homology arm (HA), murine phosphoglycerate kinase 1 5' promotor region (mPgk1 5'), puromycin resistance cassette (PuroR), phosphoglycerate kinase 1 3' UTR region (mPgk1 3'), ampicillin resistance cassette.

6.2 Buffer recipes

The LB-Agar used for bacterial culture was prepared by mixing 10 g of tryptone, 5 g of yeast extract (all from ThermoFisher Scientific), 10 g of NaCl (Sigma-Aldrich) and 15 g of agar (ThermoFisher Scientific) in 1 L of ddH_2O followed by autoclavation. After autoclavation the solution was supplemented with 100 µg/ml Ampicillin (Merck) and cast into petri dishes.

For the preparation of the 6x DNA loading dye, 100 µl of 1 M Tris-HCL pH 7.6, 1.2 ml 0.5 M EDTA, 6 ml glycerol (Merck) and a few grains of Bromphenol blue (Merck), were mixed and adjusted to a volume of 10 ml with ddH_2O.

The 5x SDS loading buffer used for protein isolation and western blot analysis, was prepared by mixing the following ingredients: 2.5 ml of 1 M Tris-HCl (pH 6.8), 1 g SDS (Sigma-Aldrich), 5 ml glycerol (Merck), 0.77 g DTT (PanReac Applichem), a pinch of Bromphenol blue (Merck), adjusted to a volume of 10 ml with ddH_2O. Before use for western blot analysis, the 5x SDS loading buffer was diluted with ddH_2O in a ratio of 1:5.

The 10x Tris/glycine blotting buffer was prepared by mixing 30 g Trizma Base and 144 g glycine (all purchased from Sigma-Aldrich) in ddH_2O, added up to a volume of 1 L followed by autoclavation. Buffer was diluted to a 1x concentration before being used for western blotting.

6.3 List of materials and primers

All the used material and reagents, as well as primers used for genotyping and quantitative PCR, are listed in **Supplementary Table 1** and **Supplementary Table 2**.

Supplementary Table 1: List of used material and reagents

Reagent/Material	Manufacturer	Catalogue number
10x Tris/Glycine/SDS Buffer for Western Blots and Native Gels	Bio-Rad	1610732
4–15% Mini-PROTEAN® TGX Precast Protein Gels, 10-well, 30 µl	Bio-Rad	4561083
Agarose	VWR	ROTH2269
Amersham Protran Premium Western blotting membranes, nitrocellulose	Merck	GE10600003
Ampicillin	Merck/Sigma-Aldrich	A5354
Aprotinin from bovine lung	Merck/Sigma-Aldrich	A6106
Benzonase Nuclease	Merck/Sigma-Aldrich	E1014
Blocking Reagent	Roche/Merck	11096176001
Bromphenol blue	Merck	8122
BsaI-HFv2 (20 U/µl)	New England BioLabs (NEB)	R3733
Buffer R (10X)	ThermoFisher Scientific	BR5
CHIR99021	Sigma-Aldrich	SML1046
Chloroform	Merck	107024
CloneR™	STEMCELL Technologies	05888
Deoxynucleotide Triphosphates (dNTPs) dCTP	Promega	U1225
Deoxynucleotide Triphosphates (dNTPs) dGTP	Promega	U1215
Deoxynucleotide Triphosphates (dNTPs) dTTP	Promega	U1235
Dithiothreitol (DTT)	AppliChem	7G011577
DPBS, no calcium, no magnesium	ThermoFisher Scientific	14190144
DS-11 FX µVolume Spectrophotometer	DeNovix	31DS-11FX-B
dTAG-13	Tocris Bioscience	6605
dTAG-13	Sigma-Aldrich/Merck	SML2601
dTAGv-1	Tocris Bioscience	6914
EcoRI (10 U/µl)	ThermoFisher Scientific	ER0271
EDTA (Titriplex III)	Merck/Sigma-Aldrich	108418
EndoFree Plasmid Maxi Kit (10)	QIAGEN	12362
Ethidium bromide	Sigma-Aldrich/Merck	1.11608
EVOS XL Core Imaging System	ThermoFisher Scientific	AMEX1000

Fetal Bovine Serum, qualified, heat inactivated, E.U.-approved	ThermoFisher Scientific	10082147
Gibco DMEM/F-12	ThermoFisher Scientific	11320033
Gibco episomal hiPSC line	ThermoFisher Scientific	A18945
Gibco RPMI 1640 Medium	ThermoFisher Scientific	21875034
Glycerol	Merck/Sigma-Aldrich	818709
Glycerol	Merck	8.81879100
Glycine	Sigma-Aldrich/Merck	G7126
Glycogen, RNA grade	Roche	10901393001
HotStarTaq DNA Polymerase (1000 U)	QIAGEN	203205
Human Stem Cell Nucleofector Kit 2	Lonza	VPH-5022
IMDM	ThermoFisher Scientific	12440053
IMDM, GlutaMAX Supplement	ThermoFisher Scientific	31980030
Invitrogen TRIzol Reagent	ThermoFisher Scientific	15596026
Leupeptin	Merck/Sigma-Aldrich	L2884
Lonza Nucleofector Transfection 2b Device	Lonza	AAB-1001
MACS Dead Cell Removal Kit	Miltenyi Biotec	130-090-101
Magnesium chloride hexahydrate	Merck	105833
Matrigel Membrane Matrix	Corning	354234
MegaCult-C Complete Kit with Cytokines	STEMCELL Technologies	04971
MethoCult H4435 Enriched	STEMCELL Technologies	04435
Monarch PCR & DNA Cleanup Kit (5 µg)	NEB	T1030S/L
mTeSR Plus 5X Supplement	STEMCELL Technologies	100-0275
mTeSR Plus Basal medium	STEMCELL Technologies	100-0274
MycoAlert Mycoplasma Detection Kit	Lonza	LT07-318
NarI (5 U/µl)	NEB	R0191
NEB 5-alpha Competent E. coli (High Efficiency)	NEB	C2987
NEBuilder HiFi DNA Assembly Cloning Kit	NEB	E5520
Nonidet P 40 Substitute	Merck/Sigma-Aldrich	74385
NotI (10 U/µl)	ThermoFisher Scientific	ER0592
One Shot Stbl3 Chemically Competent E. coli	ThermoFisher Scientific	C737303
OneTaq® DNA Polymerase	NEB	M0480
Oxoid Agar Bacteriological	ThermoFisher Scientific	LP0011B
Oxoid Tryptone Soya Agar 500g,	ThermoFisher Scientific	CM0129
Oxoid™ Yeast Extract Powder	ThermoFisher Scientific	LP0021B
PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa	ThermoFisher Scientific	26619
Penicillin-Streptomycin (5,000 U/mL)	ThermoFisher Scientific	15070063
Pepstatin A	Merck/Sigma-Aldrich	P5318

Peqlab peqGOLD, DNA Ladder, 100 bp Plus	VWR	25-2020
PMSF	Merck/Sigma-Aldrich	P7626
Ponceau S solution	Sigma-Aldrich/Merck	P7170
Potassium chloride	Merck	104936
Proteinase K	Merck/Sigma-Aldrich	1.24568
Puromycin -dihydrochlorid	Merck	58-58-2
QIAamp DNA Blood Mini Kit	QIAGEN	51104
Recombinant Human EPO	PeproTech	100-64
Recombinant Human G-CSF	PeproTech	300-23
Recombinant Human GM-CSF	PeproTech	300-03
Recombinant Human IL-3	PeproTech	200-03
Recombinant Human IL-6	PeproTech	200-06
Recombinant Human SCF	PeproTech	300-07
Recombinant Human TPO	PeproTech	300-18
S.O.C. Medium	ThermoFisher Scientific	15544034
SDS	Sigma-Aldrich/Merck	75746
SgsI (AscI) (10 U/µl)	ThermoFisher Scientific	ER1891
Sodium chloride	Merck/Sigma-Aldrich	106404
Sodium chloride	Merck/Sigma-Aldrich	106404
STEMdiff Hematopoietic Kit	STEMCELL Technologies	05310
STEMgrid-6	STEMCELL Technologies	27000
StemPro Accutase Cell Dissociation Reagent	ThermoFisher Scientific	A1110501
StemPro™-34 SFM (1X)	ThermoFisher Scientific	10639011
Synthemax II-SC Substrate	Corning	3535
T4 DNA Ligase	Promega	M1801
Tris (hydroxymethyl)aminomethane	Merck/Sigma-Aldrich	108381
TRIS (Trizma) base	Sigma-Aldrich/Merck	T1503
Triton X-100	Merck/Sigma-Aldrich	1.12298.0101
Trypan Blue solution	Sigma-Aldrich	93595
Tween-20	Merck/Sigma-Aldrich	P9416
Water DNase-, RNase-free	Ambion	9914G7
XbaI (10 U/µl)	ThermoFisher Scientific	ER0681
XhoI (10 U/µl)	ThermoFisher Scientific	ER0692
Y-27632 (Dihydrochloride)	STEMCELL Technologies	72308

Supplementary Table 2: List of primers for genotyping and quantitative PCR (all from Microsynth)

Primer name	**Sequence 5'->3'**
ABL1-ENF1003	TGGAGATAACACTCTAAGCATAACTAAAGGT
ABL1-ENR1063	GATGTAGTTGCTTGGGACCCA
Donor_loxP2_for	ACGAAGTTATTAGGTCCCTCTCG
Donor_loxP2_rev	GGTTCTCGAGAGGGACCTAA
DUSP6_ex1_2_F1	CCTGGAAGGTGGCTTCAGTA
DUSP6_ex2_R1	ACCATCCGAGTCTGTTGCAC
GUSB-ENF1102	GAAAATATGTGGTTGGAGAGCTCATT

GUSB-ENR1162	CCGAGTGAAGATCCCCTTTTTA
HOXA9_ex1_F1	GGGGTGACTGTCCCACGCTT
HOXA9_ex2_R1	CCGAGTGGAGCGCGCATGAA
KMT2Aex10in10-R1	ATGCCTTACCTCTACATGCCC
KMT2Aex11-R1	GGAAGGGCTCACAACAGACT
KMT2Aex9-F1	TGCAGGCACTTTGAACATCC
KMT2Aex9-F2	GGAGTCCACAGGATCAGAGTG
KMT2Ain10-R2	AGTTTTTGGTCACTAGAGGGCT
KMT2Ain8-F2	GGTTTCTTCCTTGTTGCTTTTCCCT
KMT2Ain9-F2	AGGTGTTGAAAGAGGAAATCAGCA
KMT2Ain9-R1	ACTGCTCAGGATGGATAGGAATG
KMT2Ain9-R2	CTCTGATACTGCTCAGGATGGAT
mKO2midF2	CGACGGAGTTCTGAAGGGTG
MLLT3ex6-R1	GTTGCCTGGTCTGGGATGGT
MLLT3ex8-R1	GGAGGTTCGTGATGTAGGGG
MLLT3in5-F1	TTGGAACCAACCCAAATGCCCA
MLLT3in5-F2	GCTGGAAACCATCATTCTCAGC
MLLT3in5-F4	GGCGACATTTTCCCTTGACTC
MLLT3in5-R1	TCATGTTGTCTTTTGTTGTGCCC
MLLT3in6-R2	GGATGGGCTGTGTCTATAATCCTT
mPgk1_3UTR_F1	ACCAAATTAAGGGCCAGCTCA
mPgk1proR1	GCCTACCGGTGGATGTGGAAT
NANOG_F1	TTCCTTCCTCCATGGATCTG
NANOG_R1	AAGTGGGTTGTTTGCCTTTG
NRASex1-F1	ATTTTTCCCGGCTGTGGTCC
NRASex1-F3	CGCCGACTGATTACGTAGCG
NRASex3-4-R1	CTCGCTTAATCTGCTCCCTGT
NRASin2-R1	ACAGTCTCGCTACTATGGCCT
NRASin2-R2	GTGTGGTAGGCAGGACAAGTT
NRASin2-R3	ATGGCAACAGGACTTTTACTTGT
OCT4_F1	GAGGAGTCCCAGGACATCAA
OCT4_R1	CATCGGCCTGTGTATATCCC
PuroRmidF2	CAACCTCCCCTTCTACGAGC
SOX2_F1	CAACGGCAGCTACAGCATGATGC
SOX2_R1	CCGTTCATGTAGGTCTGCGAGCTG
SPRY2_ex1_F2	GAGTGTTCATCAGCGGGGAA
SPRY2_ex2_R1	GGAAGTGTGGTCACTCCAGC

List of Figures

List of tables

List of supplementary figures

List of abbreviations

3'HA	3' homology arm
5'HA	5' homology arm
6xHis	6x histidine
AAAAA	poly A
ABL1	Abelson murine leukemia viral oncogene homolog 1
AFDN	Afadin, Adherens Junction Formation Factor
AGM	aorta-gonad-mesonephros
ALL	acute lymphoblastic leukemia
AMKL	acute megakaryoblastic leukemia
AML	acute myeloid leukemia
Amp^R(R)	Ampicillin resistance
APL	acute promyelocytic leukemia
ASXL1	Additional Sex Combs Like 1
AT	AT hook
ATO	arsenic trioxide
ATRA	all-trans retinoic acid
bGH poly(A) signal	bovine growth hormone polyadenylation signal
BPR	breakpoint region
BRD	bromodomain
BSA	bovine serum albumin
C	zinc finger
C/EBPα	CCAAT/enhancer binding protein alpha
CAR-T	Chimeric Antigen Receptor T-cell
Cas9	CRISPR-associated protein 9
CBA	chicken β-actin
CBFB::MYH11	Core Binding Factor Beta::Myosin Heavy Chain 11
CBh promotor	CBA promoter with cytomegalovirus enhancer and hybrid
cDNA	complementary DNA
CDS	coding sequence
CEBPE	CCAAT/Enhancer-Binding Protein Epsilon
CFU	colony forming units
CFU-Mks	megakaryocytic colony forming units
CHIP	chromatin immunoprecipitation
CLP(s)	common lymphoid progenitor(s)
CMP(s)	common myeloid progenitor(s)
CMV	cytomegalovirus
Cre	cyclization recombinase
CRISPR	clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA

CSF(s)	colony stimulating factor(s)
CSF3R	colony Stimulating Factor 3 Receptor.
CXXC	CXXC motif
DNA	deoxyribonucleic acid
DNMT3A	DNA (cytosine-5)-methyltransferase 3A
dNTPs	deoxyribonucleotide triphosphates
DOT1L	Disruptor of Telomeric silencing 1-Like
DPBS	Dulbecco's phosphate buffer saline
DSB(s)	double-stranded DNA break(s)
dTAG	degrader tag
DTT	dithiothreitol
DUSP6	Dual Specificity Phosphatase 6
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
ELL	Elongation Factor for RNA Polymerase II
EPO	erythropoietin
ESCs	embryonic stem cells
EWSR1::WT1	Ewing Sarcoma Breakpoint Region 1::Wilms Tumor 1
ex	exon
FAB	French-American-British
FBS	fetal bovine serum
FISH	fluorescence in situ hybridization
floxed	flanked by Lox-P sites
FLT3(L)	FMS-like tyrosine kinase 3 (ligand)
FLT3-ITD	Fms-Like Tyrosine Kinase 3 - Internal Tandem Duplication
GATA	GATA binding
GATA2	GATA binding protein 2
G-CSF	granulocyte colony stimulating factor
GDP	Guanosine-5'-diphosphate
GEMM	granulocyte-erythrocyte-macrophage-megakaryocyte
GFI1	growth Factor Independent 1 Transcriptional Repressor.
GM	granulocyte/macrophage
GM-CSF	granulocyte-macrophage colony stimulating factor
GMP(s)	granulocyte-macrophage progenitor(s)
gRNA	guide RNA
GSKi	GSK3 inhibitor
GTP	Guanosine-5'-triphosphate
GUSB	Glucuronidase, beta
H(2A/2B/H3/H4)	histone subunit
H3K27me2/3	hHstone 3 Lysine 27 Di- and Trimethylation
H3K4me2/3	Histone 3 Lysine 4 Di- and Trimethylation
H3K9me2/3	Histone 3 Lysine 9 Di- and Trimethylation

HA tag	hemagglutinin epitope-tag
HA(s)	homology arm(s)
HCL	hydrochloric acid
HDR	homology-directed repair
HEM	hematopoietic differentiation
Hox	homeobox
HSC(s)	hematopoietic stem cell(s)
HSPCs	hematopoietic stem and progenitor cells
IDH2	Isocitrate Dehydrogenase 2
IL(s)	Interleukin(s)
IL-3	interleukin-3
IL6R	interleukin-6 Receptor
IL-7	interleukin 7
Il7r	interleukin-7 Receptor
IMDM	Iscoves's Modified Dulbecco's Medium
in	intron
IRES	internal ribosome entry site
IRX1	iroquois homeobox 1
JAK2	Janus Kinase 2
JAK-STAT	janus kinase-signal transducer and activator of transcription
KCl	potassium chloride
KI	knock-in
Kit	KIT Proto-Oncogene, Receptor Tyrosine Kinase
KLF4	kruppel-like factor 4
KLF5	kruppel-Like Factor 5
KMT(s)	histone methyltransferase(s)
KMT2A	histone-lysine N-methyltransferase 2A
KMT2A-PTD	KMT2A partial tandem duplications
KMT2Ar	KMT2A rearrangements
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog
LB	lysogeny broth
LMPP	lymphoid primed multipotent progenitor
LSC(s)	leukemia stem cell(s)
LT-HSC(s)	long-term HSC(s)
Lyl1	lymphoblastic leukemia-derived sequence 1
M-CSF	macrophage colony stimulating factor
MECOM	MDS1 and EVI1 Complex Locus
MEK	Mitogen-Activated Protein Kinase Kinase
MEP(s)	megakaryocyte-erythrocyte progenitor(s)
Metho	MethoCult
$MgCl_2$	magnesium chloride
mKO2	monomeric Kusabira-Orange2

MLLT10	Mixed-Lineage Leukemia Translocated to 10
MLLT3	Mixed-Lineage Leukemia Translocated To 3 protein
MMR	mismatch repair
mPgk1 3’	phosphoglycerate kinase 1 3’ promotor region
mPgk1 5’	phosphoglycerate kinase 1 5’ promotor region
MPP(s)	multipotent progenitor(s)
mRNA	messenger RNA
MYC	V-myc avian myelocytomatosis viral oncogene homolog
NGS	next generation sequencing
NHEJ	error-prone non-homologous end joining
NHEJ	non homologous end joining
NK cell	natural killer cell
NLS	nuclear localization signal
NPM1	Nucleophosmin 1
NRAS	Neuroblastoma RAS viral oncogene homolog
NTC	No template control
OCT4	octamer-binding transcription factor 4
P2A	self-cleaving peptide sequence P2A
PAGE	SDS-Poly Acrylamide Gel Electrophoresis
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PE	prime editor
PE4	prime editor system 4
pegRNA	prime editing guide RNA
PHD	plant homology domain
PI3K-AKT	phosphoinositide 3-Kinase-Protein Kinase B
PML::RARA	Promyelocytic Leukemia::Retinoic Acid Receptor Alpha
PMSF	phenylmethylsulfonyl fluoride
Pol II	RNA-Polymerase II
PRMD5	PR/SET Domain 5
PuroR	puromycin resistance cassette
qPCR	quantitative polymerase chain reaction
RAS	Rat Sarcoma Viral Oncogene Homolog.
RNA	ribonucleic acid.
ROCKi	ROCK inhibitor
RPMI	Roswell Park Memorial Institute
RT	reverse transcription/transcriptase
RUNX	runt-related transcription factor
RuvC I	Rec U Crossover
Sa	Staphylococcus aureus
SaCas9	Staphylococcus aureus Cas9
SCF	stem cell factor

scRNA-seq	single-cell RNA-sequencing
SDS	sodium dodecyl sulfate
SEC	super elongation complex
SET	Su(var)3–9, Ezh2, Trithorax methyltransferase
SETD2	SET Domain Containing 2
SNP	single nucleotide polymorphism
SOX2	SRY (sex determining Y)-box 2
Sp	Streptococcus pyogenes
SpCas9	Streptococcus pyogenes Cas9
Spi1	spleen Focus Forming Virus (SFFV) Proviral Integration Oncogene
SPRY2	Sprouty Homolog 2
SSBs	single strand breaks
ST-HSC(s)	short term HSC(s)
SV40 NLS	simian virus 40 Nuclear localization signal
TARGET	Therapeutically Applicable Research to Generate Effective Treatments
TAS	Taspase-1 cute site
TBE	Tris-Borate-EDTA
TBS	tris buffered saline
TBS-T	tris buffered saline tween 20
TET2	Ten-Eleven Translocation 2
TF(s)	transcription factor(s)
TKIs	tyrosine kinase inhibitors
TPO	thrombopoietin
tracrRNA	trans-activating crRNA
Tris	Tris(hydroxymethyl)aminomethane
TS	target site
WHO	World Health Organization
WT	wild type
YEATS	Yaf9, ENL, AF9, Taf14, and Sas5

References

Abbotts, R., & Wilson, D. M. (2017). Coordination of DNA single strand break repair. *Free Radical Biology and Medicine*, *107*, 228–244. https://doi.org/10.1016/j.freeradbiomed.2016.11.039

Alaggio, R., Amador, C., Anagnostopoulos, I., Attygalle, A. D., Araujo, I. B. de O., Berti, E., Bhagat, G., Borges, A. M., Boyer, D., Calaminici, M., Chadburn, A., Chan, J. K. C., Cheuk, W., Chng, W.-J., Choi, J. K., Chuang, S.-S., Coupland, S. E., Czader, M., Dave, S. S., … Xiao, W. (2022). The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia*, *36*(7), 1720–1748. https://doi.org/10.1038/s41375-022-01620-2

Alcalay, M., Meani, N., Gelmetti, V., Fantozzi, A., Fagioli, M., Orleth, A., Riganelli, D., Sebastiani, C., Cappelli, E., Casciari, C., Sciurpi, M. T., Mariano, A. R., Minardi, S. P., Luzi, L., Muller, H., Di Fiore, P. P., Frosina, G., & Pelicci, P. G. (2003). Acute myeloid leukemia fusion proteins deregulate genes involved in stem cell maintenance and DNA repair. *Journal of Clinical Investigation*, *112*(11), 1751–1761. https://doi.org/10.1172/JCI17595

Almosailleakh, M., & Schwaller, J. (2019). Murine Models of Acute Myeloid Leukaemia. *International Journal of Molecular Sciences*, *20*(2), 453. https://doi.org/10.3390/ijms20020453

Alonso, S., & Dow, L. E. (2021). Engineering chromosome rearrangements in cancer. *Disease Models & Mechanisms*, *14*(9), dmm049078. https://doi.org/10.1242/dmm.049078

Andrews, F. H., Strahl, B. D., & Kutateladze, T. G. (2016). Insights into newly discovered marks and readers of epigenetic information. *Nature Chemical Biology*, *12*(9), 662–668. https://doi.org/10.1038/nchembio.2149

Antunes, E. T. B., & Ottersbach, K. (2020). The MLL/SET family and haematopoiesis. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1863*(8), 194579. https://doi.org/10.1016/j.bbagrm.2020.194579

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., & Liu, D. R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, *576*(7785), 149–157. https://doi.org/10.1038/s41586-019-1711-4

Araki, K., Imaizumi, T., Okuyama, K., Oike, Y., & Yamamura, K. -i. (1997). Efficiency of Recombination by Cre Transient Expression in Embryonic Stem Cells: Comparison of Various Promoters. *Journal of Biochemistry*, *122*(5), 977–982. https://doi.org/10.1093/oxfordjournals.jbchem.a021860

Aung, M. M. K., Mills, M. L., Bittencourt-Silvestre, J., & Keeshan, K. (2021). Insights into the molecular profiles of adult and paediatric acute myeloid leukaemia. *Molecular Oncology*, *15*(9), 2253–2272. https://doi.org/10.1002/1878-0261.12899

Avellino, R., & Delwel, R. (2017). Expression and regulation of C/EBPα in normal myelopoiesis and in malignant transformation. *Blood*, *129*(15), 2083–2091. https://doi.org/10.1182/blood-2016-09-687822

Barabé, F., Gil, L., Celton, M., Bergeron, A., Lamontagne, V., Roques, É., Lagacé, K., Forest, A., Johnson, R., Pécheux, L., Simard, J., Pelloux, J., Bellemare-Pelletier, A., Gagnon, E., Hébert, J., Cellot, S., & Wilhelm, B. T. (2017). Modeling human MLL-AF9 translocated acute myeloid leukemia from single donors reveals RET as a potential therapeutic target. *Leukemia*, *31*(5), 1166–1176. https://doi.org/10.1038/leu.2016.302

Bertrand, J. Y., Chi, N. C., Santoso, B., Teng, S., Stainier, D. Y. R., & Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature*, *464*(7285), 108–111. https://doi.org/10.1038/nature08738

Bertuccio, S. N., Leardini, D., Messelodi, D., Anselmi, L., Manente, F., Ragni, F., Serravalle, S., Masetti, R., & Pession, A. (2022). Are Induced Pluripotent Stem Cells a Step towards Modeling Pediatric Leukemias? *Cells*, *11*(3), 476. https://doi.org/10.3390/cells11030476

Bolouri, H., Farrar, J. E., Triche, T., Ries, R. E., Lim, E. L., Alonzo, T. A., Ma, Y., Moore, R., Mungall, A. J., Marra, M. A., Zhang, J., Ma, X., Liu, Y., Liu, Y., Auvil, J. M. G., Davidsen, T. M., Gesuwan, P., Hermida, L. C., Salhia, B., … Meshinchi, S. (2018). The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nature Medicine*, *24*(1), 103–112. https://doi.org/10.1038/nm.4439

Brunet, E., & Jasin, M. (2018). Induction of Chromosomal Translocations with CRISPR-Cas9 and Other Nucleases: Understanding the Repair Mechanisms That Give Rise to Translocations. In Y. Zhang (Ed.), *Chromosome Translocation* (Vol. 1044, pp. 15–25). Springer Singapore. https://doi.org/10.1007/978-981-13-0593-1_2

Bunn, H. F. (2013). Erythropoietin. *Cold Spring Harbor Perspectives in Medicine*, *3*(3), a011619–a011619. https://doi.org/10.1101/cshperspect.a011619

Busuttil, R. A., Rubio, M., Dollé, M. E. T., Campisi, J., & Vijg, J. (2003). Oxygen accelerates the accumulation of mutations during the senescence and immortalization of murine cells in culture. *Aging Cell*, *2*(6), 287–294. https://doi.org/10.1046/j.1474-9728.2003.00066.x

Calvanese, V., Nguyen, A. T., Bolan, T. J., Vavilina, A., Su, T., Lee, L. K., Wang, Y., Lay, F. D., Magnusson, M., Crooks, G. M., Kurdistani, S. K., & Mikkola, H. K. A. (2019). MLLT3 governs human haematopoietic stem-cell self-renewal and engraftment. *Nature*, *576*(7786), 281–286. https://doi.org/10.1038/s41586-019-1790-2

Calvo, C., Fenneteau, O., Leverger, G., Petit, A., Baruchel, A., & Méchinaud, F. (2021). Infant Acute Myeloid Leukemia: A Unique Clinical and Biological Entity. *Cancers*, *13*(4), 777. https://doi.org/10.3390/cancers13040777

Chao, M. P., Gentles, A. J., Chatterjee, S., Lan, F., Reinisch, A., Corces, M. R., Xavy, S., Shen, J., Haag, D., Chanda, S., Sinha, R., Morganti, R. M., Nishimura, T., Ameen, M., Wu, H., Wernig, M., Wu, J. C., & Majeti, R. (2017). Human AML-iPSCs Reacquire Leukemic Properties after Differentiation and Model Clonal Variation of Disease. *Cell Stem Cell*, *20*(3), 329-344.e7. https://doi.org/10.1016/j.stem.2016.11.018

Chen, D., Tang, T.-X., Deng, H., Yang, X.-P., & Tang, Z.-H. (2021). Interleukin-7 Biology and Its Effects on Immune Cells: Mediator of Generation, Differentiation, Survival, and

Homeostasis. *Frontiers in Immunology*, *12*, 747324. https://doi.org/10.3389/fimmu.2021.747324

Chen, P. J., Hussmann, J. A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.-F., Chen, C., Nelson, J. W., Newby, G. A., Sahin, M., Osborn, M. J., Weissman, J. S., Adamson, B., & Liu, D. R. (2021). Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell*, *184*(22), 5635-5652.e29. https://doi.org/10.1016/j.cell.2021.09.018

Chen, T., Wang, F., Wu, M., & Wang, Z. Z. (2015). Development of Hematopoietic Stem and Progenitor Cells From Human Pluripotent Stem Cells: HSC DEVELOPMENT IN hPSCs. *Journal of Cellular Biochemistry*, *116*(7), 1179–1189. https://doi.org/10.1002/jcb.25097

Chen, X., Burkhardt, D. B., Hartman, A. A., Hu, X., Eastman, A. E., Sun, C., Wang, X., Zhong, M., Krishnaswamy, S., & Guo, S. (2019). MLL-AF9 initiates transformation from fast-proliferating myeloid progenitors. *Nature Communications*, *10*(1), 5767. https://doi.org/10.1038/s41467-019-13666-5

Chen, X., Janssen, J. M., Liu, J., Maggio, I., 't Jong, A. E. J., Mikkers, H. M. M., & Gonçalves, M. A. F. V. (2017a). In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nature Communications*, *8*(1), 657. https://doi.org/10.1038/s41467-017-00687-1

Chen, X., Janssen, J. M., Liu, J., Maggio, I., 't Jong, A. E. J., Mikkers, H. M. M., & Gonçalves, M. A. F. V. (2017b). In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nature Communications*, *8*(1), 657. https://doi.org/10.1038/s41467-017-00687-1

Cheng, H., Zheng, Z., & Cheng, T. (2020). New paradigms on hematopoietic stem cell differentiation. *Protein & Cell*, *11*(1), 34–44. https://doi.org/10.1007/s13238-019-0633-0

Collins, C. T., & Hess, J. L. (2016a). Role of HOXA9 in leukemia: Dysregulation, cofactors and essential targets. *Oncogene*, *35*(9), 1090–1098. https://doi.org/10.1038/onc.2015.174

Collins, C. T., & Hess, J. L. (2016b). Role of HOXA9 in leukemia: Dysregulation, cofactors and essential targets. *Oncogene*, *35*(9), 1090–1098. https://doi.org/10.1038/onc.2015.174

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, *339*(6121), 819–823. https://doi.org/10.1126/science.1231143

Corral, J., Lavenir, I., Impey, H., Warren, A. J., Forster, A., Larson, T. A., Bell, S., McKenzie, A. N. J., King, G., & Rabbitts, T. H. (1996). An Mll–AF9 Fusion Gene Made by Homologous Recombination Causes Acute Leukemia in Chimeric Mice: A Method to Create Fusion Oncogenes. *Cell*, *85*(6), 853–861. https://doi.org/10.1016/S0092-8674(00)81269-6

Creutzig, U., Van Den Heuvel-Eibrink, M. M., Gibson, B., Dworzak, M. N., Adachi, S., De Bont, E., Harbott, J., Hasle, H., Johnston, D., Kinoshita, A., Lehrnbecher, T., Leverger, G., Mejstrikova, E., Meshinchi, S., Pession, A., Raimondi, S. C., Sung, L., Stary, J., Zwaan, C. M., … Reinhardt, D. (2012). Diagnosis and management of acute myeloid leukemia in children and adolescents: Recommendations from an international expert panel. *Blood*, *120*(16), 3187–3205. https://doi.org/10.1182/blood-2012-03-362608

Dang, S. M., Kyba, M., Perlingeiro, R., Daley, G. Q., & Zandstra, P. W. (2002). Efficiency of embryoid body formation and hematopoietic development from embryonic stem cells in different culture systems. *Biotechnology and Bioengineering*, *78*(4), 442–453. https://doi.org/10.1002/bit.10220

De Bruijn, M., & Dzierzak, E. (2017). Runx transcription factors in the development and function of the definitive hematopoietic system. *Blood*, *129*(15), 2061–2069. https://doi.org/10.1182/blood-2016-12-689109

de Rooij, J., Zwaan, C., & van den Heuvel-Eibrink, M. (2015). Pediatric AML: From Biology to Clinical Management. *Journal of Clinical Medicine*, *4*(1), 127–149. https://doi.org/10.3390/jcm4010127

Doman, J. L., Sousa, A. A., Randolph, P. B., Chen, P. J., & Liu, D. R. (2022). Designing and executing prime editing experiments in mammalian cells. *Nature Protocols*. https://doi.org/10.1038/s41596-022-00724-4

Donada, A., Basso-Valentina, F., Arkoun, B., Monte-Mor, B., Plo, I., & Raslova, H. (2020). Induced pluripotent stem cells and hematological malignancies: A powerful tool for disease modeling and drug development. *Stem Cell Research*, *49*, 102060. https://doi.org/10.1016/j.scr.2020.102060

Doré, L. C., & Crispino, J. D. (2011). Transcription factor networks in erythroid cell and megakaryocyte development. *Blood*, *118*(2), 231–239. https://doi.org/10.1182/blood-2011-04-285981

Drexler, H. G., Quentmeier, H., & MacLeod, R. A. F. (2004). Malignant hematopoietic cell lines: In vitro models for the study of MLL gene alterations. *Leukemia*, *18*(2), 227–232. https://doi.org/10.1038/sj.leu.2403236

Dzierzak, E., & Speck, N. A. (2008). Of lineage and legacy: The development of mammalian hematopoietic stem cells. *Nature Immunology*, *9*(2), 129–136. https://doi.org/10.1038/ni1560

Elbaz, O., & Shaltout, A. (2000). Implication of Granulocyte-Macrophage Colony Stimulating Factor (GM-CSF) and Interleukin-3 (IL-3) in Children with Acute Myeloid Leukaemia (AML). *Hematology*, *5*(5), 383–388. https://doi.org/10.1080/10245332.2000.11746533

Fortschegger, K., Husa, A.-M., Schinnerl, D., Nebral, K., & Strehl, S. (2021). Expression of RUNX1-JAK2 in Human Induced Pluripotent Stem Cell-Derived Hematopoietic Cells Activates the JAK-STAT and MYC Pathways. *International Journal of Molecular Sciences*, *22*(14), 7576. https://doi.org/10.3390/ijms22147576

Gao, X., Xu, C., Asada, N., & Frenette, P. S. (2018). The hematopoietic stem cell niche: From embryo to adult. *Development*, *145*(2), dev139691. https://doi.org/10.1242/dev.139691

Gaussmann, A., Wenger, T., Eberle, I., Bursen, A., Bracharz, S., Herr, I., Dingermann, T., & Marschalek, R. (2007). Combined effects of the two reciprocal t(4;11) fusion proteins MLL·AF4 and AF4·MLL confer resistance to apoptosis, cell cycling capacity and growth transformation. *Oncogene*, *26*(23), 3352–3363. https://doi.org/10.1038/sj.onc.1210125

Geddis, A. E. (2010). Megakaryopoiesis. *Seminars in Hematology*, *47*(3), 212–219. https://doi.org/10.1053/j.seminhematol.2010.03.001

Gore, A. V., Athans, B., Iben, J. R., Johnson, K., Russanova, V., Castranova, D., Pham, V. N., Butler, M. G., Williams-Simons, L., Nichols, J. T., Bresciani, E., Feldman, B., Kimmel, C. B., Liu, P. P., & Weinstein, B. M. (2016). Epigenetic regulation of hematopoiesis by DNA methylation. *eLife*, *5*, e11813. https://doi.org/10.7554/eLife.11813

Gore, A. V., & Weinstein, B. M. (2016). DNA methylation in hematopoietic development and disease. *Experimental Hematology*, *44*(9), 783–790. https://doi.org/10.1016/j.exphem.2016.04.013

Goyama, S., & Kitamura, T. (2017). Epigenetics in normal and malignant hematopoiesis: An overview and update 2017. *Cancer Science*, *108*(4), 553–562. https://doi.org/10.1111/cas.13168

Gritz, E., & Hirschi, K. K. (2016). Specification and function of hemogenic endothelium during embryogenesis. *Cellular and Molecular Life Sciences*, *73*(8), 1547–1567. https://doi.org/10.1007/s00018-016-2134-0

Gröschel, S., Sanders, M. A., Hoogenboezem, R., de Wit, E., Bouwman, B. A. M., Erpelinck, C., van der Velden, V. H. J., Havermans, M., Avellino, R., van Lom, K., Rombouts, E. J., van Duin, M., Döhner, K., Beverloo, H. B., Bradner, J. E., Döhner, H., Löwenberg, B., Valk, P. J. M., Bindels, E. M. J., … Delwel, R. (2014). A Single Oncogenic Enhancer Rearrangement Causes Concomitant EVI1 and GATA2 Deregulation in Leukemia. *Cell*, *157*(2), 369–381. https://doi.org/10.1016/j.cell.2014.02.019

Heisterkamp, N., Jenster, G., Kioussis, D., Pattengale, P. K., & Groffen, J. (1991). Humanbcr-abl gene has a lethal effect on embryogenesis. *Transgenic Research*, *1*(1), 45–53. https://doi.org/10.1007/BF02512996

Heuts, B. M. H., Arza-Apalategi, S., Alkema, S. G., Tijchon, E., Jussen, L., Bergevoet, S. M., Van Der Reijden, B. A., & Martens, J. H. A. (2023). Inducible MLL-AF9 Expression Drives an AML Program during Human Pluripotent Stem Cell-Derived Hematopoietic Differentiation. *Cells*, *12*(8), 1195. https://doi.org/10.3390/cells12081195

Hirakawa, H., Lee, Y., & Ding, L. (2022). The Fetal Hematopoietic Niche: Components and Mechanisms for Hematopoietic Stem Cell Emergence and Expansion. *Current Stem Cell Reports*, *8*(1), 14–23. https://doi.org/10.1007/s40778-021-00202-9

Husmann, D., & Gozani, O. (2019). Histone lysine methyltransferases in biology and disease. *Nature Structural & Molecular Biology*, *26*(10), 880–889. https://doi.org/10.1038/s41594-019-0298-7

Hyrenius-Wittsten, A., Pilheden, M., Sturesson, H., Hansson, J., Walsh, M. P., Song, G., Kazi, J. U., Liu, J., Ramakrishan, R., Garcia-Ruiz, C., Nance, S., Gupta, P., Zhang, J., Rönnstrand, L., Hultquist, A., Downing, J. R., Lindkvist-Petersson, K., Paulsson, K., Järås, M., … Hagström-Andersson, A. K. (2018). De novo activating mutations drive clonal evolution and enhance clonal fitness in KMT2A-rearranged leukemia. *Nature Communications*, *9*(1), 1770. https://doi.org/10.1038/s41467-018-04180-1

Ihry, R. J., Worringer, K. A., Salick, M. R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J., Sondey, M., Ye, C., Randhawa, R., Kulkarni, T., Yang, Z., McAllister, G., Russ, C., Reece-Hoyes, J., Forrester, W., Hoffman, G. R., Dolmetsch, R., & Kaykas, A. (2018). P53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nature Medicine*, *24*(7), 939–946. https://doi.org/10.1038/s41591-018-0050-6

Issa, G. C., Aldoss, I., DiPersio, J., Cuglievan, B., Stone, R., Arellano, M., Thirman, M. J., Patel, M. R., Dickens, D. S., Shenoy, S., Shukla, N., Kantarjian, H., Armstrong, S. A., Perner, F., Perry, J. A., Rosen, G., Bagley, R. G., Meyers, M. L., Ordentlich, P., … Stein, E. M. (2023). The menin inhibitor revumenib in KMT2A-rearranged or NPM1-mutant leukaemia. *Nature*, *615*(7954), 920–924. https://doi.org/10.1038/s41586-023-05812-3

Jacobs, K., Zambelli, F., Mertzanidou, A., Smolders, I., Geens, M., Nguyen, H. T., Barbé, L., Sermon, K., & Spits, C. (2016). Higher-Density Culture in Human Embryonic Stem Cells Results in DNA Damage and Genome Instability. *Stem Cell Reports*, *6*(3), 330–341. https://doi.org/10.1016/j.stemcr.2016.01.015

Jafari, M., Ghadami, E., Dadkhah, T., & Akhavan-Niaki, H. (2019). PI3k/AKT signaling pathway: Erythropoiesis and beyond. *Journal of Cellular Physiology*, *234*(3), 2373–2385. https://doi.org/10.1002/jcp.27262

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Kabra, A., & Bushweller, J. (2022). The Intrinsically Disordered Proteins MLLT3 (AF9) and MLLT1 (ENL) – Multimodal Transcriptional Switches With Roles in Normal Hematopoiesis, MLL Fusion Leukemia, and Kidney Cancer. *Journal of Molecular Biology*, *434*(1), 167117. https://doi.org/10.1016/j.jmb.2021.167117

Kardel, M. D., & Eaves, C. J. (2012). Modeling human hematopoietic cell development from pluripotent stem cells. *Experimental Hematology*, *40*(8), 601–611. https://doi.org/10.1016/j.exphem.2012.04.001

Karnoub, A. E., & Weinberg, R. A. (2008). Ras oncogenes: Split personalities. *Nature Reviews Molecular Cell Biology*, *9*(7), 517–531. https://doi.org/10.1038/nrm2438

Kassim, A. A., & Savani, B. N. (2017). Hematopoietic stem cell transplantation for acute myeloid leukemia: A review. *Hematology/Oncology and Stem Cell Therapy*, *10*(4), 245–251. https://doi.org/10.1016/j.hemonc.2017.05.021

Katsumura, K. R., Bresnick, E. H., & the GATA Factor Mechanisms Group. (2017). The GATA factor revolution in hematology. *Blood*, *129*(15), 2092–2102. https://doi.org/10.1182/blood-2016-09-687871

Katti, A., Diaz, B. J., Caragine, C. M., Sanjana, N. E., & Dow, L. E. (2022). CRISPR in cancer biology and therapy. *Nature Reviews Cancer*, *22*(5), 259–279. https://doi.org/10.1038/s41568-022-00441-w

Klaver-Flores, S., Zittersteijn, H. A., Canté-Barrett, K., Lankester, A., Hoeben, R. C., Gonçalves, M. A. F. V., Pike-Overzet, K., & Staal, F. J. T. (2021). Genomic Engineering in

Human Hematopoietic Stem Cells: Hype or Hope? *Frontiers in Genome Editing*, *2*, 615619. https://doi.org/10.3389/fgeed.2020.615619

Koedijk, J. B., Van Der Werf, I., Calkoen, F. G., Nierkens, S., Kaspers, G. J. L., Zwaan, C. M., & Heidenreich, O. (2021). Paving the Way for Immunotherapy in Pediatric Acute Myeloid Leukemia: Current Knowledge and the Way Forward. *Cancers*, *13*(17), 4364. https://doi.org/10.3390/cancers13174364

Kotini, A. G., Carcamo, S., Cruz-Rodriguez, N., Olszewska, M., Wang, T., Demircioglu, D., Chang, C.-J., Bernard, E., Chao, M. P., Majeti, R., Luo, H., Kharas, M. G., Hasson, D., & Papapetrou, E. P. (2023). *Patient-Derived iPSCs Faithfully Represent the Genetic Diversity and Cellular Architecture of Human Acute Myeloid Leukemia*.

Krivtsov, A. V., Hoshii, T., & Armstrong, S. A. (2017). Mixed-Lineage Leukemia Fusions and Chromatin in Leukemia. *Cold Spring Harbor Perspectives in Medicine*, *7*(11), a026658. https://doi.org/10.1101/cshperspect.a026658

Kuijk, E., Jager, M., Van Der Roest, B., Locati, M. D., Van Hoeck, A., Korzelius, J., Janssen, R., Besselink, N., Boymans, S., Van Boxtel, R., & Cuppen, E. (2020). The mutational impact of culturing human pluripotent and adult stem cells. *Nature Communications*, *11*(1), 2493. https://doi.org/10.1038/s41467-020-16323-4

Kumar, A. R., Li, Q., Hudson, W. A., Chen, W., Sam, T., Yao, Q., Lund, E. A., Wu, B., Kowal, B. J., & Kersey, J. H. (2009). A role for MEIS1 in MLL-fusion gene leukemia. *Blood*, *113*(8), 1756–1758. https://doi.org/10.1182/blood-2008-06-163287

Ladines-Castro, W., Barragán-Ibañez, G., Luna-Pérez, M. A., Santoyo-Sánchez, A., Collazo-Jaloma, J., Mendoza-García, E., & Ramos-Peñafiel, C. O. (2016). Morphology of leukaemias. *Revista Médica Del Hospital General de México*, *79*(2), 107–113. https://doi.org/10.1016/j.hgmx.2015.06.007

Laurenti, E., & Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature*, *553*(7689), 418–426. https://doi.org/10.1038/nature25022

Li, H., Luo, Q., Shan, W., Cai, S., Tie, R., Xu, Y., Lin, Y., Qian, P., & Huang, H. (2021). Biomechanical cues as master regulators of hematopoietic stem cell fate. *Cellular and Molecular Life Sciences*, *78*(16), 5881–5902. https://doi.org/10.1007/s00018-021-03882-y

Liggett, L. A., & Sankaran, V. G. (2020). Unraveling Hematopoiesis through the Lens of Genomics. *Cell*, *182*(6), 1384–1400. https://doi.org/10.1016/j.cell.2020.08.030

Lim, W. F., Inoue-Yokoo, T., Tan, K. S., Lai, M. I., & Sugiyama, D. (2013). Hematopoietic cell differentiation from embryonic and induced pluripotent stem cells. *Stem Cell Research & Therapy*, *4*(3), 71. https://doi.org/10.1186/scrt222

Liu, X., Ye, Q., Zhao, X.-P., Zhang, P.-B., Li, S., Li, R.-Q., & Zhao, X.-L. (2019). RAS mutations in acute myeloid leukaemia patients: A review and meta-analysis. *Clinica Chimica Acta*, *489*, 254–260. https://doi.org/10.1016/j.cca.2018.08.040

Liu, Y., Cheng, H., Gao, S., Lu, X., He, F., Hu, L., Hou, D., Zou, Z., Li, Y., Zhang, H., Xu, J., Kang, L., Wang, Q., Yuan, W., Gao, S., & Cheng, T. (2014). Reprogramming of MLL-AF9

leukemia cells into pluripotent stem cells. *Leukemia*, *28*(5), 1071–1080. https://doi.org/10.1038/leu.2013.304

Luciano, M., Krenn, P. W., & Horejs-Hoeck, J. (2022). The cytokine network in acute myeloid leukemia. *Frontiers in Immunology*, *13*, 1000996. https://doi.org/10.3389/fimmu.2022.1000996

Luo, Z., Lin, C., & Shilatifard, A. (2012). The super elongation complex (SEC) family in transcriptional control. *Nature Reviews Molecular Cell Biology*, *13*(9), 543–547. https://doi.org/10.1038/nrm3417

Marchand, T., & Pinho, S. (2021). Leukemic Stem Cells: From Leukemic Niche Biology to Treatment Opportunities. *Frontiers in Immunology*, *12*, 775128. https://doi.org/10.3389/fimmu.2021.775128

Marcotte, E. L., Spector, L. G., Mendes-de-Almeida, D. P., & Nelson, H. H. (2021). The Prenatal Origin of Childhood Leukemia: Potential Applications for Epidemiology and Newborn Screening. *Frontiers in Pediatrics*, *9*, 639479. https://doi.org/10.3389/fped.2021.639479

Mercher, T., & Schwaller, J. (2019). Pediatric Acute Myeloid Leukemia (AML): From Genes to Models Toward Targeted Therapeutic Intervention. *Frontiers in Pediatrics*, *7*, 401. https://doi.org/10.3389/fped.2019.00401

Metcalf, D. (2008). Hematopoietic cytokines. *Blood*, *111*(2), 485–491. https://doi.org/10.1182/blood-2007-03-079681

Meyer, C., Burmeister, T., Gröger, D., Tsaur, G., Fechina, L., Renneville, A., Sutton, R., Venn, N. C., Emerenciano, M., Pombo-de-Oliveira, M. S., Barbieri Blunck, C., Almeida Lopes, B., Zuna, J., Trka, J., Ballerini, P., Lapillonne, H., De Braekeleer, M., Cazzaniga, G., Corral Abascal, L., … Marschalek, R. (2018). The MLL recombinome of acute leukemias in 2017. *Leukemia*, *32*(2), 273–284. https://doi.org/10.1038/leu.2017.213

Meyer, C., Larghero, P., Almeida Lopes, B., Burmeister, T., Gröger, D., Sutton, R., Venn, N. C., Cazzaniga, G., Corral Abascal, L., Tsaur, G., Fechina, L., Emerenciano, M., Pombo-de-Oliveira, M. S., Lund-Aho, T., Lundán, T., Montonen, M., Juvonen, V., Zuna, J., Trka, J., … Marschalek, R. (2023). The KMT2A recombinome of acute leukemias in 2023. *Leukemia*, *37*(5), 988–1005. https://doi.org/10.1038/s41375-023-01877-1

Milan, T., Canaj, H., Villeneuve, C., Ghosh, A., Barabé, F., Cellot, S., & Wilhelm, B. T. (2019). Pediatric leukemia: Moving toward more accurate models. *Experimental Hematology*, *74*, 1–12. https://doi.org/10.1016/j.exphem.2019.05.003

Milholland, B., Dong, X., Zhang, L., Hao, X., Suh, Y., & Vijg, J. (2017). Differences between germline and somatic mutation rates in humans and mice. *Nature Communications*, *8*(1), 15183. https://doi.org/10.1038/ncomms15183

Milne, T. A. (2017). Mouse models of MLL leukemia: Recapitulating the human disease. *Blood*, *129*(16), 2217–2223. https://doi.org/10.1182/blood-2016-10-691428

Molina, B., Chavez, J., & Grainger, S. (2021). Zebrafish models of acute leukemias: Current models and future directions. *WIREs Developmental Biology*, *10*(6). https://doi.org/10.1002/wdev.400

Muhammad, K. A., Nur, A. A., Nurul, H. S., Narazah, M. Y., & Siti, R. A. R. (2018). Dual-specificity phosphatase 6 (DUSP6): A review of its molecular characteristics and clinical relevance in cancer. *Cancer Biology & Medicine*, *15*(1), 14. https://doi.org/10.20892/j.issn.2095-3941.2017.0107

Nabet, B., Ferguson, F. M., Seong, B. K. A., Kuljanin, M., Leggett, A. L., Mohardt, M. L., Robichaud, A., Conway, A. S., Buckley, D. L., Mancias, J. D., Bradner, J. E., Stegmaier, K., & Gray, N. S. (2020). Rapid and direct control of target protein levels with VHL-recruiting dTAG molecules. *Nature Communications*, *11*(1), 4687. https://doi.org/10.1038/s41467-020-18377-w

Nabet, B., Roberts, J. M., Buckley, D. L., Paulk, J., Dastjerdi, S., Yang, A., Leggett, A. L., Erb, M. A., Lawlor, M. A., Souza, A., Scott, T. G., Vittori, S., Perry, J. A., Qi, J., Winter, G. E., Wong, K.-K., Gray, N. S., & Bradner, J. E. (2018). The dTAG system for immediate and target-specific protein degradation. *Nature Chemical Biology*, *14*(5), 431–441. https://doi.org/10.1038/s41589-018-0021-8

Nakano, T., Kodama, H., & Honjo, T. (1996). In Vitro Development of Primitive and Definitive Erythrocytes from Different Precursors. *Science*, *272*(5262), 722–724. https://doi.org/10.1126/science.272.5262.722

Nowak, R., Oelschlägel, U., Gurth, H., Range, U., Albrecht, S., Krebs, U., Hietschold, V., & Ehninger, G. (1999). RELATIONS BETWEEN IL-3-INDUCED PROLIFERATION AND IN VITRO CYTOKINE SECRETION OF BONE MARROW CELLS FROM AML PATIENTS. *Cytokine*, *11*(6), 435–442. https://doi.org/10.1006/cyto.1998.0445

Omole, A. E., & Fakoya, A. O. J. (2018). Ten years of progress and promise of induced pluripotent stem cells: Historical origins, characteristics, mechanisms, limitations, and potential applications. *PeerJ*, *6*, e4370. https://doi.org/10.7717/peerj.4370

Orkin, S. H., & Zon, L. I. (2008). Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell*, *132*(4), 631–644. https://doi.org/10.1016/j.cell.2008.01.025

Papapetrou, E. P. (2019). Modeling Leukemia with Human Induced Pluripotent Stem Cells. *Cold Spring Harbor Perspectives in Medicine*, *9*(12), a034868. https://doi.org/10.1101/cshperspect.a034868

Pei, W., Feyerabend, T. B., Rössler, J., Wang, X., Postrach, D., Busch, K., Rode, I., Klapproth, K., Dietlein, N., Quedenau, C., Chen, W., Sauer, S., Wolf, S., Höfer, T., & Rodewald, H.-R. (2017). Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature*, *548*(7668), 456–460. https://doi.org/10.1038/nature23653

Perlman, R. L. (2016). Mouse Models of Human Disease: An Evolutionary Perspective. *Evolution, Medicine, and Public Health*, eow014. https://doi.org/10.1093/emph/eow014

Pettersson, M., & Crews, C. M. (2019). PROteolysis TArgeting Chimeras (PROTACs)—Past, present and future. *Drug Discovery Today: Technologies*, *31*, 15–27. https://doi.org/10.1016/j.ddtec.2019.01.002

Pietilä, I., & Vainio, S. (2005). The embryonic aorta-gonad-mesonephros region as a generator of haematopoietic stem cells. *APMIS*, *113*(11–12), 804–812. https://doi.org/10.1111/j.1600-0463.2005.apm_368.x

Pietras, E. M., Reynaud, D., Kang, Y.-A., Carlin, D., Calero-Nieto, F. J., Leavitt, A. D., Stuart, J. M., Göttgens, B., & Passegué, E. (2015). Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell*, *17*(1), 35–46. https://doi.org/10.1016/j.stem.2015.05.003

Pommert, L., & Tarlock, K. (2022). The evolution of targeted therapy in pediatric AML: Gemtuzumab ozogamicin, FLT3/IDH/BCL2 inhibitors, and other therapies. *Hematology*, *2022*(1), 603–610. https://doi.org/10.1182/hematology.2022000358

Quessada, J., Cuccuini, W., Saultier, P., Loosveld, M., Harrison, C. J., & Lafage-Pochitaloff, M. (2021a). Cytogenetics of Pediatric Acute Myeloid Leukemia: A Review of the Current Knowledge. *Genes*, *12*(6), 924. https://doi.org/10.3390/genes12060924

Quessada, J., Cuccuini, W., Saultier, P., Loosveld, M., Harrison, C. J., & Lafage-Pochitaloff, M. (2021b). Cytogenetics of Pediatric Acute Myeloid Leukemia: A Review of the Current Knowledge. *Genes*, *12*(6), 924. https://doi.org/10.3390/genes12060924

Ranzoni, A. M., Tangherloni, A., Berest, I., Riva, S. G., Myers, B., Strzelecka, P. M., Xu, J., Panada, E., Mohorianu, I., Zaugg, J. B., & Cvejic, A. (2021). Integrative Single-Cell RNA-Seq and ATAC-Seq Analysis of Human Developmental Hematopoiesis. *Cell Stem Cell*, *28*(3), 472-487.e7. https://doi.org/10.1016/j.stem.2020.11.015

Rees, H. A., & Liu, D. R. (2018). Base editing: Precision chemistry on the genome and transcriptome of living cells. *Nature Reviews Genetics*, *19*(12), 770–788. https://doi.org/10.1038/s41576-018-0059-1

Reinhardt, D., Antoniou, E., & Waack, K. (2022). Pediatric Acute Myeloid Leukemia—Past, Present, and Future. *Journal of Clinical Medicine*, *11*(3), 504. https://doi.org/10.3390/jcm11030504

Reuven, N., & Shaul, Y. (2022). Selecting for CRISPR-Edited Knock-In Cells. *International Journal of Molecular Sciences*, *23*(19), 11919. https://doi.org/10.3390/ijms231911919

Richter, M., Piwocka, O., Musielak, M., Piotrowski, I., Suchorska, W. M., & Trzeciak, T. (2021). From Donor to the Lab: A Fascinating Journey of Primary Cell Lines. *Frontiers in Cell and Developmental Biology*, *9*, 711381. https://doi.org/10.3389/fcell.2021.711381

Rodrigues, C. P., Shvedunova, M., & Akhtar, A. (2021). Epigenetic Regulators as the Gatekeepers of Hematopoiesis. *Trends in Genetics*, *37*(2), 125–142. https://doi.org/10.1016/j.tig.2020.09.015

Rowe, R. G., Lummertz Da Rocha, E., Sousa, P., Missios, P., Morse, M., Marion, W., Yermalovich, A., Barragan, J., Mathieu, R., Jha, D. K., Fleming, M. D., North, T. E., & Daley,

G. Q. (2019). The developmental stage of the hematopoietic niche regulates lineage in MLL-rearranged leukemia. *Journal of Experimental Medicine*, *216*(3), 527–538. https://doi.org/10.1084/jem.20181765

Rubnitz, J. E., & Inaba, H. (2012). Childhood acute myeloid leukaemia. *British Journal of Haematology*, *159*(3), 259–276. https://doi.org/10.1111/bjh.12040

Ruiz, J. P., Chen, G., Haro Mora, J. J., Keyvanfar, K., Liu, C., Zou, J., Beers, J., Bloomer, H., Qanash, H., Uchida, N., Tisdale, J. F., Boehm, M., & Larochelle, A. (2019). Robust generation of erythroid and multilineage hematopoietic progenitors from human iPSCs using a scalable monolayer culture system. *Stem Cell Research*, *41*, 101600. https://doi.org/10.1016/j.scr.2019.101600

Sachs, L. (1996). The control of hematopoiesis and leukemia: From basic biology to the clinic. *Proceedings of the National Academy of Sciences*, *93*(10), 4742–4749. https://doi.org/10.1073/pnas.93.10.4742

Salci, K. R., Lee, J.-H., Laronde, S., Dingwall, S., Kushwah, R., Fiebig-Comyn, A., Leber, B., Foley, R., Dal Cin, A., & Bhatia, M. (2015). Cellular Reprogramming Allows Generation of Autologous Hematopoietic Progenitors From AML Patients That Are Devoid of Patient-Specific Genomic Aberrations. *Stem Cells*, *33*(6), 1839–1849. https://doi.org/10.1002/stem.1994

Schmetzer, H. M., Gerhartz, H. H., & Wilmanns, W. (1999). GM-CSF stimulates proliferation of clonal leukemic bone marrow cells in acute myeloid leukemia (AML) in vitro. *Annals of Hematology*, *78*(10), 449–455. https://doi.org/10.1007/s002770050597

Seita, J., & Weissman, I. L. (2010). Hematopoietic stem cell: Self-renewal versus differentiation: Hematopoietic stem cell. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *2*(6), 640–653. https://doi.org/10.1002/wsbm.86

Shiba, N. (2023). Comprehensive molecular understanding of pediatric acute myeloid leukemia. *International Journal of Hematology*, *117*(2), 173–181. https://doi.org/10.1007/s12185-023-03533-x

Shojaee, S., Caeser, R., Buchner, M., Park, E., Swaminathan, S., Hurtz, C., Geng, H., Chan, L. N., Klemm, L., Hofmann, W.-K., Qiu, Y. H., Zhang, N., Coombes, K. R., Paietta, E., Molkentin, J., Koeffler, H. P., Willman, C. L., Hunger, S. P., Melnick, A., … Müschen, M. (2015). Erk Negative Feedback Control Enables Pre-B Cell Transformation and Represents a Therapeutic Target in Acute Lymphoblastic Leukemia. *Cancer Cell*, *28*(1), 114–128. https://doi.org/10.1016/j.ccell.2015.05.008

Sive, J. I., & Göttgens, B. (2014). Transcriptional network control of normal and leukaemic haematopoiesis. *Experimental Cell Research*, *329*(2), 255–264. https://doi.org/10.1016/j.yexcr.2014.06.021

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., & Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science*, *351*(6268), 84–88. https://doi.org/10.1126/science.aad5227

Slukvin, I. I. (2013). Hematopoietic specification from human pluripotent stem cells: Current advances and challenges toward de novo generation of hematopoietic stem cells. *Blood*, *122*(25), 4035–4046. https://doi.org/10.1182/blood-2013-07-474825

Somasundaram, R., Prasad, M. A. J., Ungerbäck, J., & Sigvardsson, M. (2015). Transcription factor networks in B-cell differentiation link development to acute lymphoid leukemia. *Blood*, *126*(2), 144–152. https://doi.org/10.1182/blood-2014-12-575688

Staerk, J., & Constantinescu, S. N. (2012). The JAK-STAT pathway and hematopoietic stem cells from the JAK2 V617F perspective. *JAK-STAT*, *1*(3), 184–190. https://doi.org/10.4161/jkst.22071

Stavropoulou, V., Kaspar, S., Brault, L., Sanders, M. A., Juge, S., Morettini, S., Tzankov, A., Iacovino, M., Lau, I.-J., Milne, T. A., Royo, H., Kyba, M., Valk, P. J. M., Peters, A. H. F. M., & Schwaller, J. (2016). MLL-AF9 Expression in Hematopoietic Stem Cells Drives a Highly Invasive AML Expressing EMT-Related Genes Linked to Poor Outcome. *Cancer Cell*, *30*(1), 43–58. https://doi.org/10.1016/j.ccell.2016.05.011

Symeonidou, V., & Ottersbach, K. (2021). HOXA9/IRX1 expression pattern defines two subgroups of infant MLL-AF4-driven acute lymphoblastic leukemia. *Experimental Hematology*, *93*, 38-43.e5. https://doi.org/10.1016/j.exphem.2020.10.002

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, *131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019

Tan, J., Zhao, L., Wang, G., Li, T., Li, D., Xu, Q., Chen, X., Shang, Z., Wang, J., & Zhou, J. (2018). Human MLL-AF9 Overexpression Induces Aberrant Hematopoietic Expansion in Zebrafish. *BioMed Research International*, *2018*, 1–9. https://doi.org/10.1155/2018/6705842

Terns, M. P., & Terns, R. M. (2011). CRISPR-based adaptive immune systems. *Current Opinion in Microbiology*, *14*(3), 321–327. https://doi.org/10.1016/j.mib.2011.03.005

Testa, U. (n.d.). Germana Castelli Elvira Pelosi Ugo Testa. *OncoTargets and Therapy*.

Torres, R., Rodriguez-Perales, S., & Ramirez, J. (2015). The Use of Innovative Tools to Reproduce Human Cancer Translocations: Lessons from the CRISPR/Cas System. *Current Biotechnology*, *3*(4), 273–278. https://doi.org/10.2174/2211550103666141023221106

Tsavachidou, D., Coleman, M. L., Athanasiadis, G., Li, S., Licht, J. D., Olson, M. F., & Weber, B. L. (2004). SPRY2 Is an Inhibitor of the Ras/Extracellular Signal-Regulated Kinase Pathway in Melanocytes and Melanoma Cells with Wild-Type ***BRAF*** but Not with the V599E Mutant. *Cancer Research*, *64*(16), 5556–5559. https://doi.org/10.1158/0008-5472.CAN-04-1669

Vanoli, F., Tomishima, M., Feng, W., Lamribet, K., Babin, L., Brunet, E., & Jasin, M. (2017). CRISPR-Cas9–guided oncogenic chromosomal translocations with conditional fusion protein expression in human mesenchymal cells. *Proceedings of the National Academy of Sciences*, *114*(14), 3696–3701. https://doi.org/10.1073/pnas.1700622114

Voon, D. C.-C., Hor, Y. T., & Ito, Y. (2015). The RUNX complex: Reaching beyond haematopoiesis into immunity. *Immunology*, *146*(4), 523–536. https://doi.org/10.1111/imm.12535

Wang, J., Kong, G., Liu, Y., Du, J., Chang, Y.-I., Tey, S. R., Zhang, X., Ranheim, E. A., Saba-El-Leil, M. K., Meloche, S., Damnernsawad, A., Zhang, J., & Zhang, J. (2013). NrasG12D/+ promotes leukemogenesis by aberrantly regulating hematopoietic stem cell functions. *Blood*, *121*(26), 5203–5207. https://doi.org/10.1182/blood-2012-12-475863

Wang, T., Pine, A. R., Kotini, A. G., Yuan, H., Zamparo, L., Starczynowski, D. T., Leslie, C., & Papapetrou, E. P. (2021a). Sequential CRISPR gene editing in human iPSCs charts the clonal evolution of myeloid leukemia and identifies early disease targets. *Cell Stem Cell*, *28*(6), 1074-1089.e7. https://doi.org/10.1016/j.stem.2021.01.011

Wang, T., Pine, A. R., Kotini, A. G., Yuan, H., Zamparo, L., Starczynowski, D. T., Leslie, C., & Papapetrou, E. P. (2021b). Sequential CRISPR gene editing in human iPSCs charts the clonal evolution of myeloid leukemia and identifies early disease targets. *Cell Stem Cell*, *28*(6), 1074-1089.e7. https://doi.org/10.1016/j.stem.2021.01.011

Watcham, S., Kucinski, I., & Gottgens, B. (2019). New insights into hematopoietic differentiation landscapes from single-cell RNA sequencing. *Blood*, *133*(13), 1415–1426. https://doi.org/10.1182/blood-2018-08-835355

Weinstein, I. B., & Joe, A. (2008). Oncogene Addiction. *Cancer Research*, *68*(9), 3077–3080. https://doi.org/10.1158/0008-5472.CAN-07-3293

Winters, A. C., & Bernt, K. M. (2017). MLL-Rearranged Leukemias—An Update on Science and Clinical Approaches. *Frontiers in Pediatrics*, *5*. https://doi.org/10.3389/fped.2017.00004

Yu, V. W. C., Yusuf, R. Z., Oki, T., Wu, J., Saez, B., Wang, X., Cook, C., Baryawno, N., Ziller, M. J., Lee, E., Gu, H., Meissner, A., Lin, C. P., Kharchenko, P. V., & Scadden, D. T. (2016). Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. *Cell*, *167*(5), 1310-1322.e17. https://doi.org/10.1016/j.cell.2016.10.045

Yuen, K., Liu, Y., Zhou, Y., Wang, Y., Zhou, D., Fang, J., & Xu, L. (2023). Mutational landscape and clinical outcome of pediatric acute myeloid leukemia with 11q23/ *KMT2A* rearrangements. *Cancer Medicine*, *12*(2), 1418–1430. https://doi.org/10.1002/cam4.5026

Zhang, X., Zhang, H., Chen, L., Wang, M., Xi, J., Liu, X., Xie, M., Li, D., Gulati, E. S., Gong, S., & Wang, H. (2018). Arsenic trioxide and all-trans retinoic acid (ATRA) treatment for acute promyelocytic leukemia in all risk groups: Study protocol for a randomized controlled trial. *Trials*, *19*(1), 476. https://doi.org/10.1186/s13063-018-2812-3

Zhang, X.-H., Tee, L. Y., Wang, X.-G., Huang, Q.-S., & Yang, S.-H. (2015). Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Molecular Therapy - Nucleic Acids*, *4*, e264. https://doi.org/10.1038/mtna.2015.37

Zhao, Z., Shang, P., Mohanraju, P., & Geijsen, N. (2023). Prime editing: Advances and therapeutic applications. *Trends in Biotechnology*, *41*(8), 1000–1012. https://doi.org/10.1016/j.tibtech.2023.03.004