Department of Biomedical Sciences
University of Veterinary Medicine Vienna

Institute of Medical Biochemistry
(Head: Univ.-Prof. Dr. rer.nat. Florian Grebien)

The role of *GATA2* in *CEBPA*-mutated AML

Bachelor Thesis

For obtaining the degree

Bachelor of Science (BSc)

University of Veterinary Medicine Vienna

submitted by

Edwin Rzepa

Vienna, May 2021

Supervisor:

Elizabeth Heyes, PhD, Department of Biomedical Sciences, University of Veterinary Medicine Vienna, Austria

Reviewer:

Barbara Maurer, PhD, Department of Biomedical Sciences, University of Veterinary Medicine Vienna, Austria

Acknowledgements

First of all, I want to thank Univ.-Prof. Dr. rer.nat. Florian Grebien for allowing me to be part of the work group and giving me the opportunity to conduct my bachelor thesis at this institute.

In particular, I want to thank my supervisor Elizabeth Heyes for the superb supervision and all the support and advice she has given me.

Additionally, I want to thank all the members of the FG-Team for providing a welcoming and kind working environment. I also want to thank Martin Piontek for sharing his practical knowledge and encouragement.

Lastly, I want to thank my family for all the great support during my time conducting this bachelor thesis.

Table of content

1. Introduction

1.1 The hematopoietic system

The mammalian hematopoietic system is a complex and intertwined network with many different regulatory mechanisms. It is located in the bone marrow, the spleen, the lymph nodes as well as in the thymus and ensures the production of both white and red blood cells. The development and maintenance of the blood cell pool is initiated by the proliferation and differentiation of long-term pluripotent hematopoietic stem cells (LT-HSCs) (Seita and Weissman, 2010). These cells have the ability to self-renew and give rise to short-term pluripotent hematopoietic stem cells (ST-HSCs) with decreased self-renewal potential, which further differentiate into multipotent progenitors (MPPs). MPPs can either differentiate into common lymphoid progenitors (CLPs), which give rise to all types of lymphocytes, or into common myeloid progenitors (CMPs) (Seita and Weissman, 2010) (Figure 1). These cells are also called myeloblasts and further differentiate into megakaryocyte/erythrocyte progenitors (MEPs) or granulocyte/macrophage progenitors (GMPs), which give rise to erythrocytes, platelets, macrophages and granulocytes (Figure 1) (Seita and Weissman, 2010). Although current advances in single-cell RNA sequencing challenge the strict order of the hemopoietic tree, it is obvious that the development of blood cells is organized in a hierarchical manner in which cells continuously lose their self-renewal capacity in order to further differentiate.

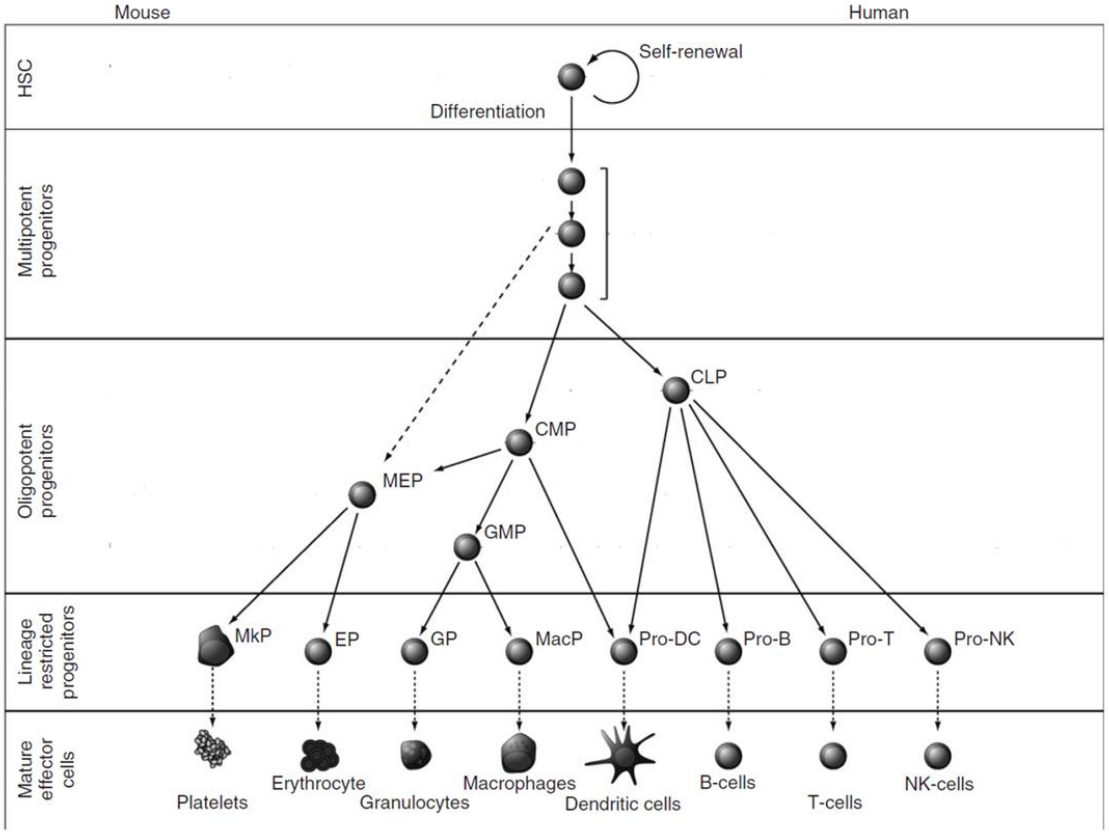


Figure 1: The hematopoietic hierarchy (edited) (Seita and Weissman, 2010)

hematopoietic stem cell (HSC), common lymphoid progenitor (CLP), common myeloid progenitor (CMP), megakaryocyte/macrophage progenitor (MEP), granulocyte/macrophage progenitor (GMP), megakaryocyte progenitor (MkP), erythrocyte progenitor (EP), granulocyte progenitor (GP), macrophage progenitor (MacP), dendritic cell (DC), natural killer (NK)

1.1.1 The hematopoietic niche

To ensure that the hematopoietic system can produce the large amount of progenitors and mature cell types, several extrinsic factors as well as intrinsic factors contribute to the correct development of blood cells subsets. For example, it was shown that hematopoietic stem cells (HSCs) interact with several other cell types that are present in the bone marrow. Studies show that the proliferation and differentiation of HSCs are heavily influenced by the interaction of HSCs with osteoblasts, stromal and vascular cells, which secrete diverse cytokines and chemokines forming a niche that ensures a correct development of the hematopoietic compartment (Orkin and Zon, 2008).

1.1.2 Role of transcription factors in haematopoiesis

The intrinsic component consists of a large network of transcription factors (TFs) which modulates diverse genetic programs determining cell proliferation and differentiation. These sets of nuclear DNA-binding proteins are involved in repression or activation of diverse genes, which promote or repress differentiation of certain lineages and/or are crucial for the effector function of mature cell types. For instance, the formation of myeloid cells is orchestrated by set of TFs including the CCAAT/enhancer binding proteins *C/EBPα*, *C/EBPβ* and *C/EBPε* together with other TFs, such as RUNX1 and PU.1 (Rosenbauer and Tenen, 2007). By regulating the expression of cytokines such as the granulocyte colony-stimulating factor (G-CSF), the C/EBP-family of TFS modulates the differentiation of granulocyte progenitors into granulocytes (D.-E. Zhang et al., 1997). In addition, C/EBPα plays a major role in the commitment of CMPs into GMPs (P. Zhang et al., 2004.). These findings suggest that many TFs have specific roles that depend on the differentiation stage of a given lineage. In addition, the expression level of a given TF must be tightly regulated to ensure correct differentiation. The expression of almost every TF is regulated through a combination of *cis*-regulatory elements in the respective promoter and enhancer sequences. These regions contain binding sites for other TFs allowing for a complex crosstalk and ensuring tight regulation of transcriptional control (de-Leon and Davidson, 2007). Furthermore, the activity of a given TF can be fine-tuned via post-transcriptional protein modifications, translational variation or through protein-protein interactions between different TFs and cofactors (Rosenbauer and Tenen, 2007). Another intrinsic element are signalling pathways which act as a link between extrinsic factors and intrinsic components, ensuring correct haematopoiesis. They are activated by extrinsic

stimuli such as growth factors and ensure a coordinated activation of diverse transcription factors. The signalling pathways instructing lineage differentiation of HSCs include pathways like the BMP and NOTCH pathway which activate important transcriptional programmes (Dzierzak and Bigas, 2018).

1.1.3 Epigenetic regulation of haematopoiesis

Another important aspect in the regulation of hematopoietic differentiation is the accessibility of diverse genes and their regulatory elements in the context of chromatin organisation and modification. Chromatin is a core component of the nucleus and consist of histones and other proteins associated to DNA, forming so-called nucleosomes. It packages DNA molecules into a more condensed form. Based on the density of the chromatin organisation, one distinguishes between heterochromatin and euchromatin. The formation of heterochromatin hinders or completely suppresses the transcription of genes by making the DNA less accessible for TFs and the transcriptional machinery. In contrast, euchromatic regions allow more access to gene regulatory elements like enhancers and thus are associated with increased transcriptional activity of a given gene. The packaging and organisation of chromatin is regulated by a set of histone- and DNA-modifying enzymes, which covalently modify histone tails and DNA to regulate transcriptional activity. These modifications, which are also called histone marks, include methylation, acetylation and many more (Kouzarides, 2007). For instance, DNA-methyltransferases (DNMts), and TET enzymes regulate gene expression in HSCs through DNA-methylation and demethylation of gene regulatory elements and thereby ensure correct differentiation (Gore and Weinstein, 2016). Notably, histone modifications and DNA methylation patterns can be linked to functional elements like enhancers (Zhou, Goren, and Bernstein, 2011). For example, it was shown that during haematopoiesis the chromatin structure changes to expose lineage-specific enhancer regions for particular blood lineages while other enhancers are only active in progenitor stages (Lara-Astiaso et al., 2014). This finding underlies the importance of chromatin dynamics and epigenetic modifications in haematopoiesis. TFs play a major role in changing the chromatin landscape through their interaction with so-called enhancer complexes, which consist of several proteins with chromatin-remodelling and -modifying activity, resulting in site-specific accessibility for TFs and thereby increasing transcriptional activity (Bonifer et al., 2008).

1.2 Acute myeloid leukaemia (AML)

Disruption or deregulation of the regulatory mechanisms of blood cell development can lead to uncontrolled proliferation and aberrant cell differentiation, which can lead to the development of haematological malignancies, such as acute myeloid leukaemia (AML). AML is characterized by the abnormal growth and proliferation of myeloid progenitors, which acquire self-renewal capabilities, leading to the accumulation of immature myeloblasts in the bone marrow and extramedullary tissue. This hinders the production of terminally differentiated blood cells, resulting in bone marrow failure, neutropenia and thrombocytopenia (Khwaja et al., 2016). AML is present across all ages but occurs predominantly in the elderly, with a median age above 69 years (Juliusson et al., 2009). The standard treatment of AML consists of induction chemotherapy using cytarabine in combination with other cytostatics over a timespan of seven to ten days followed by a consolidation therapy mainly consisting of high-dose cytarabine infusions in order to achieve full remission (Tallman, Gilliland, and Rowe, 2005). Chemotherapy is often followed by allogeneic HSC transplantation in patients with a high risk of relapse to enhance a favourable treatment outcome (Kassim and Savani, 2017). Although much research has been performed, life expectancy of AML patients has changed little over the last decade and overall survival has remained low for patients older than 60 years owing to treatment-related mortality and age-related complications (Shah et al., 2013). One reason for the slow progress in AML treatment is the large genomic heterogeneity and complexity of this disease, as well as the development of resistance against several anti-cancer drugs. In addition, relapse after successful treatment is often caused by persisting leukaemia stem cells (LSCs), indicating that these cells are derived from preleukaemic HSCs, which survive treatment and can accumulate new mutations (Jan et al., 2012, Yamashita et al., 2020). LSCs are a subpopulation of mutated and preleukaemic HSCs. They give rise to malignant clones and therefore can maintain and propagate disease even after xenotransplantation (J. C. Y. Wang and Dick, 2005). These findings suggest that LSCs can induce or maintain AML in a hierarchical manner, further highlighting the fact that AML is the result of deregulated haematopoiesis (Figure 2). Therefore, it is essential to study the mutations and defective mechanisms in LSCs to establish new and more efficient treatment options. As the term AML comprises a large group of heterogeneous haematological malignancies the World Health Organization (WHO) classified several types of AML based on phenotypes and genetic aberrations (Hwang, 2020). Many factors underlying the development of acute leukaemia have been previously described and include changes in the microenvironment, mutations in

TFs, which instruct lineage differentiation and HSC quiescence as well as mutations in epigenetic regulators (Yamashita et al., 2020).

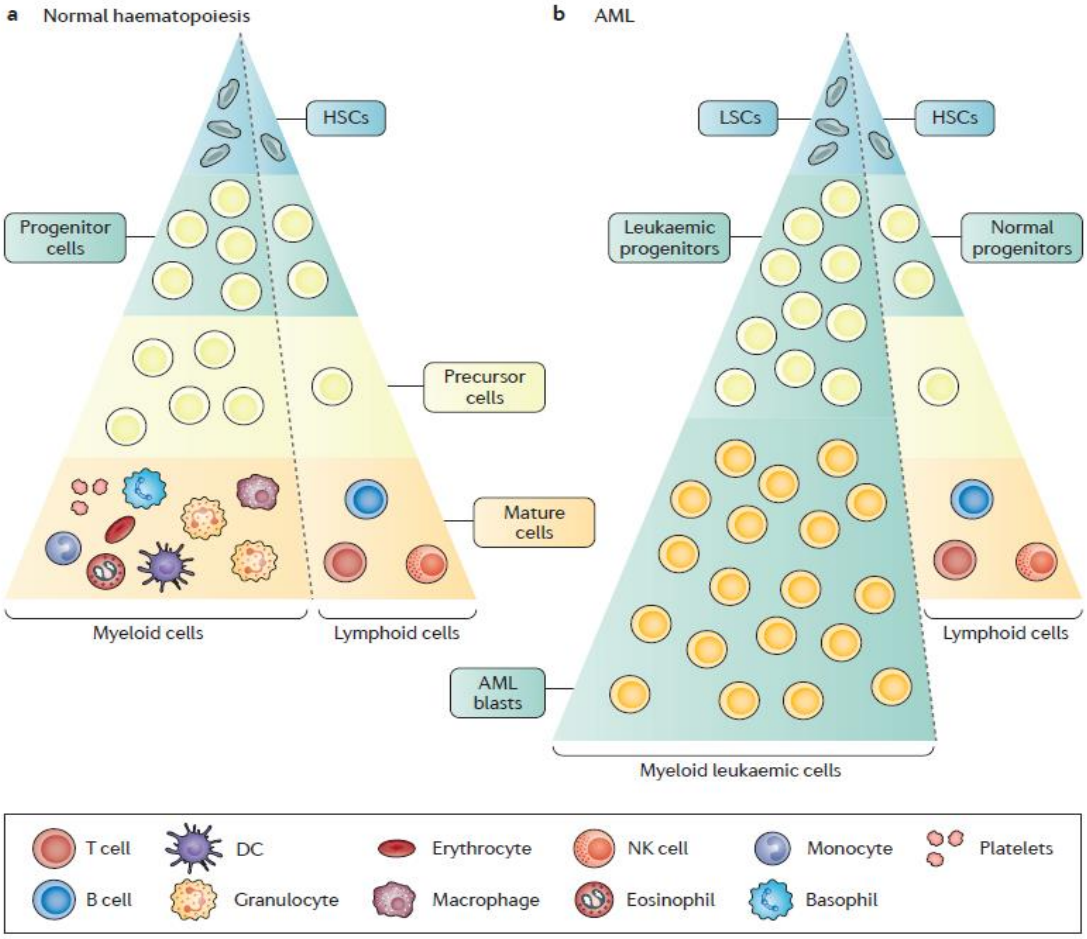


Figure 2: Normal and leukemic haematopoiesis (Khwaja et al., 2016)

(A) Normal haematopoiesis. Hematopoietic stem cells (HSCs) differentiate into progenitor cells which have a reduced self-renewal capacity. These progenitor cells then differentiate into more committed precursor cells which give rise to diverse types of blood cells.
(B) Disturbed haematopoiesis in acute myeloid leukaemia. Leukemic stem cells (LSCs) give rise to malignant leukemic progenitors and myeloblasts, which lack the ability to terminally differentiate into blood cells, such as granulocytes.
hematopoietic stem cells (HSC), leukemic stem cell (LSC), dendritic cell (DC), natural killer (NK)

1.3 Role of *CEBPA* in haematopoiesis

The CCAAT/enhancer-binding protein alpha (C/EBPα) belongs to the family of CCAAT/enhancer binding proteins, which are TFs that regulate differentiation in diverse tissues. The C/EBPα protein harbours two transcriptional activation domains (TADs), a DNA-binding region and a highly conserved leucine-rich zipper dimerization domain enabling it to dimerize and interact with co-activators and repressors in a cell-type-specific manner (Lekstrom-Himes and Xanthopoulos, 1998). C/EBPα plays a major role in instructing myeloid lineage differentiation and regulating HSC proliferation including the differentiation of CMPs into GMPs. The intronless *CEBPA* gene is located on chromosome 19q13.1 and encodes an mRNA that harbours two translation initiation codons. Thus, this mRNA can be translated into either a 42 kDa protein (p42) or into a shorter 30 kDa protein (p30) isoform.

1.3.1 *CEBPA* mutations

Mutations in the *CEBPA* gene are found across the whole coding region and represent insertions, deletions as well as missense and nonsense mutations. However, *CEBPA*-mutations cluster around two hotspots located in the N-terminus and C-terminus of the gene. C-terminal mutations affect the leucine zipper region and dimerization domain, resulting in disrupted DNA-binding and compromised dimerization of the full-length protein p42 (Figure 3) (Kato et al., 2011). On the other hand, N-terminal frame-shift mutations lead to a premature stop of translation, resulting in the increased expression of the shorter p30 isoform (Figure 3) (Pabst et al., 2001). As increased p30 expression might impair p42 function, these mutations are thought to lead to the deregulation of the TF network required for normal haematopoiesis, ultimately ending in leukemogenesis. We and others have shown that the shorter p30 isoform actively upregulates the expression of oncogenes and interacts with epigenetic regulators, concluding that p30 is a gain-of-function variant (Schmidt, Heyes, and Grebien, 2020).

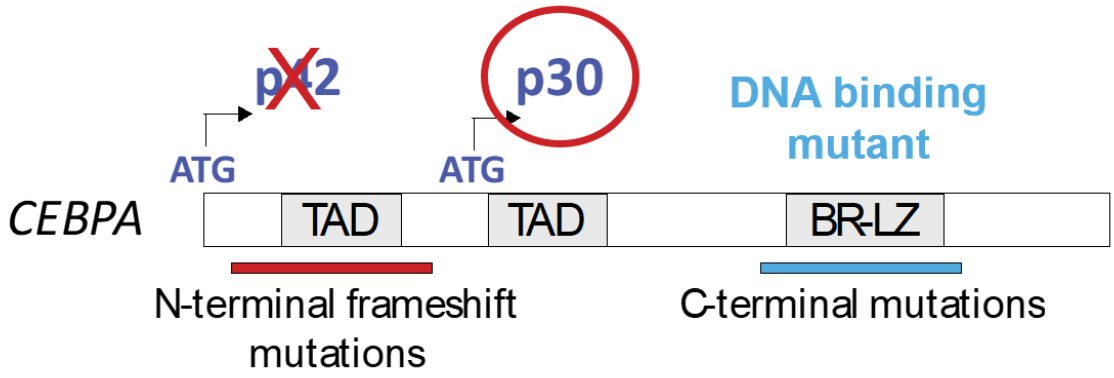


Figure 3: Mutational hotspots in the C*EBPA* gene (Elizabeth Heyes)

N-terminal frameshift mutations cause initiation of translation from the downstream internal ATG codon resulting in the production of the shorter p30 isoform. C-terminal mutations predominantly located in the leucine zipper region result in functional impairment of p42.
transcriptional activation domain (TAD), basic region-leucine zipper (BR-LZ)

1.3.2 Role of *CEBPA* mutations in human AML

CEBPA-mutated AML can be divided into two subclasses based on the mutational patterns of *CEBPA*. AML with one *CEBPA* mutation, either N- or C-terminal, on one allele, but retaining one copy fully functional p42 on the second allele, is called AML with single-mutated *CEBPA* (*CEBPA*^{sm}). AML patients with double-mutated *CEBPA* (*CEBPA*^{dm}) harbour two *CEBPA* mutations on different alleles. In most cases an N-terminal mutation on one allele is combined with a C-terminal mutation on the second allele, resulting in absent wild-type p42 expression (Wouters et al., 2009). Screening of *CEBPA* mutations in a total of 2296 AML patients revealed that *CEBPA* mutations occurred in 244 of these patients (10.6 %) with 140 (6.1 %) *CEBPA*^{sm} cases and 104 (4.5 %) *CEBPA*^{dm} cases. This study demonstrated that patients with *CEBPA*^{dm} had a more favourable outcome (Fasan et al., 2014). However, the presence of additional mutations influences the prognostic outcome. It was shown that *TET2* mutations significantly impaired the outcome of patients with *CEBPA* mutations (Fasan et al., 2014). In summary, these findings underline the importance of *CEBPA* mutations as a prognostic factor and emphasize the need to study the co-occurring mutations in order to understand the development of *CEBPA*-mutated AML.

1.3.3 Models to study AML-associated *CEBPA* mutations

In order to study the molecular mechanisms through which *CEBPA* mutations lead to leukaemia, several mouse models were developed. The use of mouse models allows for a precise investigation of the molecular mechanisms that depend on the introduced mutation. In contrast, cell lines derived from AML patients harbour a variety of additional mutations which could alter particular phenotypes of interest. Yet two cell lines exist that are derived from AML patients with *CEBPA* mutations - KO-52 ($CEBPA^{dm}$) and Kasumi-6 ($CEBPA^{sm}$ C-terminal).

1.3.3.1 The "L"-Allele

Using the Cre-LoxP-System in mice, Kirstetter et al. introduced a translational stop codon in the N-terminus of the murine *Cebpa* gene, naming the allele after its inserted LoxP sites the L-allele. Through intercrossing of mice harbouring the L-allele, the group established a homozygous $Cebpa^{p30/p30}$ (L/L) mouse model, which only expresses p30 (Kirstetter et al., 2008). L/L mice die perinatally, which complicates in vivo studies of the hematopoietic system in this mouse model. Schmidt *et al.* generated hematopoietic cell lines through the isolation and serial replating of foetal liver cells with the L/L genotype (Schmidt et al., 2019).

1.3.3.2 The "K"-Allele

The same group established a knock-in allele resulting in the duplication of the lysine residue K313 at the C-terminal DNA binding domain of the *Cebpa* gene (K-allele) (Bereshchenko et al., 2009). Through combination of the K-allele with the L-allele a new mouse model harbouring bi-allelic N- and C-terminal *Cebpa* mutations could be established (K/L-mice). As this model closely mimics the situation found in AML patients with bi-allelic *CEBPA* mutations, these models allow the study of cancer-inducing *Cebpa* mutations.

1.4 TET2 as an important epigenetic modifier

The Ten-Eleven-Translocation 2 factor (TET2) belongs to the group of TET proteins, which are responsible for the demethylation of DNA in humans and in mice (Ito et al., 2012). Its catalytic domain is located near the C-terminus and consists of a cysteine rich domain and a double-stranded β-helix fold domain which encompasses three Fe-binding sites as well as

one 2-oxoglutarate binding site (Figure 4) (Feng et al., 2019). It catalyses the conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) by using molecular oxygen as substrate to decarboxylate 2-oxoglutarate, leading to the generation of enzyme-bound Fe(IV)-oxo which oxidises 5mC and thereby leads to its hydroxylation (Feng et al., 2019). This process ultimately leads to the demethylation of 5hmC, which results in a more accessible DNA conformation. TET2 is an important epigenetic modifier. For instance, a recent study showed that TET2 interacts with WT1, a TF, which recruits TET2 to its target genes to activate their expression (Y. Wang et al., 2015).

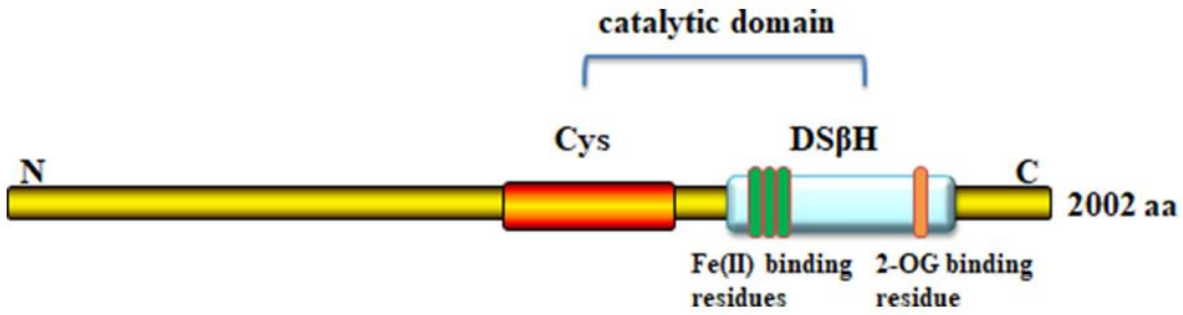


Figure 4: Protein Structure of the catalytic domain of the TET2 enzyme (Feng et al., 2019)

Depicted is the C-terminal catalytic domain consisting of a cysteine rich region (Cys) and the double strand β-helix fold domain (DSβH) containing three Fe-binding sites and one 2-oxoglutarate (2-Og) binding site.

1.4.1 *TET2* mutations in human AML

Mutations in the *TET2* gene are most frequently loss-of-function (LOF) mutations that lead to the development of myeloid malignancies and highlight the importance of *TET2* in haematopoiesis. As TET2 mutations cause global DNA hypermethylation, their oncogenic effects could be the consequence of hypermethylation of important regulatory genes that control haematopoiesis. A study using *Tet2*-deficient mice demonstrated that inactivation of *Tet2* resulted in pleiotropic alterations in both mature and immature hematopoietic compartments affecting myeloid as well as lymphoid lineages (Quivoron et al., 2011). This study also supported the idea that *TET2* mutations do not induce the development of hematopoietic malignancies alone, but rather together with the acquisition of secondary mutations, indicating that *TET2* mutations cooperate with other mutations during oncogenesis. Furthermore, a deep-sequencing study of 318 AML patients showed that *TET2* mutations occurred in 87 (27.4%) of these patients. Most mutations were frameshift or

missense mutations across the conserved regions in the TET2 gene, impairing the catalytic function of TET2 (Weissmann et al., 2012). Interestingly, this study showed that *CEBPA* and *TET2* are frequently co-mutated, which is also supported by other studies (Fasan et al., 2014). In order to investigate the effect of this co-mutation, we analysed RNA-sequencing and ATAC-sequencing data from human AML patients and from mouse models and harbouring the *CEBPA* and *TET2* co-mutations. Integrative analysis of the datasets identified *GATA2* expression as one of few genes that were down-regulated upon *TET2* mutation in all conditions (Figure 5).

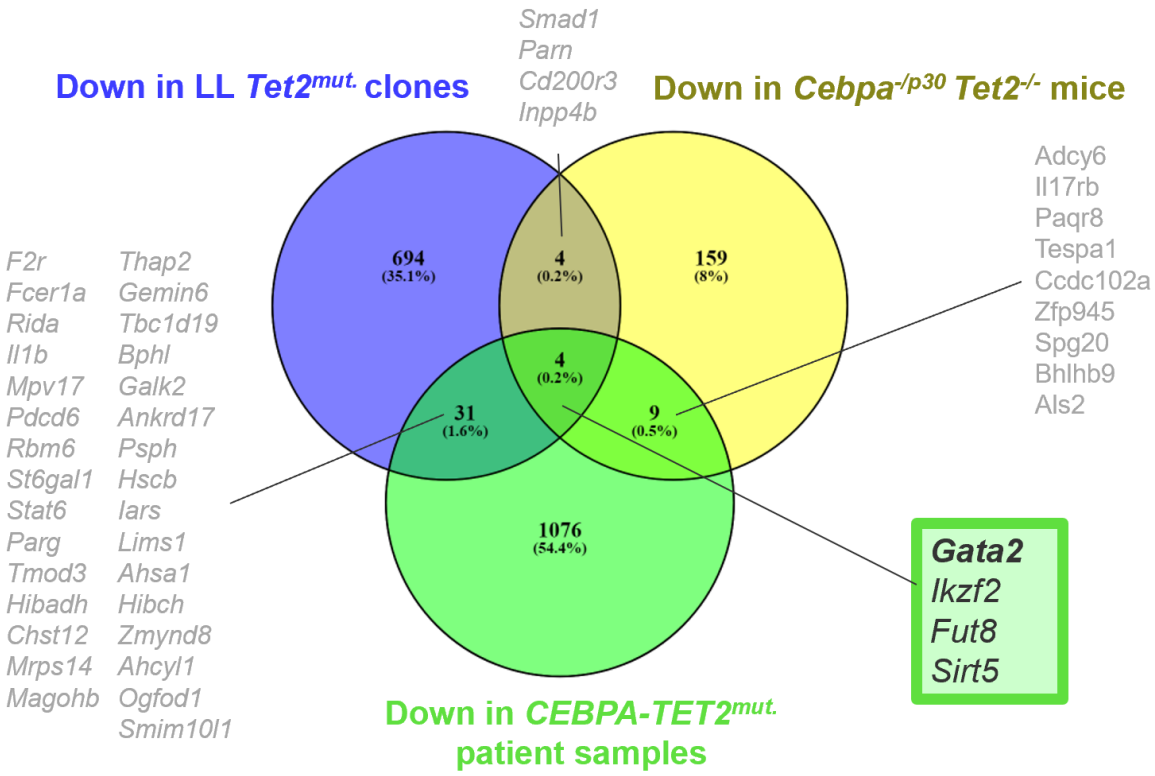


Figure 5: RNA-Sequencing data (Heyes and Wilhelmson, unpublished data)

Depicted are downregulated genes upon *TET2* mutation in *CEBPA*^{mut} *TET2*^{wt} versus *CEBPA*^{mut} *TET*^{mut} AML patients and models. The upper left blue circle depicts the absolute number of downregulated genes and their percentage in LL clones harbouring a *Tet2* mutation, labelled LL *Tet2*^{mut} clones. The upper right yellow circle depicts the absolute number of downregulated genes and their percentage in mouse models harbouring a monoallelic N-terminal *Cebpa* mutation, termed *Cebpa*$^{-/p30}$ *Tet2*$^{-/-}$ mice. The lower green circle depicts the number of downregulated genes and their percentage in patient samples with *CEBPA* and *TET2* co-mutation. Overlapping down-regulated genes are indicated.

1.5 GATA2 in haematopoiesis and leukaemia

GATA2 is a TF that plays a major role in haematopoiesis by instructing differentiation of GMPs as well as MEPs into their committed mature progeny (Orkin and Zon, 2008). It belongs to the GATA family of TFs, which bind DNA through their highly conserved zinc finger domains on the DNA consensus sequence (A/T) GATA (A/G) (Leubolt, Redondo Monte, and Greif, 2020). In a conditional knock-out mouse model, deletion of the zinc finger domains of GATA2 in hematopoietic stem cells resulted in complete loss of adult hematopoietic stem cells and their multilineage potential. This further showcases the importance of GATA2 in haematopoiesis (Menendez-Gonzalez, et al. 2019). Mutations in the *GATA2* gene which functionally impair GATA2 or reduce its expression level have been linked to several human malignancies, including AML (Ostergaard et al., 2011). Screening of *GATA2* mutations in AML patients revealed that these mutations cluster around two conserved zinc finger domains, thereby disrupting the DNA binding ability of GATA2 (Greif et al., 2012). Additionally, *GATA2* mutations also frequently co-occur with biallelic mutated *CEBPA,* resulting in erythroleukemia (Di Genua et al., 2020). A study by Schmidt et al. showed that a *GATA2* knockout led to a rapid loss of *Cebpa*-mutated cells, indicating that *Gata2* expression is essential for the survival and proliferation of these cells. *Gata2* expression was cooperatively upregulated through interaction of p30 with the histone-methyltransferase MLL1 (Schmidt et al., 2019). Unpublished data suggests that reduced *Gata2* expression might lead to a proliferative advantage of *Cebpa*-mutated cells. This demonstrates that *GATA2* has a critical effector role in leukemogenesis.

1.5.1 Enhancers that regulate *GATA2* expression

While *GATA2* mutations can lead to functional impairment of GATA2, the expression levels of *GATA2* also play a major role in the pathogenesis of leukaemia. *GATA2* expression is regulated by two enhancers located - 77 kilobases upstream and + 9.5 kilobases downstream of the *GATA2* gene body and their disruption leads to aberrant haematopoiesis (Bresnick and Johnson, 2019). Like other gene regulatory elements, the accessibility of these enhancers is dependent on epigenetic modifications, such as DNA- or histone methylation. ChIP-sequencing data revealed that p30 binds the -77 kb enhancer of *Gata2,* most likely promoting *Gata2* expression (Figure 6). Inducing a *Tet2* knockout in those cells resulted in reduced chromatin accessibility at the *Gata2* promoter and the enhancer region as

demonstrated by ATAC-sequencing data, thereby possibly decreasing *Gata2* expression (Figure 6).

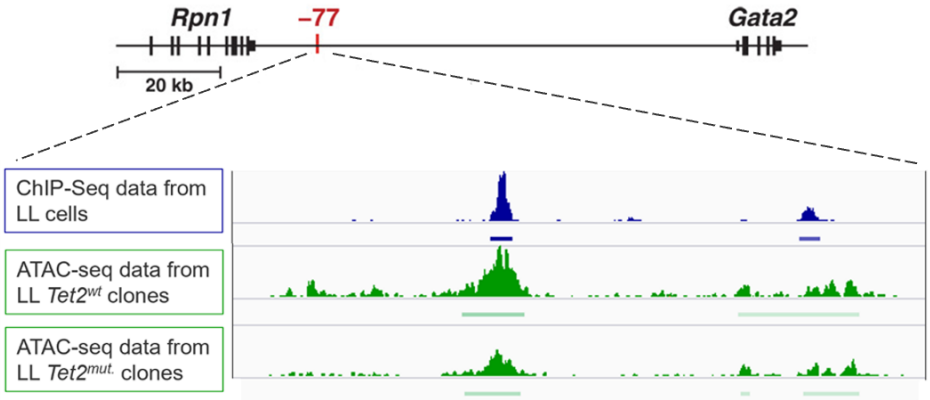


Figure 6: ATAC- and ChIP-sequencing results of the – 77 kb enhancer of *Gata2* (edited) (Johnson et al., 2015) (Elizabeth Heyes, unpublished data)

Depicted above the results is the position of the – 77 kb enhancer of *Gata2*. The peak heights depict the frequency of the ChIP/ATAC reads. The depicted data from the ChIP-sequencing was obtained by using an antibody against p30 in LL cells.

1.6 Methods to study TFs and the epigenome

1.6.1 Chromatin Immunoprecipitation (ChIP) Assay

The Chromatin Immunoprecipitation Assay (ChIP-Assay) is a widely used method to study protein-DNA interactions at specific genomic loci. It consists of three main steps. The first step involves crosslinking of proteins to DNA by formaldehyde. This helps to preserve and strengthen protein-DNA interactions, which might otherwise be lost in downstream processes. Following the fixation, the chromatin is sonicated using ultrasound, producing small DNA fragments (approximately 300-500 bp). After sonication the DNA-protein complexes are immunoprecipitated using an antibody which specifically binds to the protein of interest. The DNA is then eluted from the DNA-protein-antibody complex and can be used for diverse analyses, like next generation sequencing (NGS) or conventional real time PCR (Das et al., 2004). This method allows researchers to study and discover epigenetic marks and TF-binding sites across the whole genome (Mundade et al., 2014). Therefore, the ChIP-

assay is a suitable method to study the complex interplay of TFs and the epigenome, which orchestrates haematopoiesis.

1.6.2 CRISPR interference (CRISPRi)

While the ChIP method allows for analysis of the epigenetic landscape of the cell it is also essential to identify positions in the genome that can be investigated for their functional relevance through site-specific alterations. One way to edit the epigenome is to use dCas9. The dCas9 protein is a nuclease-deficient variant of the CRISPR-associated protein 9 (Cas9) protein, which lacks the ability to induce double-strand breaks in the DNA. As a result, it only binds to DNA and interferes with transcription through steric blockage of the RNA polymerase and thereby represses the transcription of a given gene (Brocken, Tark-Dame, and Dame, 2018). To increase the extent of transcriptional repression, dCas9 can be fused to transcriptional repressor domains like the Krüppel associated box (KRAB), which is found in many KRAB zinc finger DNA binding repressors (Groner et al., 2010). Like the original Cas9 enzyme, dCas9 is guided to its cognate target site by guide RNAs (gRNAs). As gRNAs can be easily designed to target any specific sequence in the genome, this approach allows researchers to modulate the expression of diverse genes.

1.7 Hypothesis and Aims of this Bachelor thesis

Based on the available data, we postulate that *CEBPA* and *TET2* co-mutations decrease *GATA2* expression to a specific level, which results in a proliferative advantage of AML cells. The aims of this study are:

- Confirmation that *Cebpa* and *Tet2* co-mutations lead to a decreased *Gata2* expression *in vitro*. In order to measure the expression of *Gata2* we established a quantitative Real Time PCR (qRT-PCR/qPCR) assay. For this we used RNA isolated from eleven LL *Gata2* knockout clones (LL $Gata2^{mut.}$clones) and compared their expression levels to wild type LL cells (LL $Gata2^{wt}$ clones).

- Confirmation that decreased *Gata2* expression provides a proliferative advantage to *Cebpa*-mutated cells and thereby promotes leukemogenesis. Both competition assays and a growth curve were performed in order to test whether *Cebpa*-mutated cells with lower *Gata2* expression have a proliferative advantage. Eleven LL $Gata2^{mut.}$

clones were used and their growth rates were compared to that of LL *Gata2*^{wt} cells. The competition assays were performed using four cell lines: two LL and two KL cell lines stably expressing Cas9 and dCas9. Each cell line was transduced with a construct containing the green fluorescent protein (GFP) gene and one of 13 single guide RNAs (sgRNAs) designed to target the - 77kb enhancer of *Gata2* with the goal of introducing mutations in the sequence that disrupt *Gata2* expression.

- Determination of the effect of *Gata2* expression levels on the proliferation of *Cebpa*-mutated cells. Due to the essential role of GATA2 in differentiation and self-renewal, we postulate that only a modest reduction of Gata2 levels provides a proliferative advantage to AML cells, while its complete elimination is incompatible with AML cell growth. Thus, we compared the growth rate of eleven LL *Gata2*$^{mut.}$ clones with their expression of *Gata2*.

- Investigation of the relationship between *Gata2* expression and decreased TET2-dependent p30 binding to the - 77 kb enhancer of the *Gata2 gene*. For this we performed a ChIP-qPCR with the aim to show that a *Tet2* knockout leads to decreased p30 binding to the - 77 kb enhancer of *Gata2*, using four LL *Tet2* knockout clones and four LL wild-type clones.

2. Materials and Methods

2.1 Cell culture

2.1.1 Reagents for Cell culture

Reagent	Concentration	Manufacturer	Catalogue Number
Gibco RPMI 1640 Media	1X	Fisher Scientific, U.S	21875034
Gibco Dulbecco's Modified Eagle Media (DMEM)	1X	Fisher Scientific, U.S	11500416
Gibco L-Glutamine	200 mM	Fisher Scientific, U.S	11500626
Gibco Penicillin-Streptomycin	5000 U/ml	Fisher Scientific, U.S	11528876
Fetal Bovine Serum	-	Cytiva Hyclone, U.S	SH30073.03
Gibco Trypsin-EDTA, phenol red	0.25 %	Fisher Scientific, U.S	11560626
Gibco Dulbecco's Phosphate Buffered Saline	1X	Fisher Scientific, U.S	14190144
PBS	ThermoFisher Scientific, U.S	1X	14190144
Polybrene Transfection Reagent	-	Merck Chemicals and Life Science, Germany	TR-1003-G

Table 1: List of reagents used in cell culture

2.1.2 Equipment for cell culture

Equipment	Manufacturer	Catalogue Number
Cell Culture Flask 250 ml	Greiner Bio One, Germany	658175
Cell Culture Dish 100/20 mm	Greiner Bio One, Germany	664160
Cell Culture Multiwell Plate 24 well	Greiner Bio One, Germany	662160

Cell Culture Multiwell Plate 6 well	Greiner Bio One, Germany	657160
Micro test plate 96 well	Sarstedt, Germany	82.1582.001
Thermo Scientific BBD 6220 CO^2 Incubator	Thermo Scientific U.S	10423582
Allegra X-12 and X-12R Benchtop Centrifuges	Beckman Coulter U.S	-
IntelliCyt iQue Screener PLUS	Sartorius	-
Syringe Filter 45	TPP Switzerland	99745

Table 2: List of equipment for cell culture

2.1.3 Maintenance of LL and KL-cells

For this study, we used the LL-cell line (*Cebpa*$^{p30/p30}$) and the KL-cell line (*Cebpa*$^{p30/C\text{-}mut.}$) which were previously established from mouse models for *CEBPA*-mutated AML (Kirstetter et al., 2008, Bereshchenko et al., 2009). The cells were cultivated in suspension in Gibco RPMI 1640 Media (Fisher Scientific) supplemented with 10 % fetal bovine serum (FBS) (Cytiva Hyclon), 1 % Gibco L-Glutamine (200 mM) (Fisher Scientific), 1 % Gibco Penicillin-Streptomycin (5000 U/ml) (Fisher Scientific) and murine Interleukin-3 (5 ng/ml). Cells were kept in 250 ml cell culture flasks (Greiner Bio One) and incubated at 37 °C with 5 % CO_2 and 95 % humidity (BBD 6220 CO_2 Incubator, Thermo Scientific). Cells were examined under the light microscope and split three times a week after counting them.

2.1.4 Maintenance of LentiX cells

For virus production, the cell line Lenti-X 293T was used. The cells were cultivated in Gibco Dulbecco's Modified Eagle Media (DMEM) (Fisher Scientific) supplemented with 10 % FBS (Cytiva Hyclon), 2 % Gibco L-Glutamine (200 mM) (Fisher Scientific) and 1 % Gibco Penicillin-Streptomycin (5000 U/ml) (Fisher Scientific) in 100/20 mm dishes (Greiner Bio One). Cells were incubated at 37 °C with 5 % CO^2 and 95 % humidity (BBD 6220 CO_2 Incubator, Thermo Scientific). After examination in the light microscope the cells were split. For splitting, cells were washed in phosphate-buffered saline (PBS) (Fisher Scientific) and treated with Gibco Trypsin-EDTA (0.25 %), phenol red (Fisher Scientific) in order to disrupt

the adhesion of cells to the dish surface. After incubation with trypsin, the cells were resuspended in fresh media and transferred into a new 100/20 mm dish.

2.1.5 Cell pelleting

The cells were pelleted by centrifugation at 300 x g for 5 minutes at 21 °C then washed with PBS and centrifuged again at 300 x g for 5 minutes at 21 °C and stored for further use at -20 °C or -150 °C (Allegra X-12 Centrifuge, Beckman Coulter).

2.1.6 Generation of LL *Gata2*$^{mut.}$ knockout clones

To mimic the cooperative downregulation of *Gata2* expression by the p30 isoform with TET2 loss, a *Gata2* knockout, using CRISPR-Cas9 with sgRNAs targeting the *Gata2* gene, was generated prior to this bachelor thesis. The recombinant endonuclease Cas9, a crRNA targeting *Gata2* and a tracrRNA were introduced into LL-cells as a ribonucleoprotein (RNP) by electroporation. After electroporation the cells were passaged in methylcellulose. Clones that outgrew the cell pool were screened for mutations and stored at - 180 °C for further use. In total, eleven LL *Gata2*$^{mut.}$ clones and the LL *Gata2*^{wt} cells were used to conduct the experiments described in this thesis.

2.1.7 Generation of LL *Tet2*^{mut} knockout Clones and LL *Tet2*^{wt} clones

In order to analyse the effects of a *Cebpa* and *Tet2* co-mutation, a *Tet2* knockout was introduced into LL-cells in advance of this bachelor thesis. The knockout was established by electroporation of LL cells with RNPs containing recombinant SpCas9, a crRNA targeting *Tet2* and a tracrRNA. The sgRNAs were designed to guide the Cas9 enzyme to exon 3 of the *Tet2* gene in order to induce a mutation which should lead to a reduced expression or functional impairment of the TET2 protein. In total, four wild type LL-clones and four *Tet2* knockout clones were established (Table 3). The *Tet2* knockout was verified by PCR as well as by sequencing of PCR amplicons in advance of this bachelor thesis. The data were analysed using the TIDE web tool (Brinkman et al., 2014).

Clone name	Clone Type	Knockout
LL *Tet2*^{wt} clone 1 (LL_2)	Wild type	-

LL *Tet2*^{wt} clone 2 (LL_3)	Wild type	-
LL *Tet2*^{wt} clone 3 (LL_6)	Wild type	-
LL *Tet2*^{wt} clone 4 (LL_8)	Wild type	-
LL *Tet2*$^{mut.}$ clone 1 (4_9)	*Tet2* Knockout	heterozygous (+1/wt)
LL *Tet2*$^{mut.}$ clone 2 (4_15)	*Tet2* Knockout	homozygous (+1/+1)
LL *Tet2*$^{mut.}$ clone 3 (4_19)	*Tet2* Knockout	homozygous (+1/+1)
LL *Tet2*$^{mut.}$ clone 4 (4_20)	*Tet2* Knockout	homozygous (+1/+1)

Table 3: List of *Tet2* clones

2.2 *Gata2* knockout genotyping PCR

2.2.1 Reagents for PCR

Reagents	Concentration	Manufacturer	Catalogue Number
TaKaRa LA Taq DNA Polymerase (Mg^{2+} free buffer)	125 U/vial	TaKaRa Bio, Japan	RR002A
$MgCl^{2+}$	25 mM	TaKaRa Bio, Japan	-
dNTP Mixture	2.5 mM	TaKaRa Bio, Japan	4030
LA PCR Buffer (Mg2+ free)	10X	TaKaRa Bio, Japan	9152AM
Dimethyl sulfoxide (DMSO)	99.5 %	Sigma Aldrich, U.S	D4540-500ML
Agarose BioReagent, for molecular biology	99 %	Sigma Aldrich, U.S	A9539-500G
GelGreen Nucleic Acid Gel Stain	10000X	Biotium, U.S	41004
Gel Loading Dye, Purple	6X	NEB, U.K	B7024S
Quick-Load Purple 100 bp DNA Ladder	50 µg/ml	NEB, U.K	N0551S
UltraPure DNA Typing Grade TAE	50X	LifeTech, Austria	24710030

Buffer-1 L

Table 4: List of reagents for PCR

2.2.2 Equipment for PCR

Equipment	Manufacturer	Catalogue Number
S1000 Thermal Cycler with 96-Well Fast Reaction Module	Bio Rad, U.S	1852196
Spark Multimode Microplate Reader	TECAN, Switzerland	-
MiniPex 3 in 1 Kit	IMP, Austria	-
Quick-gDNA Miniprep Kit	Zymo Research, U.S	D3025

Table 5:List of equipment for PCR

In order to validate the *Gata2* knockout a genotyping polymerase chain reaction (PCR) was designed to amplify the target regions in which the mutation was induced. The amplified regions were sequenced and analysed using the TIDE web tool (Brinkman et al., 2014). Primers were designed using the UCSC Genome Browser and the NCBI tool “primer blast”. The primer sequences are listed in Table 6. The cells were harvested by centrifugation and the DNA was extracted using a Quick-gDNA Miniprep Kit (Zymo Research) following the manufacturer's protocol using 500 µl gDNA lysis buffer and eluting the DNA in 60 µl distilled water. DNA concentration was measured with a Spark multimode microplate reader (Tecan). For the PCR reaction the LA Taq DNA polymerase (TaKaRa) was used according to the manufacturer’s protocol (Table 7). The PCR was performed using a thermocycler (S1000 Thermal Cycler, Bio Rad) with the settings and conditions listed below in Table 7 and Table 8.

Primer	Sequence	Manufacturer
mGata2_01L_fw	TTTCCGGGTAACTTGCTGCT	IDT
mGata2_01L_rev	TCAGGTGGTGAAGTGTCTGC	IDT

Table 6: List of primers used to amplify the region of interest

Reagents	Stock	Final concentration/Volume
LA PCR Buffer (Mg^{2+} free)	10X	1X
$MgCl^2$	25 mM	2.5 mM
dNTP mixture	25 mM	2.5 mM
forward Primer	10 µM	1 µM
reverse Primer	10 µM	1 µM
LA Taq DNA Polymerase	5 units/µl	1 unit/µl
DMSO	-	2.5 µl
ddH_2O	-	15 µl
cell DNA	-	100 ng

Table 7: PCR reaction mix

Step	Temperature	Time	Cycles
Denaturation	94 °C	5 min	
	94 °C	15 sec	34x
Annealing	57 °C	20 sec	34x
Elongation	68 °C	25 sec	34x
	68 °C	5 min	

Table 8: PCR conditions and settings

To verify that the region of interest was amplified, an aliquot of the PCR product was separated by gel electrophoresis on a 1.5 % agarose gel. The gel was made by resuspending 1.5 g agarose (Sigma-Aldrich) in 100 ml TRIS-Acetate-EDTA-buffer (TAE-buffer) (LifeTech) and heating it up until the agarose was completely dissolved. In order to stain the DNA 10 µl of 2000X GelGreen (Biotium) was added. The PCR product was loaded with 6X purple gel loading dye (NEB) in a ratio of 1:6. To verify the band size a 100-base pair ladder was used (NEB). Gel electrophoresis was performed at 110 V for 35 minutes. The PCR samples were purified using the MiniPEX 3 in 1 PCR purification kit (IMP). The purification was performed following the manufacturer's protocol and DNA was eluted in 30 µl elution buffer. For sequencing, the primer mGata2_01L_rev and 80 ng of the PCR product were used and sent to Microsynth AG. The results were analysed using the web tool TIDE (Brinkman et al., 2014). TIDE uses a decomposition algorithm to determine insertions and

deletions and their frequency from resulting mixed chromatograms by comparing them to the wild-type, untargeted sequence (Brinkman et al., 2014).

2.3 Growth curve of LL *Gata2*$^{mut.}$ clones

In order to investigate any changes in proliferation kinetics upon decreased *Gata2* expression, we performed a growth curve over a time-span of 40 days. The goal was to determine the growth rate and the cumulative cell number of eleven LL *Gata2*$^{mut.}$ clones in comparison with LL *Gata2*^{wt} cells. The LL *Gata2*^{mut} clones and LL *Gata2*^{wt} cells were cultured in 6-well cell culture plates before starting the growth curve (Greiner Bio-One). After acquiring enough cells, the LL*Gata2*$^{mut.}$ clones and the LL *Gata2*^{wt} cells were seeded in 24-well cell culture plates (Greiner Bio-One). Each clone was seeded in triplicates using 500 000 cells per well in a total of 1 ml on day one. Cells were split three times a week after cell counting into new 24-well plates in a total of 1 ml. For measuring the cell number 50 µl of cell suspension was transferred into a 96-well plate (Sarstedt) and cells were counted using a flow cytometer (IntelliCyt iQue Screener PLUS, Sartorius).

2.4 *Gata2* Expression Analysis using qPCR

2.4.1 Reagents and Chemicals

Reagent	Concentration	Manufacturer	Catalogue Number
SsoAdvanced Universal SYBR Green Supermix	2X	Bio Rad, U.S	1725271
Invitrogen UltraPure DNase/RNase-Free Distilled Water	-	Thermo Scientific, U.S	10977035

Table 9: List of reagents for qPCR

2.4.2 Equipment

Product	Manufacturer	Catalogue Number
Hard-Shell 96-well PCR plates	Bio Rad, U.S	HSP9601
RevertAid H Minus First	Thermo Scientific, U.S	K1631

Strand cDNA Synthesis Kit		
CFX96 Touch Real-Time PCR Detection System	Bio Rad, U.S	1855195
CFX Manager Software	Bio Rad, U.S	1845000
S1000 Thermal Cycler with 96-Well Fast Reaction Module	Bio Rad, U.S	1852196
RNeasy Mini Kit	Qiagen, Netherlands	74104
Spark Multimode Microplate Reader	TECAN, Switzerland	

Table 10: List of equipment for qPCR

To assess the expression levels of *Gata2* in the eleven LL *Gata2*$^{mut.}$ clones as well as in the LL *Gata2*^{wt} cells, a real time quantitative PCR (qPCR) assay was established. For amplifying the cDNA coding for *Gata2*, the forward primer Gata2_2_fw and Gata2_2_rev were used (Table 11). In order to normalize the RNA amount of *GATA2*, the expression of the housekeeping gene *Glyceraldehyde 3-phosphate dehydrogenase* (*Gapdh*) was also quantified. For *Gapdh* the forward primer GAPDH-RT-mouse-F and the reverse primer GAPDH-RT-mouse-R were used (Table 11).

Primer Name	**Primer Sequence**	**Manufacturer**
Gata2_2_fw_	ACAGGCCACTGACCATGAAG	IDT
Gata2_2_rev	AAGGGCGGTGACTTCTCTTG	IDT
GAPDH-RT-mouse-F	AGAAGGTGGTGAAGCAGGCAT	IDT
GAPDH-RT-mouse-R	CGGCATCGAAGGTGGAAGAGT	IDT

Table 11: List of primers for qPCR

Cells were pelleted as outlined above with the exception that pellets were shock frozen in liquid nitrogen before storing them at - 120 °C. The RNA pellets were obtained before the start of the growth curve and at two time points during the growth curve. RNA was extracted using the RNeasy Plus Mini RNA extraction kit (Qiagen) following the manufacturers protocol eluting the RNA in 50 µl RNase free water. cDNA was synthesized using the RevertAid H Minus First Strand cDNA synthesis kit (Thermo Scientific) using 1 µg of sample RNA and 1 µl

of oligo $(dT)_{18}$ primer. cDNA synthesis was performed in a thermocycler (Bio Rad) following the manufacturers protocol including the step for GC-rich RNA. qPCR was performed in a hard-shell 96-Well PCR Plate (Bio-Rad) using a CFX96 Touch Real-Time PCR Detection System (Bio-Rad) in triplicates. The analysis was conducted using the built-in software CFX Manager (Bio-Rad). The data was then further analysed using the 2-ddC(t) method. In order to control for contamination, non-template controls for each primer pair were also analysed in triplicates. The used reaction mix and conditions are listed in Table 12 and Table 13.

Reagent	Stock	Volume
SsoAdvanced Universal SYBR Green Supermix	2X	5 µl
cDNA	200 ng/µl	1 µl
forward primer	10 µM	0.5 µl
reverse primer	10 µM	0.5 µl
UltraPure DNase/RNase-Free Distilled Wate	-	3 µl

Table 12: qPCR reaction mix

Step	Temperature	Time	Cycles
Denaturation	95 °C	2 min	
Denaturation	95 °C	10 sec	45 x
	60 °C	25 sec	45 x
	60 °C	5 sec	
Melting curve	95 °C	5 sec	

Table 13: qPCR conditions and settings

2.5 Cloning of sgRNAs into sgRNA-expression vectors

2.5.1 Cloning reagents

Reagent	Concentration	Manufacturer	Catalogue Number
BsmBI	10000 U/ml	NEB, U.S	R0580
NEB Buffer 3.1	10X	NEB, U.S	B7203S
Antarctic	5000 U/ml	NEB, U.S	M0289S

Phosphatase			
Antarctic Phosphatase Reaction Buffer	10X	NEB, U.S	B0289S
Agarose BioReagent for molecular biology	99 %	Sigma-Aldrich, U.S	A9539
Quick-Load 1 kb DNA Ladder	50 µg/ml	NEB, U.S	N0552S
GelGreen Nucleic Acid Gel Stain	10000X	Biotium, U.S	41004
UltraPure DNA Typing Grade TAE Buffer-1 L	50X	LifeTech, Austria	24710030
T4 Polynucleotide Kinase	10000 U/ml	NEB, U.S	M0201S
T4 Polynucleotide Kinase Reaction Buffer	10X	NEB, U.S	B0201S
T4 DNA Ligase	400000 U/ml	NEB, U.S	M0202S
T4 DNA Ligase Reaction Buffer	10X	NEB, U.S	B0202S

Table 14: List of cloning reagents

2.5.2 Equipment

Product	Manufacturer	Catalogue Number
Thermomixer Comfort 5355	Eppendorf, Germany	5355 000.011
MiniPex 3 in 1 Kit	IMP, Austria	-
Spark Multimode Microplate Reader	TECAN, Switzerland	-
S1000 Thermal Cycler with 96-Well Fast Reaction Module	Bio Rad, U.S	1852196

Table 15: List of equipment for cloning

KL- and LL cells expressing Cas9 and dCas9 were transduced with a lentiviral vector carrying sgRNAs that target the - 77 kb enhancer of *Gata2*. sgRNAs were designed using the web tool CHOPCHOP (Labun et al., 2019) (Table 16).

Oligo	**Sequence**	**Manufacturer**
mGata2_77enh1_rev	caccgTAGGATCGTCCCCATGAAAG	IDT
mGata2_77enh1_rev	aaacCTTTCATGGGGACGATCCTA	IDT
mGata2_77enh2_fwd	caccgTGCATGAATTCCGGTCTCAA	IDT
mGata2_77enh2_rev	aaacTTGAGACCGGAATTCATGCA	IDT
mGata2_77enh3_fwd	caccgTGGTCAGGTGGCGCTTATCA	IDT
mGata2_77enh3_rev	aaacTGATAAGCGCCACCTGACCA	IDT
mGata2_77enh4_fwd	caccgTCACCCGCTCCACGGTGACT	IDT
mGata2_77enh4_rev	aaacAGTCACCGTGGAGCGGGTGAc	IDT
mGata2_77enh5_fwd	caccgGTCGCTGGGCCATTACATTC	IDT
mGata2_77enh5_rev	aaacGAATGTAATGGCCCAGCGACc	IDT
mGata2_77enh6_fwd	caccgCGCCTCATTCTTCGGCAGAC	IDT
mGata2_77enh6_rev	aaacGTCTGCCGAAGAATGAGGCGc	IDT
mGata2_77enh7_fwd	caccgAGGCCCTGGCTATGTTATAC	IDT
mGata2_77enh7_rev	aaacGTATAACATAGCCAGGGCCTc	IDT
mGata2_77enh8_fwd	caccgTGACGTAGCAAGCTGAGCGC	IDT
mGata2_77enh8_rev	aaacGCGCTCAGCTTGCTACGTCAc	IDT
mGata2_77enh9_fwd	caccgCTCCCCCATGGGGTATGTCG	IDT
mGata2_77enh9_rev	aaacCGACATACCCCATGGGGGAGc	IDT
mGata2_77enh10_fwd	caccgCTCCCCCATGGGGTATGTCG	IDT
mGata2_77enh10_rev	aaacCGACATACCCCATGGGGGAGc	IDT
mGata2_77enh11_fwd	caccgCCACCAACCTAGCAGGGATC	IDT
mGata2_77enh11_rev	aaacGATCCCTGCTAGGTTGGTGGc	IDT
mGata2_77enh12_fwd	caccgCGCACAGCCTCCCTTAATTA	IDT
mGata2_77enh12_rev	aaacTAATTAAGGGAGGCTGTGCGc	IDT
mGata2_77enh13_fwd	caccgGAGCGACCTTTGCAGCAGCA	IDT
mGata2_77enh13_rev	aaacTGCTGCTGCAAAGGTCGCTCc	IDT

Renilla-CRISPR guideRNA_1	caccg GGTATAATACACCGCGCTAC	IDT
Renilla-CRISPR guideRNA_2	aaac GTAGCGCGGTGTATTATACC c	IDT
Rpa3_e1.3_f	caccgGCTGGCGTTGACGCGCGCTT	IDT
Rpa3_e1.3_r	aaacAAGCGCGCGTCAACGCCAGCc	IDT
PSMD11-6_fw	caccgTACCGCTCTCACACCGTCCC	IDT
PSMD11-6_rev	aaacGGGACGGTGTGAGAGCGGTAc	IDT

Table 16: List of sgRNAs

The designed sgRNA were cloned into the Lentiguide_Puro_IRES_GFP (LGPIG) backbone. The backbone was digested with the restriction enzyme *BsmBI* (NEB) at 55 °C for 2 hours in an incubator (Eppendorf) (Table 17). After digestion the backbone was dephosphorylated using the Antarctic Phosphatase (NEB) and incubated at 37 °C for 1 hour (Table 17).

Reagent	**Concentration**	**Volume**
LGPIG (Backbone)	1.768 ng/µl	5 µg
BsmBI	10000 U/ml	3 µl
NEB Buffer 3.1	10X	5 µl
ddH_2O	-	40 µl
Antarctic Phosphatase	5000 U/ml	2 µl
Antarctic Phosphatase Reaction Buffer	10X	6 µl

Table 17: Reaction mix for backbone digest

The digested backbone was loaded on a 7 % agarose gel with GelGreen in a ratio of 1:10000 (Biotium). Gel electrophoresis was performed at 110 V for 35 minutes. The band size was determined using the Quick-Load Purple 1 kb DNA Ladder (NEB, U.K). The digested backbone was excised from the gel and purified using the MiniPEX 3 in 1 gel extraction kit (IMP). The extraction was performed following the manufacturers instructions and DNA was eluted in 30 µl of the supplied elution buffer.

Phosphorylation and annealing reactions were performed in a thermocycler (Bio Rad) under the conditions listed below (Table 19) using T4 Polynucleotide Ligase (NEB) (Table18).

Reagent	Concentration	Volume
forward oligo	100 µM	1 µl
reverse oligo	100 µM	1 µl
T4 DNA Polynucleotide Kinase Reaction Buffer	10X	1 µl
T4 Polynucleotide Kinase	10000 U/ml	0.5 µl
ddH_2O	-	6.5 µl

Table 18: Reaction mix for oligo annealing

Temperature	Time
37 °C	30 min
95 °C	5 min
ramp down to 25 °C at a rate of 5 °C/min	

Table 19: Conditions and settings for oligo annealing

The ligation reaction was performed at room temperature for 10 minutes (Table 20). In addition, a negative control ligation using only ddH_2O instead of oligos was included.

Reagent	Manufacturer	Volume
LGPIG (digested)	50 ng/µl	1 µl
annealed oligo (diluted 1:200)	0.2 µM	1 µl
T4 DNA Ligase Reaction Buffer	10X	1 µl
T4 DNA Ligase	400000 U/ml	0.5 µl
ddH_2O	-	6.5 µl

Table 20: Ligation reaction mix

2.6 Transformation

2.6.1 Transformation reagents

Reagent	Concentration	Manufacturer	Catalogue Number
LB-Medium (Lennox)	98.9 %	Carl Roth, Germany	X964.4
LB-Agar (Lennox), granuliert	98 %	Carl Roth, Germany	6671.2
Carbenicillin Dinatriumsalz	88 %	Carl Roth, Germany	6344.2

Table 21: List of reagents for transformation

2.6.2 Equipment

Product	Manufacturer	Catalogue Number
Heratherm General Protocol Microbiological Incubator	ThermoFisher Scientific, U.S	51029334
3015 Analogue Orbital Shaker	LAUDA, Germany	3015
Cell Culture Dish 100/20 mm	Greiner Bio One, Germany	664160
Flacon 14 ml Round Bottom High Clarity PP Test Tube	FALCON, Austria	352059
MiniPex 3 in 1 Kit	IMP, Austria	-
Spark Multimode Microplate Reader	TECAN, Switzerland	-
One Shot Stbl3 Chemically Competent E. coli	ThermoFisher Scientific, U.S	C7373-03

Table 22: List of equipment for transformation

Transformation was performed using Stbl3 competent bacteria (ThermoFisher Scientific). Bacterial cells were thawed on ice and 2 µl of each ligation reaction as well as the negative control were added and mixtures were incubated for 20 minutes on ice. After incubation the

bacteria were heat-shocked at 45 °C for 45 seconds and incubated for 2 minutes on ice. 200 µl of lysogeny broth (LB) media (Carl Roth) was added followed by an incubation at 37 °C for 40 minutes in a shaking incubator at 300 rpm (Eppendorf). 100 µl of the bacteria were seeded on a LB agar plate (Carl Roth) containing carbenicillin dinatriumsalt (100 mg/ml) (Carl Roth) and incubated at 37 °C overnight in a general protocol microbiological incubator (Thermo Scientific). Two single colonies from each plate were picked and inoculated in 14 ml polypropylene round-bottom tubes (Falcon) containing 5 ml LB media and 5 µl Carbenicillin (1:1000). The inoculated bacteria were incubated at 37 °C in an incubator (Thermo Scientific) and shaken at 150 rpm using an orbital shaker (GFL) overnight. Plasmid isolation was performed using the MiniPex 3 in 1 Plasmid prep Kit (IMP) following the manufacturer's instructions. The DNA was eluted in 30 µl supplied elution buffer and DNA concentration was measured using a Spark multimode microplate reader (Tecan).

2.7 Transfection of LentiX cells

In order to produce lentiviral particles containing the 13 sgRNAs the cell line LentiX was used. The helper plasmids psPAX2 (Addgene) coding for the HIV-1 gag-polyprotein and polymerase and pMD2.G (Addgene) coding for VSV-G were used. To exclude cytotoxicity induced by unspecific Cas9 endonuclease activity a construct containing a sgRNA targeting the gene *Renilla luciferase* (*Renilla*) was also transfected in the viral producer cell line. As the *Renilla luciferase* gene is not present in the target cell line, the sgRNA targeting this gene should therefore not induce any mutations. In order to control for Cas9 and dCas9 functionality, two constructs containing sgRNAs targeting essential genes were also transfected into the virus producer cell line. For controlling Cas9 functionality, we used an sgRNA targeting the *replication protein a 3* (*Rpa3*)-gene, which is essential for stabilizing single stranded DNA intermediates during DNA replication. For dCas9, an sgRNA targeting the *Psmd11* gene coding for the 26S proteasome non-ATPase regulatory subunit 11 was used. LentiX cells were seeded in 12-well cell culture plates the day before transfection. The cells were transfected the next day using 0.3 µg of the helper plasmid psPAX2, 0.2 µg of the helper plasmid pMD2.G and 0.69 µg of each sgRNA expression vector mixed in DMEM without supplements. The mixes were incubated with polyethyleneimine (PEI) for 20 minutes and added drop wise to each well. Media was changed the next day. Virus harvest was performed on day four by sterile filtering the supernatant through a 0.45 µm syringe filter

(TPP). The second harvest was performed on the next day using the same materials by pooling the supernatant with the first harvest.

2.8 Infection and Cell count for Competition Assay

Before lentiviral infection the KL- and LL Cas9 and dCas9 cells were seeded in 24 well plates in a concentration of 1 million cells per ml (Greiner Bio-one). The KL and LL cells were infected by adding the harvested virus supplemented with polybrene (Merck) to each cell line in a ratio of 1:5. After adding the virus the cells were spinoculated at 1000 x g for 45 minutes at 21 °C. On day two the cells were spinoculated again using the same amount of virus. On the third day the virus was removed by discarding the supernatant. On day five the first cell count was conducted. The cells were kept in a Biosafety level 2 (BSL2) laboratory for 14 days before they were transferred into a BSL1 laboratory. The cell count was conducted in triplicates by transferring 200 µl of cell suspension into a 96-well plate (Sarstedt) and centrifuging it at 300 x g for 5 minutes at room temperature. The supernatant was discarded and the cells were resuspended in 50 µl PBS. Finally, the cells were counted using a flow cytometer (Sartorius). All work in the BSL2 laboratory was conducted by my supervisor, E.H.

2.9 Chromatin immunoprecipitation (ChIP)

2.9.1 Buffer Reagents

Reagent	Manufacturer	Concentration	Catalogue number
HEPES Buffer	Szabo Scandic, Austria	1 M	LONBE17-737E
Sodium Chloride	Carl Roth, Germany	99.5 %	HN00.3
EDTA	Alfa Aesar, U.S	99 %	A10713.30
EGTA	Lactan Chemikalien und Laborgeräte GmbH, Austria	99 %	3054.1
Formaldehyde Methanol-free	ThermoFisher Scientific, U.S	16 %	28908
Glycine	CHEM-LAB, Belgium	99 %	CL00.0712.1000
SDS	SIGMA-Aldrich, U.S	97 %	11667289001
Triton X-100	PanReac AppliChem,	1 M	APP A4975,0100

	Germany		
Tris	SIGMA-Aldrich, U.S	99 %	T1503
BSA	PAN-BIOTECH, GmbH Germany	99 %	P06-139350
PBS	ThermoFisher Scientific, U.S	1X	14190144
HEPES solution sterile-filtered, BioReagent,	SIGMA-Aldrich, U.S	1 M	H0887-100ML
Lithium chloride	SIGMA-Aldrich, U.S	99 %	L9650
NP40	Calbiochem, Germany	1.06 g/ml	492016
Sodium deoxychelate	SIGMA-Aldrich, U.S	98 %	30970-25G

Table 23: List of buffer reagents for the ChIP-Assay

2.9.2 Other reagents

Reagent	Manufacturer	Concentration	Catalogue number
Dynabeads Protein G	ThermoFisher Scientific, U.S	-	10003D
RNase A, DNase and protease-free	ThermoFisher Scientific, U.S	10 mg/ml	EN0531
Glycogen	Roche, Switzerland	20 mg	10901393001
Proteinase K Solution	Invitrogen, ThermoFisher Scientific U.S	20 mg/ml	25530049

Table 24: List of additional reagents for downstream ChIP steps

2.9.3 Antibody

Name	Target	Antibody type	Manufacturer
C/EBP α antibody (C-18) sc-9314	CEBPA	goat polyclonal IgG	Santa Cruz Biotechnology

Table 25: Used antibody for the ChIP-Assay

2.9.4 Crosslinking

To perform the ChIP a total of 25 million cells of each clone were seeded in 250 ml flasks. The cells were pelleted by centrifugation and resuspended in PBS. Crosslinking was initiated by adding a 11 % formaldehyde solution followed by a 10-minute incubation on a rotator (Table 26). Formaldehyde cross-linking was quenched by adding a glycine solution (2.5 M) and incubating the samples for 5 minutes on a rotator.

Reagent	Stock	Final Concentration
Hepes pH 7.4	1 M	50 mM
Sodium chloride	5 M	100 mM
EDTA	0.5 M	1 mM
EGTA	0.5 M	0.5 mM
Formaldehyde	16 %	11 %
ddH_2O	-	Rest

Table 26: Composition of the 11 % formaldehyde solution

2.9.5 Cell lysis

After performing the cross-linking, the cells were centrifuged at 350 x g for 5 minutes at 21 °C followed by a wash step with PBS and another centrifugation at 350 x g for 5 minutes at 21 °C to remove residual formaldehyde and glycine. The cells were lysed by using a 1 % SDS lysis buffer followed by 30-minute incubation on ice (Table 27).

Reagent	Stock	Final Concentration
SDS	10 %	1 %
EDTA	0.5 M	10 mM
Tris pH 8	1 M	50 mM
ddH_2O	-	Rest

Table 27: Composition of the 1 % SDS lysis buffer

2.9.6 Sonication and clearing of sheared Chromatin

The sonication was performed in a Bioruptor Plus sonication device (Diagenode). The samples were sonicated for a duration of 20 cycles each cycle consisting of 30 sec ultrasound on and 30 sec ultrasound off. After sonication the lysates were diluted in a ratio of 1:10 in dilution buffer (Table 28) and 75 µl of a 10 % Triton X-100 solution was added to each

diluted sample. The lysates were then incubated for 5 minutes on ice before being centrifuged at 10000 x g for 10 minutes at 4 °C. Finally, the supernatant containing the sheared chromatin was transferred into a new tube and incubated on a rotator at 4 °C overnight with 5 µg of the C-18 C/EBPα antibody (Santa Cruz Biotechnology). Before adding the antibody, an aliquot of each sample was removed to store input DNA, which is used for normalization purposes. Also, the concentration of the enriched DNA is not solely dependent on the presence of the target protein/epitope, which is bound by the antibody to precipitate the DNA, but also on the accessibility of the target protein. The conditions can vary from sample to sample and strongly influence the enrichment of a given DNA sequence/fragment. Another aspect, which has to be considered, is the initial DNA amount in the sample before immunoprecipitation. Therefore, the enrichment of a given DNA fragment is always normalized to the input DNA.

Reagent	Stock	Final Concentration
SDS	10 %	0.01 %
Triton X-100	10 %	1.1 %
EDTA	0.5 M	1.2 mM
Tris pH 8	1 M	16.7 mM
NaCl	5 M	167 mM
ddH_2O	-	Rest

Table 28: Composition of the dilution buffer

2.9.7 Blocking of magnetic beads and immunoprecipitation

Before conducting the immunoprecipitation, beads were blocked to reduce unspecific binding. For this, beads were washed in a blocking solution and collected using a magnetic stand DynaMag-2 Magnet (ThermoFisher Scientific) (Table 29). The blocked beats were then added to the chromatin samples containing the C-18 C/EBPα antibody (Santa Cruz Biotechnology) and the mixture was incubated at 4 °C for 4 hours.

Reagent	Stock	Final Concentration
PBS pH 7.4	10X	1X

BSA	100 %	0.5 %

Table 29: Composition of the blocking solution

2.9.8 Washing of beads and elution of protein-DNA complexes

After incubation the magnetic beads coupled with the C/EBPα antibody were washed five times with a wash buffer and collected using the magnetic stand (Table 30). Following the wash steps the magnetic beads were resuspended in TE solution containing 50 mM NaCl (Table 31). Next the samples were centrifuged at 960 x g for 30 seconds at 4 °C. The supernatant was discarded, and the beads were mixed in elution buffer followed by an incubation at 65 °C for 15 minutes at 900 rpm in an incubator (Eppendorf) (Table 32). After incubation the beads were centrifuged at 16000 x g for 30 seconds at room temperature and the supernatant was transferred into a new tube. The remaining beads were resuspended in elution buffer and again incubated and centrifuged under same conditions as mentioned in the previous step. The supernatant was pooled with the previous one. To reverse the protein-DNA cross-linking a 5 M NaCl-solution was added to each sample followed by an incubation at 65 °C for 14 hours.

Reagent	**Stock**	**Final Concentration**
Hepes-KOH pH 7.4	1 M	50 mM
Lithium chloride	5 M	500 mM
EDTA	0.5 M	0.5 M
NP-40	10 %	10 %
Sodium deoxycholate	10 %	10 %
ddH_2O	-	Rest

Table 30: Composition of the wash buffer

Reagent	**Stock**	**Final Concentration**
TE	1X	0.99X
Sodium chloride	5 M	50 mM

Table 31: Composition of the TE solution

Reagent	Stock	Final Concentration
Tris-HCL pH 8.0	1 M	50 mM
EDTA	0.5 M	10 mM
SDS	10 %	1 %
ddH2O	-	Rest

Table 32: Composition of the elution buffer

2.9.9 DNA Purification

To remove/digest residual RNA, a TE solution and RNase A (0.2 mg/ml) was added and the samples were incubated at 37 °C for 1 hour at 600 rpm. Next, protein digestion was performed by adding 4 µl of proteinase K and incubating the samples at 55 °C for 2 hours at 600 rpm. In the last step the DNA was purified using a MiniPEX 3 in1 PCR purification kit (IMP) by following the manufacturers instruction.

2.10 ChIP-qPCR analysis

2.10.1 Reagents for ChIP-qPCR

Reagents	Manufacturer	Catalogue Number
SsoAdvanced Universal SYBR Green Supermix	Bio Rad, U.S	1725271
UltraPure DNase/RNase-Free Distilled Water	Invitrogen, Fisher Scientific, U.S	10977035

Table 33: List of reagents for ChIP-qPCR

2.10.2 Equipment for ChIP-qPCR analysis

Product	Manufacturer	Catalogue Number
Hard-Shell 96-Well PCR Plate	Bio Rad, U.S	HSP9601B
CFX96 Touch Real-Time PCR Detection System	Bio Rad, U.S	1855195

Table 34: List of equipment for the ChIP-qPCR analysis

After performing the ChIP, a qPCR analysis was performed to identify enriched DNA fragments/sequences. The reagents were pipetted into a 96 well plate and the qPCR analysis was performed in duplicates in a CFX96 Touch Real-Time PCR detection system (Bio Rad) using the following conditions listed below (Table 36) (Table 37). We designed primers to verify C/EBPα binding at the - 77 kb enhancer (Cepba1_77enh_2). As a positive region, we used a region (Cebpa1_1) where C/EBPAα-binding was previously detected (Schmidt et al., 2019). We also included a negative region (neg_reg77enh_1) 10 kb upstream of the – 77 kb enhancer region, where no C/EBPα- binding is expected. All primers are listed in Table 35. The results were analysed using the CFX Manager software (Bio Rad,). The data was then further analysed using the 2-ddC(t) method.

Primer ID	Sequence	Manufacturer
Cebpa1_77enh_2_fwd	AATTCTGGTCAACCGCAAGC	IDT
Cebpa1_77enh_2_rev	TCTCGCATCCGTTACTTGCC	IDT
negreg_77enh_1_fwd	GCCGCCCAGGATGTCTATTA	IDT
negreg_77enh_1_rev	TCGAGGCCGGATTTCCATTT	IDT
Cebpa_1_fwd1	CGCAGGAAAACACACCAGTC	IDT
Cebpa_1_rev1	AAACACAGCCTTCCACGTCT	IDT

Table 35: List of primers used for the ChIP-qPCR

Reagent	Stock	Volume
forward Primer	10 µM	1.25 µl
reverse Primer	10 µM	1.25 µl
DNA	-	0.7 µl
SsoAdvanced Universal SYBR Green Supermix	2X	5 µl
UltraPure DNase/RNase-Free Distilled Water	-	1.8 µl

Table 36: ChIP-qPCR reaction mix

Step	Temperature	Time	Cycles
Denaturation	95 °C	2 min	
Denaturation	95 °C	10 sec	45 x
	60 °C	25 sec	45 x
	60 °C	5 sec	
Melting curve	95 °C	5 sec	

Table 37: ChIP-qPCR conditions and settings

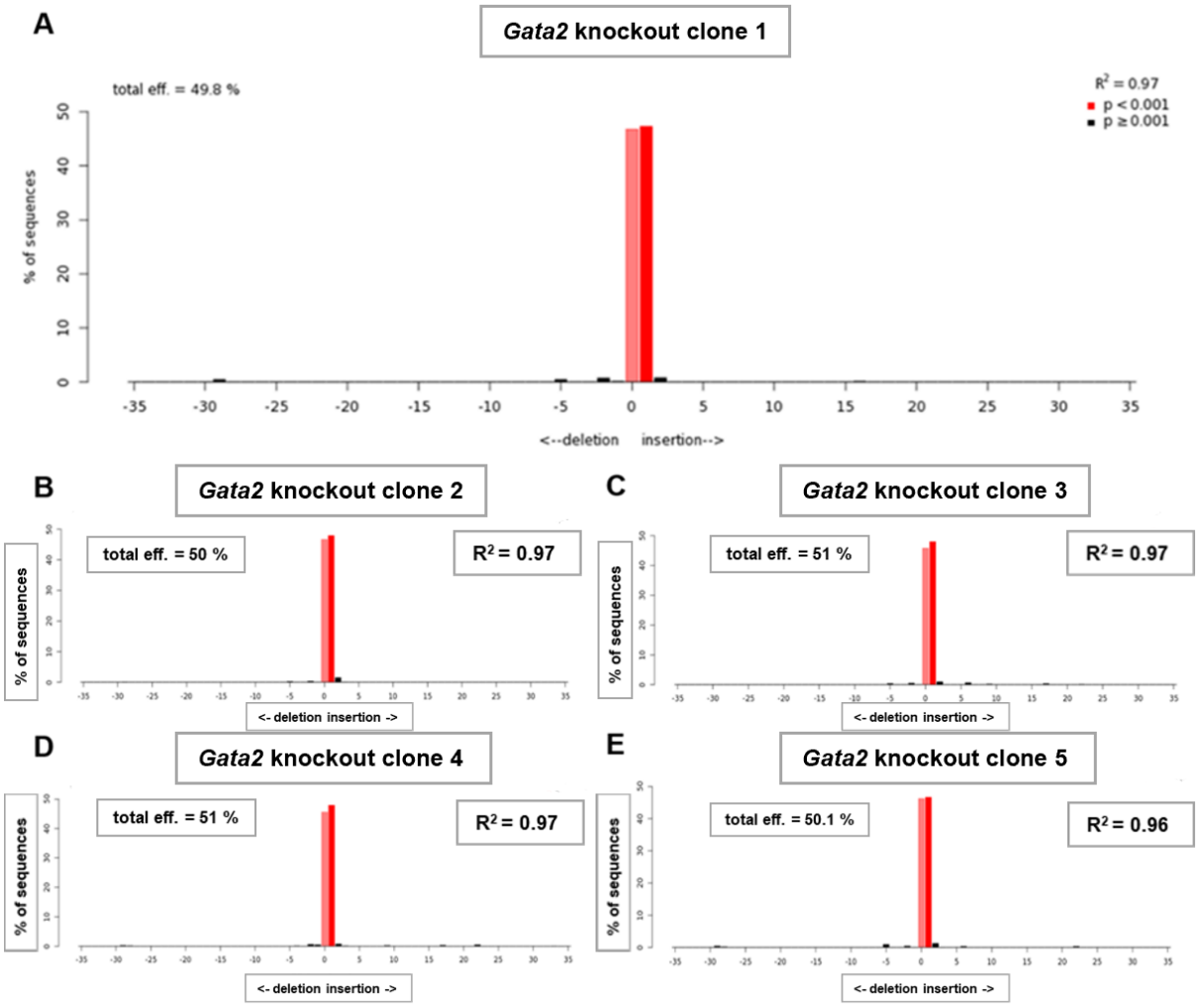
2.11 Statistical data analysis

Statistical data analysis of the obtained experimental data was performed with Microsoft Excel (Microsoft, U.S) and GraphPad Prism (GraphPad, U.S) software.

3. Results

3.1 Validation of heterozygous *Gata2* knock-out in LL *Gata2*$^{mut.}$ clones

To validate *Gata2* knock-outs in LL cells that were previously targeted with sgRNAs against *Gata2*, we extracted the genomic DNA of eleven LL *Gata2*$^{mut.}$ clones, amplified the region of interest via PCR and sequenced the PCR product. Using the TIDE web tool, we analysed the *Gata2* gene sequence of interest and could verify heterozygous frameshift mutations in *Gata2* in all LL *Gata2*$^{mut.}$ clones (Figure 7) (Brinkman et al., 2014). All clones harboured a + 1 insertion expect for LL *Gata2*$^{mut.}$ clone 9 which harboured a + 2 insertion. Using the bioinformatic programme Geneious we could verify that the + 1 insertion of a cytosine nucleotide between two thymines led to a premature stop codon in the zinc finger 2 region of *Gata2* (Figure 7-9). Based on these results, we hypothesize that this mutation leads to a functional impairment of Gata2.



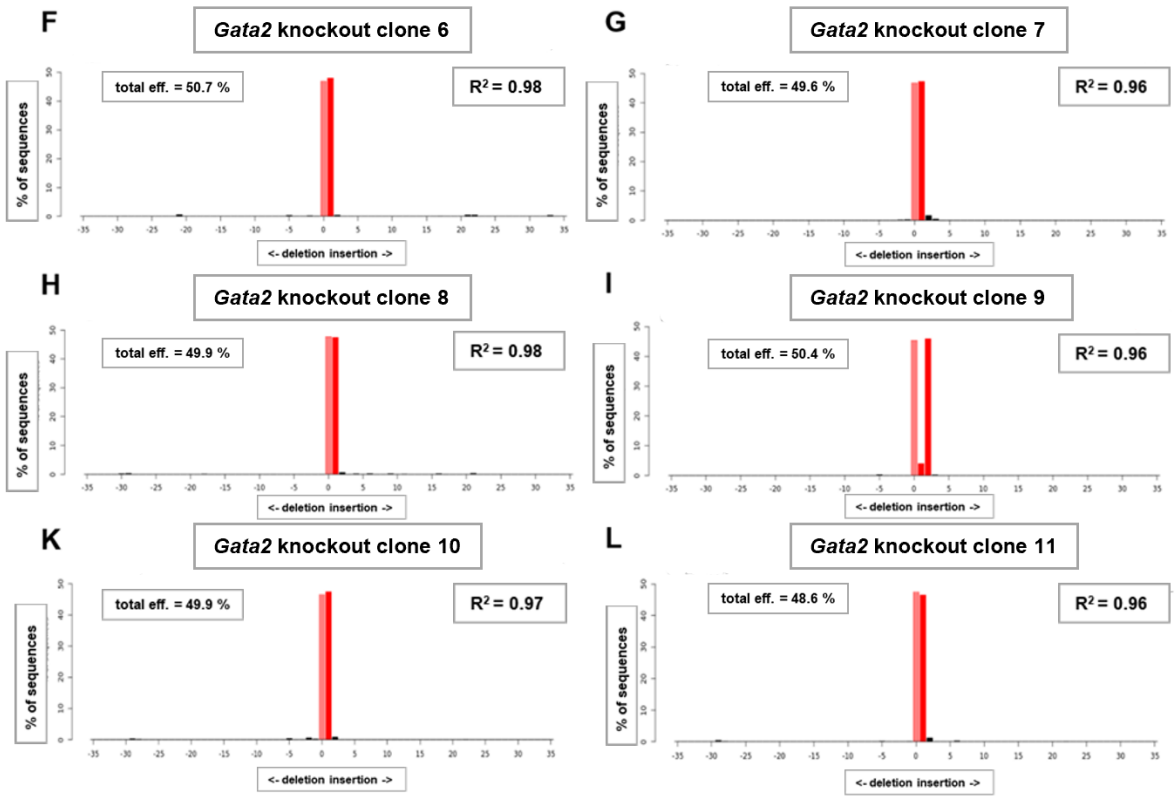


Figure 7: TIDE analysis of the eleven LL *Gata2*^{mut} clones

(A) TIDE analysis of LL *Gata2*$^{mut.}$ clone 1

(B)-(L) TIDE analysis of LL *Gata2*$^{mut.}$ clones 2-11

TIDE analysis was performed on sequencing data derived from PCR-amplicons of the expected cut site of the endonuclease in the *Gata2* gene. The x axis displays the number of estimated nucleotide changes in form of deletions or insertions (indel) based on the decomposition algorithm of TIDE. The frequency of the given indel is displayed on the y-axis.

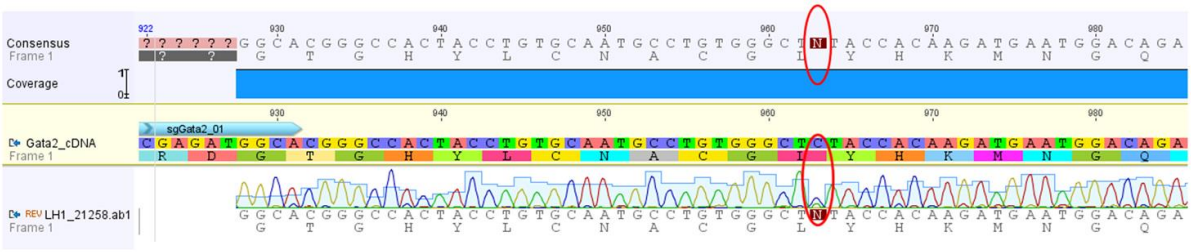


Figure 8: Site of the + 1 insertion in cDNA from a LL *Gata2*$^{mut.}$ clone

The first row depicts the consensus sequence of wild type *Gata2* cDNA. The second row depicts the cDNA from an exemplary LL *Gata2*$^{mut.}$ clone harbouring the 1 + insertion. The + 1 insertion of a cytosine residue is marked by two red circles. In the third row a chromatogram is depicted which consists of the sequencing data derived from the exemplary *LL Gata2*$^{mut.}$ clone.

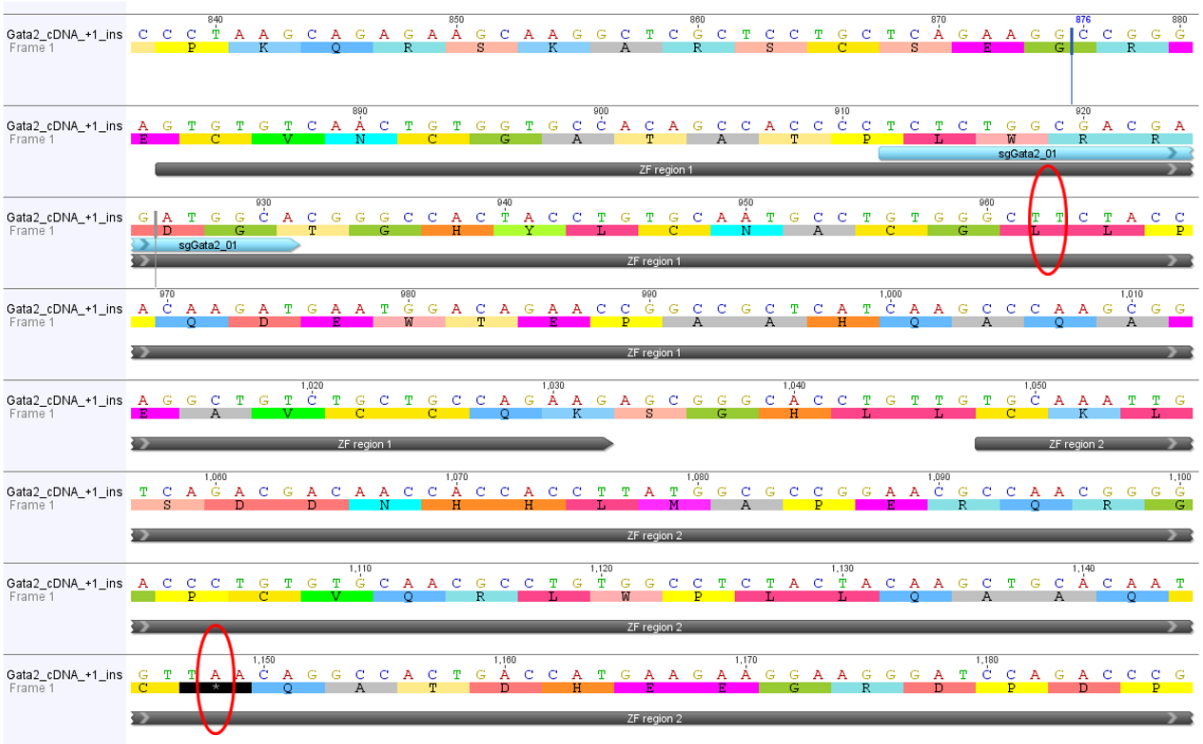
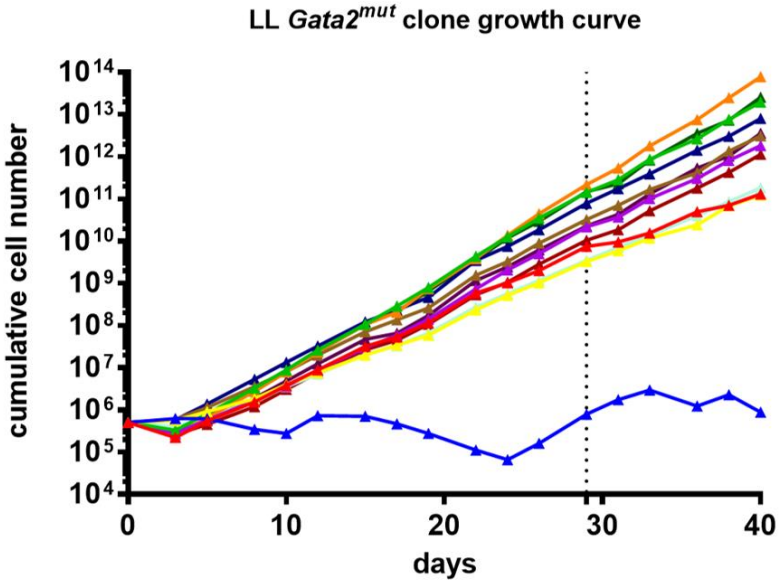


Figure 9: Site of the + 1 insertion leading to a preliminary stop codon

Depicted is the cDNA sequence derived from the *Gata2* mRNA as well as the amino acid encoded by each codon. The sequence of the zinc finger region one and two are marked by black bars and annotated with “ZF region 1” and “ZF region 2”. In addition, the targeting site of the sgRNA is marked by a blue bar and annotated “sgGata2_01”. The + 1 insertion site is marked by a red circle. The preliminary stop codon is marked by red circle and additionally indicated by a black starlet. Note that this is the sequence of wild type *Gata2* cDNA.

3.2 LL *Gata2*$^{mut.}$ clone growth curve

We performed a growth curve with eleven *Gata2* knock out clones and observed no significant difference in terms of growth rate when the cells were cultivated in IL-3 at a concentration of 5 µg/ml (Figure 10). On day 29 the cells were cultivated in IL-3 at a concentration of 1 µg/ml in order to determine if the lack of difference in growth rate of *Gata2*-mutated versus wildtype cells was dependent on high IL-3 levels that favour optimal growth. In fact, the growth rate appeared to increase in all the eleven LL *Gata2*$^{mut.}$ clones while remaining stable in the LL *Gata2*^{wt} cells upon cultivation in lower IL-3 concentrations. The LL *Gata2*$^{mut.}$ Clone 1 showed a very low and heterogeneous growth rate over the 40-day timespan.



		IL-3 (5 µg/ml)		IL-3 (1 µg/ml)	
Legend	Clone	growth rate	SD	growth rate	SD
	LL *Gata2* Clone 1	1.01	0.74	2.34	1.54
	LL *Gata2* Clone 2	1.93	0.71	2.82	0.53
	LL *Gata2* Clone 3	2.81	1.17	3.36	0.40
	LL *Gata2* Clone 4	2.38	1.10	3.06	0.43
	LL *Gata2* Clone 5	2.94	1.23	3.97	0.51
	LL *Gata2* Clone 6	2.01	0.84	2.63	0.31
	LL *Gata2* Clone 7	2.45	1.25	3.05	0.42
	LL *Gata2* Clone 8	2.69	1.66	3.04	0.52
	LL *Gata2* Clone 9	2.57	1.58	3.57	0.80
	LL *Gata2* Clone 10	2.27	1.08	3.21	0.52
	LL *Gata2* Clone 11	2.80	1.20	3.77	1.03
	LL *Gata2* wild type cells	2.26	1.18	2.23	0.67

Figure 10: LL *Gata2*$^{mut.}$ clone growth curve

The growth curve was obtained by measuring the cell number in triplicates over a time span of 40 days. Note that on day 29 the IL-3 concentration was reduced from 5 µg/ml to a concentration of 1 µg/ml as indicated by a dotted line. The table below the growth curve includes the legend as well as the mean growth rate of the respective LL *Gata2*^{mut} clone in 5 and 1 µlg/ml IL-3 and the standard deviation.

3.3 *Gata2* expression levels

While the growth curve revealed no proliferative advantage of the eleven *Gata2* mutated clones, the qPCR analysis revealed reduced *Gata2* expression in six clones. No changes in *Gata2* expression relative to LL Gata2^{wt} cells was observed in the LL *Gata2*$^{mut.}$ clone 2 and the LL *Gata2*$^{mut.}$ clone 7. Interestingly, LL *Gata2*$^{mut.}$ clone 4 and clone 11 showed higher *Gata2* expression levels than LL *Gata2*^{wt} cells. In order to test for correlation between growth rate and *Gata2* expression a non-parametric Spearman correlation test was performed. No correlation between *Gata2* expression levels and growth rate could be found (Figure 11-12).

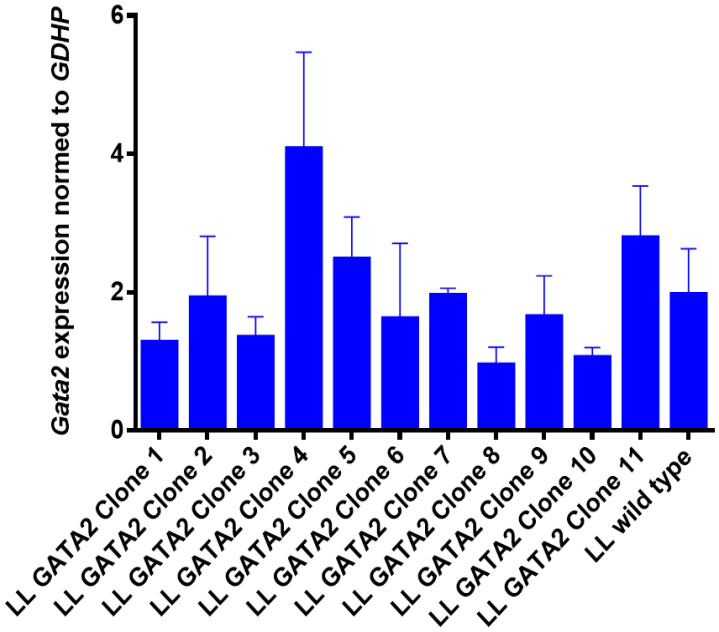
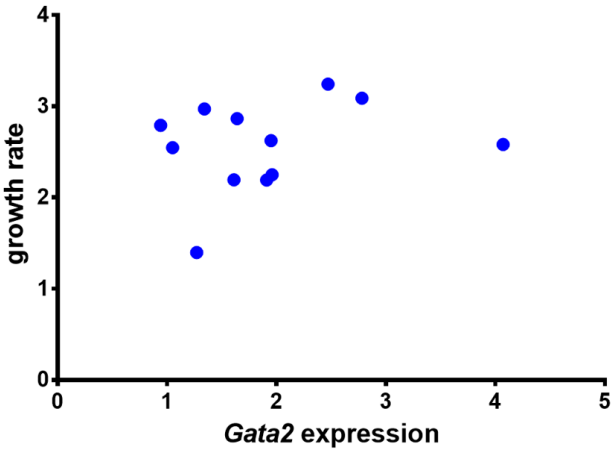


Figure 11: *Gata2* RNA levels of all LL *Gata2*$^{mut.}$ clones and Gata2 wild type LL cells

RNA samples were obtained at three time points during the 40-day timespan of the growth curve. qPCR analysis was performed on each set of RNA samples individually. Displayed in the graph above are the mean values of the *Gata2* mRNA levels of the data derived from all of the three qPCR analyses.



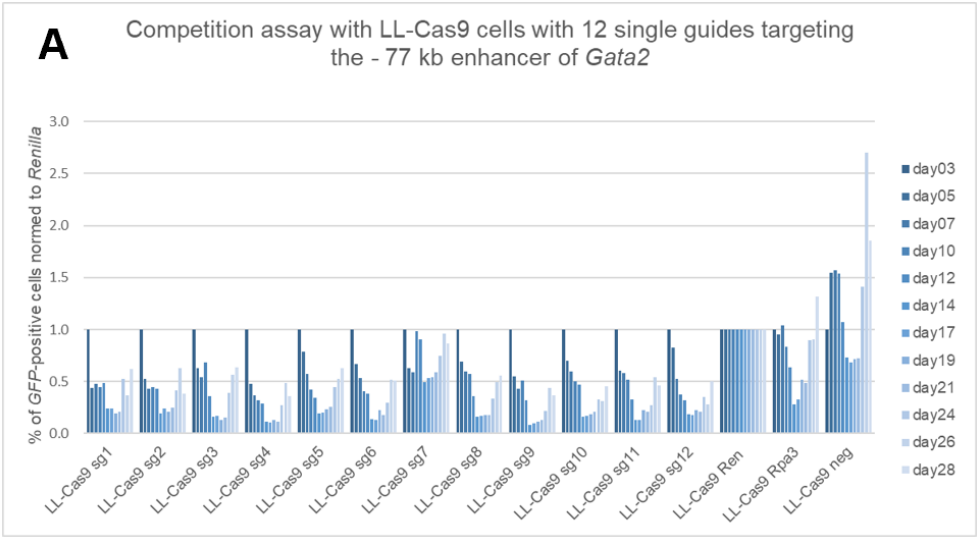
	Gata2 expression vs. growth rate
Spearman r	
r	0.2937
95% confidence interval	-0.3542 to 0.7511
P value	
P (two-tailed)	0.3545
P value summary	ns
Exact or approximate P value?	Exact
Significant? (alpha = 0.05)	No
Number of XY Pairs	12

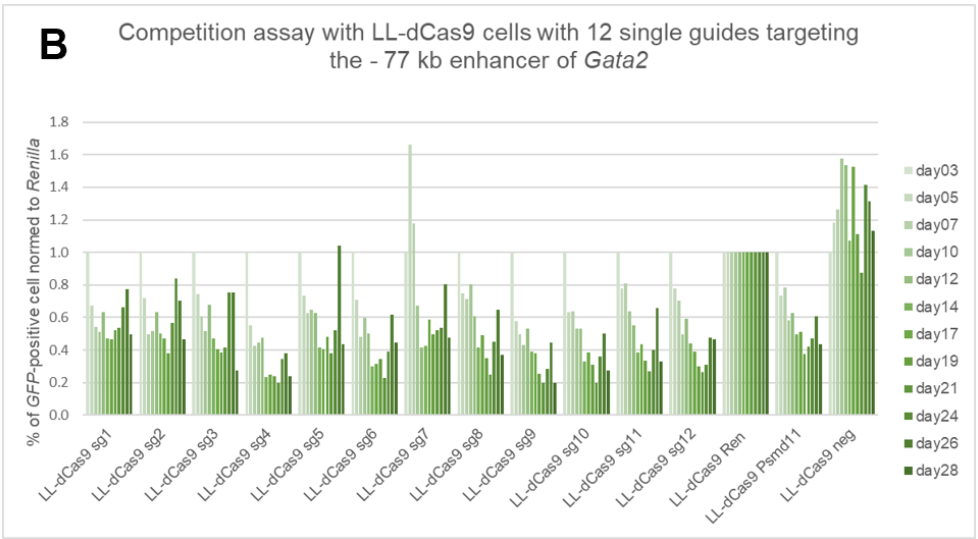
Figure 12: **Correlation between growth rate and *Gata2* expression**

The y-axis depicts the growth rate obtained from the growth curve shown in Figure 10 in both IL-3 concentrations (5 µg/ml and 1 µg/ml). The x-axis depicts the *Gata2* expression levels obtained from the qPCR analysis. A Spearman correlation test was performed by using the program GraphPad Prism. The results of the test are depicted in the table to the right of the correlation plot. Note that this was a nonparametric correlation test.

3.4 Competition assays in LL and KL cells targeting the - 77 kb enhancer of *Gata2*

We performed four competition assays in LL and KL cells, which stably express Cas9 or dCas9-KRAB. In total, we used twelve sgRNAs to target the – 77 kb enhancer of *Gata2*. The competition assays showed no proliferative outgrowth of infected cells as depicted in Figure 13 A-D. In contrary, the percentages of GFP-positive cells were reduced 5 days post-infection and stayed low until the end of the measurements in all the used cell lines. The positive control targeting the essential gene *Rpa3* in the LL-Cas9 cells and KL-Cas9 cells initially resulted in a proliferative disadvantage of the infected cells, which was lost on day 21.





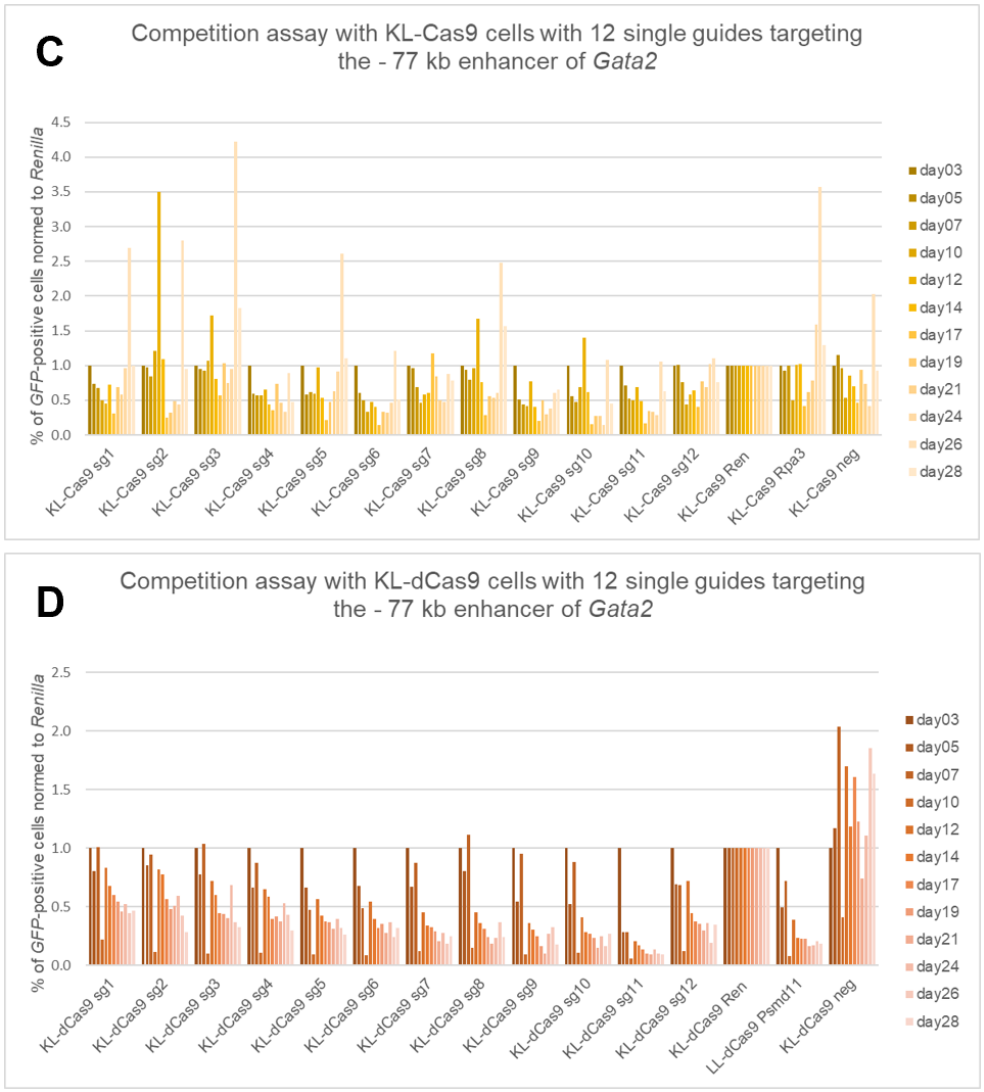
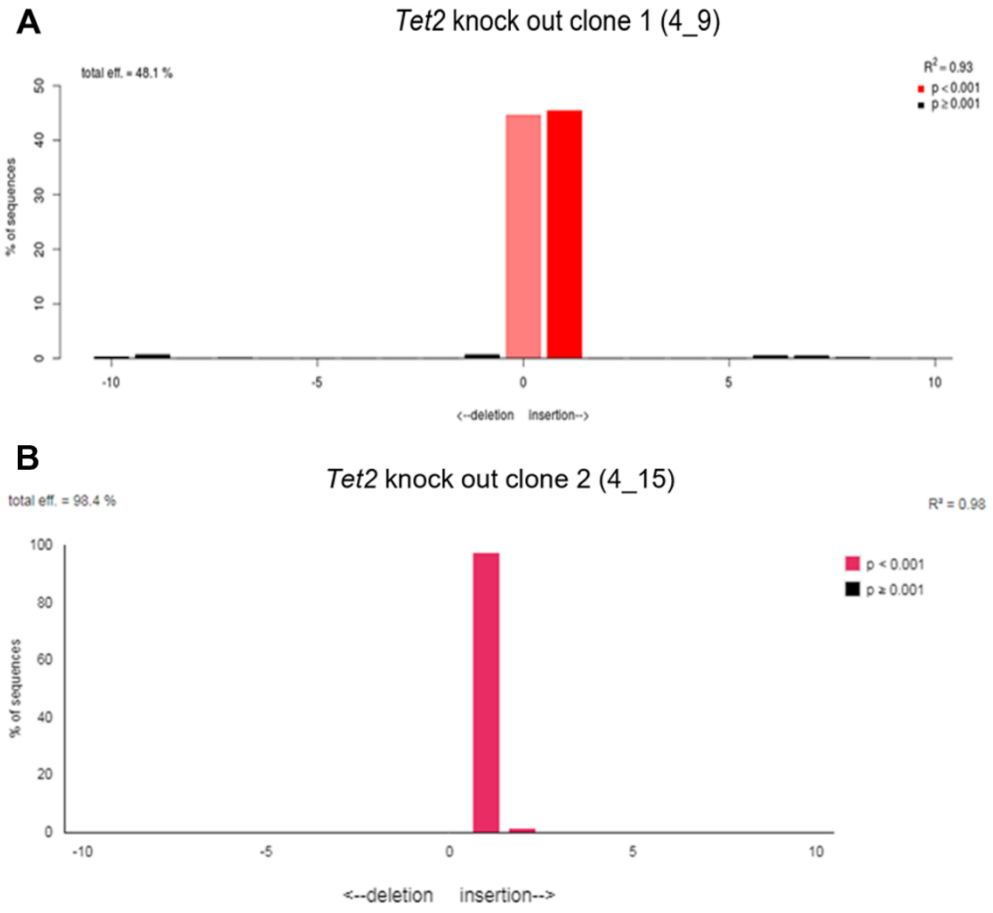


Figure 13: **Competition assays with LL/KL-Cas9/dCas9 cells with 12 single guides targeting the – 77 kb enhancer of *Gata2***

(A)-/D) Competition assay with LL-Cas9 (A), LL dCas9 (B), KL Cas9 (C) and KL dCas9 (D) cells
The cells were counted twelve times over a time span of 28 days. The first cell count was started on day 3 post-transduction. The sgRNAs are coupled to *GFP* and successfully transduced cells are *GFP*-positive and can be measured via flow cytometry. Note that only the percentage of living *GFP*-positive cells is displayed and normalized to the percentage of *GFP*-positive cells harbouring the sgRNA which targets the *Renilla* gene. Additionally, cells were transduced with a construct containing only the *GFP* gene and were labelled with LL/KL-Cas9/dCas9 “neg”.

3.5 *Tet2* knock-out validation

As previously mentioned, a *Tet2* knock-out validation was performed prior to this bachelor project. The sequencing results were analysed using the TIDE web tool (Figure 6) and show that all clones have a homozygous + 1 insertion in the exon 3 of the *Tet2* gene except for *Tet2*$^{mut.}$ clone 1 which had a heterozygous + 1 insertion.



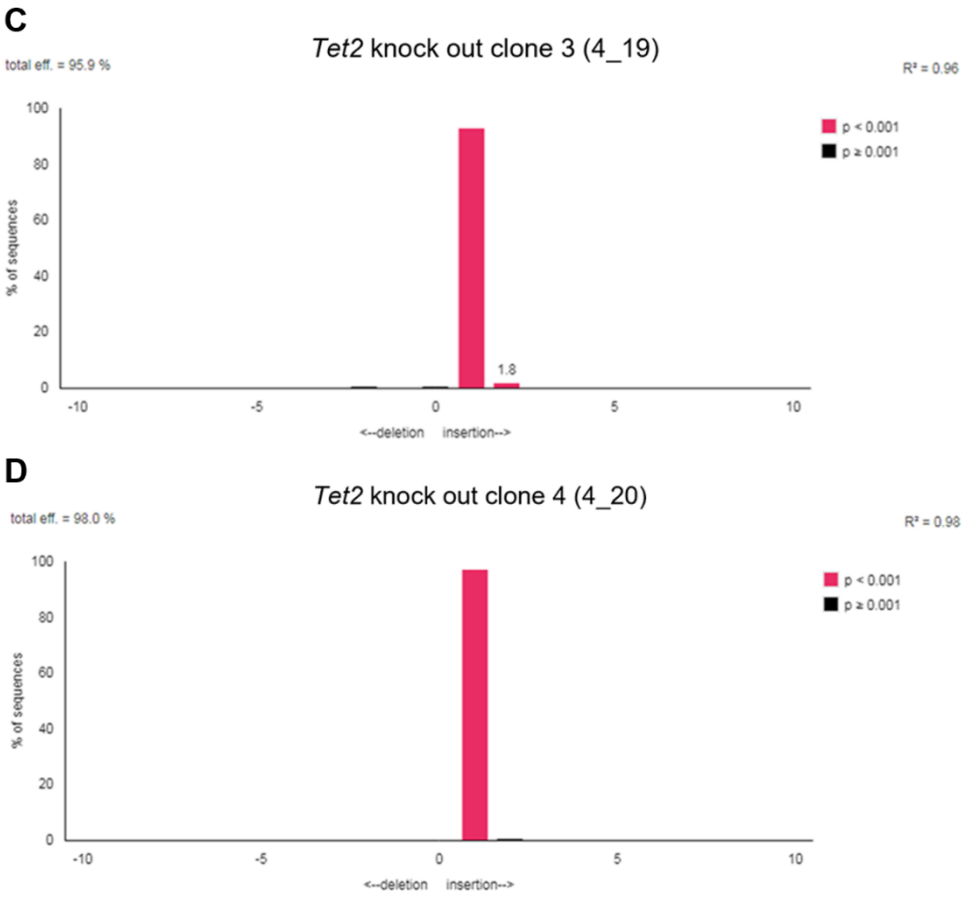


Figure 14: TIDE analysis of the four LL *Tet2*^{mut} knock out clones

(A)-(D) TIDE analysis of LL *Tet2*^{mut} clone 1-4

TIDE analysis was performed on sequencing data derived from PCR-amplicons of the expected cut site of the endonuclease. The x axis displays the number of estimated nucleotide changes in form of deletions or insertions based on the decomposition algorithm of TIDE. The frequency of the given indel is displayed on the y-axis. Note that the *Tet2* knock-out validation was performed in advance of this bachelor thesis.

3.6 p30-binding at the –77 kb enhancer of *Gata2*

We performed a Chromatin Immunoprecipitation Assay (ChIP) using an antibody targeting the C-terminus of C/EBPα to retrieve p30-bound DNA fragments. After conducting the ChIP, we performed a qPCR to quantify the enriched DNA that was bound by p30. The ChIP-qPCR revealed p30 binding at the -77 kb enhancer of *Gata2* as depicted in Figure 15. We observed no decrease of p30 binding in LL *Tet2*$^{mut.}$ clones compared to LL *Tet2* wildtype clones. Moreover, the LL *Tet2*$^{mut.}$ clones showed a trend of higher p30 binding as it can be seen by the amount of enriched DNA.

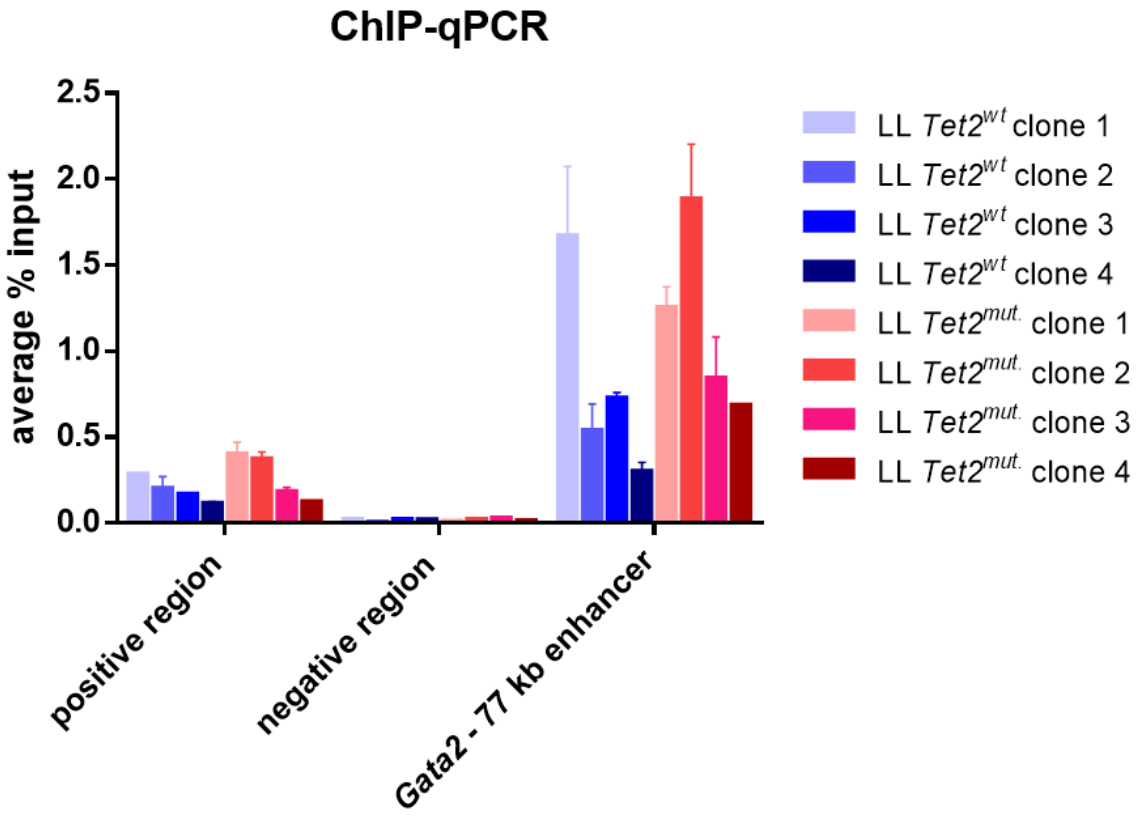


Figure 15: **ChIP-qPCR analysis of four LL *Tet2* wild-type clones and four LL *Tet2*^{mut} clones**

Depicted is the PCR-amplified p30-bound DNA as an average percentage of input DNA. In order to amplify and quantify the enriched DNA a total of three primer pairs were used. As a positive control, we used a region in which p30 is expected to bind based on previous data bind. The negative region is 1 kb upstream of the - 77 kb *Gata2* enhancer. qPCR analysis was performed on DNA obtained from the ChIP assay.

4. Discussion

In this project, we aimed to identify possible effector functions of *GATA2* in *CEBPA*-*TET2* co-mutated AML. We postulated that the reduction of *Gata2* expression to a certain level causes a proliferative advantage in the *CEBPA*-*TET2* mutated scenario. We performed a growth curve and a competition assay to determine if *Gata2* knockout leads to a proliferative advantage in *Cebpa*-mutated cells, as this would recapitulate a potential effect of the *Tet2* mutation. We also hypothesized that the reduction of *Gata2* expression is the result of reduced accessibility of p30 to the – 77 kb enhancer of *Gata2* upon *Tet2* knockout. This was investigated by ChIP-qPCR to test p30-binding in the – 77 kb enhancer in the presence or absence of Tet2.

We observed that the reduction of *Gata2* expression did not lead to a proliferative advantage in LL *Gata2*$^{mut.}$ clones over LL *Gata2*^{wt} cells (Figure 10). One possible explanation could be that the LL cells rely on the cytokine IL-3 for growth and the culture in presence of 5 ng/ml IL-3 might mask an effect of the *Gata2* mutation (Schmidt et al., 2019). Therefore, the IL-3 concentration was reduced to 1 ng/ml on day 29. In this way, we aimed to test if reduction of *Gata2* expression led to an advantage in suboptimal growth conditions. In fact, the growth rate increased in all LL *Gata2*$^{mut.}$ clones but not in the LL *Gata2*$^{wt.}$ cells, which had the lowest growth rate in culture with a lower IL-3 concentration. This is in line with the pro-proliferative role of IL-3 on myeloid cells. It could be that high concentrations of IL-3 activate diverse genetic programmes which increase AML cell proliferation in general and thereby masks the effect of a mild proliferative advantage caused by *Gata2* knockout. Although the changes in growth rates in the LL *Gata2*$^{mut.}$ clones compared to the LL *Gata2*^{wt} cells were not statistically significant, this finding supports the idea that the presence of IL-3 can obscure potential effects of *Gata2* mutation. Moreover, GATA2 is phosphorylated by a mitogen-activated protein kinase (MAPK) which is activated by IL-3 (Towatari et al., 1995). Although it is unclear how phosphorylation affects the activity of GATA2, it is possible that this mechanism explains a direct effect of IL-3 on the activity of GATA2. To strengthen our hypothesis that the *Gata2* mutation and expression reduction leads to a growth advantage, the growth curve should be performed again at lower IL-3 concentrations and over a longer time span.

We were not able to detect a correlation between growth rate and *Gata2* expression in the LL *Gata2*$^{mut.}$ clones. qPCR-analysis revealed that *Gata2* mRNA levels were indeed reduced in the LL *Gata2*^{mut} clones 1, 3, 6, 8 and 10 but were relatively unchanged in the LL *Gata2*^{mut}

clone 2 and 7 and even increased in LL *Gata2*^{mut} clone 4, 5 and 11 (Figure 11). Analysis of the *Gata2* sequence in the *Gata2*-mutated clones revealed that the + 1 insertion led to a premature stop codon. As the *Gata2*-mutation is heterozygous in most clones, *Gata2* can still be expressed at various levels from the remaining wild-type allele. This would explain the unchanged *Gata2* expression in three of the LL *Gata2* knock out clones.

Furthermore, *Gata2* expression can oscillate based on the cell cycle phase. A study revealed that *Gata2* expression is increased during S phase and reduced during G_1 and M phase (Koga et al., 2007). As samples might have been collected from cells in different cell cycle phases, this could explain increased expression in the LL *Gata2*^{mut} clone 2 and 7 compared to the LL *Gata2*^{wt} cells (Figure 11). Additionally, the CRISPR/Cas9-induced reduction of *Gata2* levels could be compensated by a negative feedback loop, which responds to reduced *Gata2* levels by increasing *Gata2* expression from the remaining non-mutated allele. But as we only measured *Gata2* mRNA expression we cannot make any quantitative or qualitative statements in terms of expression on the protein level. A western blot analysis could give further insights into changes in *GATA2* expression.

As previously described, we performed competition assays with LL and KL cells expressing either Cas9 or dCas9 using eleven sgRNAs to target the – 77 kb enhancer of *Gata2.* We found that cells transduced with sgRNAs targeting the *Gata2* enhancer had a proliferative disadvantage rather than an advantage in early stages of the experiment (Figure 14 A-D). We did not have enough cells to determine the level of *Gata2*-expression in the cells that were targeted. It could be that targeting of the – 77 kb enhancer of *Gata2* with CRISPR or CRISPRi leads to a complete knock out or knock down of *Gata2,* respectively*,* resulting in a proliferative disadvantage, as it was described in a previous study (Schmidt et al., 2019). Another aspect which has to be considered is that the expression of *Gata2* is also regulated through the + 9.5 kb enhancer, which is further downstream of *Gata2.* Activation of this enhancer might compensate for reduced expression in case of mutations in the – 77 kb enhancer (Mehta et al., 2017).

Another reason for the unexpected results could be technical issues and suboptimal experimental conditions. The LL and KL cell lines are difficult to infect with lentiviruses compared to other cell lines due to their susceptibility to infection-related cell stress. Therefore, it is essential to use the right amount of virus to successfully infect the cells without stressing them. The use of higher virus concentrations resulted in high numbers of

dead cells. Therefore, the competition assay had to be repeated with a lower virus concentration. As the transduction rate was relatively low (2-14 %) the CRISPR/CRISPRi-targeted subpopulations might be too small to outgrow the non-infected cells. Furthermore, targeting of the essential gene *Rpa3* resulted in unexpected fluctuations of GFP-positive cells, which might indicate aberrant Cas9 functionality or clonal outgrowth of escaper cells that express *GFP*, but are resistant to *Rpa3*-targeting. As previously mentioned, the LL cell line, as well as the KL cell line, have a relatively low growth rate. For example, in the process of establishing LL *Tet2*^{mut} clones it took 6 months until proliferative outgrowth of the LL T*et2*^{mut} clones was observed. This means that a proliferative outgrowth might only be detected if the competition assays are conducted over a longer time span.

It has been shown that *Tet2* is an important epigenetic regulator that demethylates diverse enhancers as well as promoters. Loss of function mutations in Tet2 lead to global DNA hypermethylation and disturbed gene expression (Nakajima and Kunimoto, 2014; Rasmussen et al., 2015). Moreover, the hypermethylation of DNA mainly leads to reduced TF-binding (Yin et al., 2017). We performed ChIP-qPCR experiment using an anti-C/EBPa antibody with four LL *Tet2*^{wt} and four LL *Tet2*$^{mut.}$ clones. The aim was to investigate p30-binding in the – 77 kb enhancer of *Gata2* as this locus showed reduced accessibility upon *Tet2* knockout. The results from the ChIP-qPCR indicate that there is no difference in p30-binding to the -77 kb enhancer of *Gata2* in the LL *Tet2*^{mut} clones compared to the LL *Tet2*^{wt} clones (Figure 15). Furthermore, the LL *Tet2*^{mut} clones 1, 2 and 3 exhibited even more p30-binding compared to the LL *Tet2*^{wt} clones 2, 3 and 4 (Figure 15). This is contrasting the ATAC-sequencing data performed on the LL *Tet2*^{mut} clones, in which reduced accessibility of the – 77 kb enhancer of *Gata2* was observed (Figure 6). However, the possibility remains that p30 binds to methylated DNA and then recruits TET2. In fact, it has previously been demonstrated that C/EBPα can interact with methylated DNA (Zhu, Wang, and Qian, 2016; Rishi et al., 2010). In this scenario, p30 would recruit TET2 to its target site via direct protein-protein interactions. This was recently shown in a study in which C/EBPα recruited TET2 to its target regions to demethylate diverse enhancers, thereby regulating enhancer activity (Sardina et al., 2018). Such a model would imply that p30 binds to the -77 kb enhancer of *Gata2* independently of its methylation state. This would explain the results we obtained in the ChIP-qPCR experiments. However, we do not have any information about the DNA methylation status of the -77 kb enhancer region in our LL Tet2^{wt} clones and LL Tet2^{mut}

clones to prove that the unchanged p30-binding is connected to unaltered DNA methylation. However, our collaboration partners performed Bisulphate sequencing in a related AML mouse model (*Cebpa*$^{p30/-}$ *Tet2*$^{-/-}$ versus *Cebpa*$^{p30/-}$ *Tet2*$^{+/+}$ cells) and observed no difference in the methylation state of the -77 kb enhancer of *Gata2* between *Cebpa*$^{p30/-}$ *Tet2*$^{-/-}$ versus *Cebpa*$^{p30/-}$ *Tet2*$^{+/+}$ cells.

It is possible that the LL *Tet2*^{wt} clones acquired an unknown mutation that leads to decreased reduced p30-binding in the -77 kb enhancer. To exclude this possibility, we performed ChIP-qPCR on the LL *Tet2*^{wt} pool from which the four LL *Tet2*^{wt} clones were derived. p30-binding to the enhancer of the parental LL cells was comparable to the LL *Tet2*^{wt} clones.

In summary, we did not find that reduced *Gata2* expression levels lead to a significant proliferative advantage of the cells used in this project. Furthermore, we found that p30 binding to the -77 kb enhancer of *Gata2* was not dependent on the presence of functional TET2 in our system. In conclusion, more experiments need to be performed to confirm preliminary data indicating that C/EBPα p30 and TET2 functionally cooperate. For example, ChIP-seq of TET2 in LL cells expressing shRen versus shCebpa will shed light on the hypothesis that knockdown of *Cebpa* affects the binding of TET2.

As depicted in Figure 5, a *CEBPA* and *TET2* co-mutation is also associated with reduced expression of three other factors: IKZF2, FUT8 and SIRT5. Interestingly IKZF2 is an important regulator of hematopoietic stem cell renewal and myeloid differentiation (Park et al. 2019). Therefore, it could be possible that p30 and TET2 interact through reduced expression of IKZF2. Nevertheless, *GATA2* remains a possible effector in the development of AML and even a potential drug target (Menendez-Gonzalez et al., 2019).

5. Summary

Acute myeloid leukaemia (AML) is a hematopoietic malignancy, which can arise from mutations in diverse differentiation-driving transcription factors leading to uncontrolled proliferation and aberrant differentiation of myeloid progenitors at the expense of terminally differentiated blood cells. One essential driver of myeloid differentiation is the CCAAT/enhancer-binding protein alpha (C/EBPα) transcription factor, which is mutated in 10-15 % of all AML patients. The *CEBPA* gene encodes a full-length p42 and a shorter p30 isoform by the use of two translation initiation codons in the *CEBPA* mRNA. Both mono- and biallelic *CEBPA* mutations (*CEBPA*$^{mo/bi}$) frequently lead to aberrant production of the p30 isoform, which has been shown to act as a gain-of-function variant. However, the exact molecular mechanisms leading to oncogenic transformation by CEBPA mutations are still unknown. Moreover, *CEBPA* is frequently co-mutated with *TET2* in 44.4 % of *CEBPA*^{mo} patients and in 34.8 % of *CEBPA*^{bi} patients. TET2 is an important epigenetic modifier that demethylates DNA and thereby can regulate the chromatin accessibility. RNA sequencing data from human AML patients as well as from murine cell lines and AML mouse models for *CEBPA*-*TET2* mutations showed that *GATA2*, a transcription factor that ensures the differentiation and production of granulocytes and macrophages, is downregulated upon *CEBPA*-*TET2* co-mutation. The aim of this study was to investigate how the level of *Gata2* expression affects proliferation of *Cebpa*-mutated cells and how mutated *Cebpa* cooperates with *Tet2* mutations to reduce *Gata2* expression. For this, we performed competition assays with eleven *Cebpa*-mutated clones that carry heterozygous *Gata2* mutations using CRISPR and CRISPRi to target the -77 kb enhancer of *Gata2.* We measured *Gata2* expression via qPCR and determined the effect of Gata2 expression on cell viability. In addition, we performed ChIP-qPCR on four *Tet2* knock-out clones and four wild type clones to investigate if reduced *Gata2* expression is the result of reduced binding of p30 to the *Gata2* -77 kb enhancer upon *Tet2* knockout. We found no conclusive correlation between *Gata2* expression levels and proliferation in our cell line models. As we found no difference in p30-binding in the *Gata2* gene region in co-mutated cells, we hypothesize that the two mutations cooperate via a different mechanism.

6.Zusammenfassung

Die akute myeloische Leukämie (AML) ist eine hämatopoetische Krebserkrankung, die durch Mutationen in verschiedenen differenzierungstreibenden Transkriptionsfaktoren entstehen kann. Bei AML kommt es zur unkontrollierten Proliferation und aberranter Differenzierung von myeloischen Vorläuferzellen. Dadurch wird die Produktion von ausdifferenzierten Blutzellen gestört. Ein essentieller Transkriptionsfaktor der myeloischen Differenzierung ist das CCAAT/Enhancer-Binding Protein Alpha (C/EBPα), das bei 10 bis 15 % aller AML-Patienten mutiert ist. Das *CEBPA*-Gen kodiert für eine vollständige p42-Isoform und eine kürzere p30-Isoform durch Verwendung von zwei internen Translations-Inititationscodons in der *CEBPA* mRNA. Sowohl mono- als auch biallelische *CEBPA*-Mutationen (*CEBPA*$^{mo/bi}$) führen häufig zu einer gestörten Produktion der p42-Isoform, wodurch nur mehr die p30-Isoform produziert wird, die sich als Gain-of-function-Variante erwiesen hat. Die genauen Mechanismen, durch welche mutiertes C/EBPα zur malignen Transformation führen, sind jedoch unbekannt. 44,4 % der *CEBPA*^{mo}-Patienten und 34,8 % der *CEBPA*^{bi}-Patienten weisen auch Mutationen im *TET2* Gen auf. TET2 ist ein wichtiger epigenetischer Regulator, der DNA demethyliert und dadurch die Zugänglichkeit des Chromatins regulieren kann. RNA-Sequenzierungsdaten von menschlichen AML-Patienten sowie von murinen Zelllinien und Mausmodellen für *CEBPA-TET2*-Mutationen zeigten, dass *GATA2*, ein essentieller Transkriptionsfaktor, der die Differenzierung und Produktion von Granulozyten und Makrophagen sicherstellt, bei einer *CEBPA-TET2*-Ko-Mutation herunterreguliert wird. Ziel der Studie war es zu untersuchen, wie der Grad der *Gata2*-Expression die Proliferation von *Cebpa*-mutierten Zellen beeinflusst und wie *Cebpa* mit *Tet2*-Mutationen kooperiert, um die *Gata2*-Expression zu reduzieren. Dazu führten wir Competition-Assays mit elf *Cebpa*-mutierten Klonen, die eine heterozygote *Gata2*-Mutationen aufwiesen, durch und verwendeten CRISPRn und CRISPRi, um Mutation im – 77 kb Enhancer von *Gata2* zu induzieren. Wir haben die *Gata2*-Expression mittels qPCR gemessen und ihren Effekt auf die Wachstumsrate bestimmt. Zusätzlich führten wir eine ChIP-qPCR an vier *Tet2*-Knockout-Klonen und vier Wildtyp-Klonen durch um zu untersuchen, ob die reduzierte *Gata2*-Expression das Ergebnis einer reduzierten Bindung von p30 an den – 77 kb Enhancer des *GATA2* Gens nach einem *TET2*-Knockout ist. Wir fanden keine Korrelation zwischen *Gata2*-Expressionslevels und Proliferation in unseren Zelllinienmodellen. Da wir keinen Unterschied in der p30-Bindung in der *Gata2*- Genregion in ko-mutierten Zellen beobachten konnten, vermuten wir, dass die beiden Mutationen über einen anderen Mechanismus kooperieren.

7.List of abbreviations

Abbreviation	Complete word
2-Og	2-oxoglutarate
5hmC	5-methylcytosine
5mC	5-hydroxymethylcytosine
AML	acute myeloid leukaemia
BR-LZ	basic region-leucine zipper
BSA	bis(trimethylsilyl)acetamide
BSL	Biosafety level
Cas9	CRISPR associated protein 9
C/EBP	CCAAT/enhancer binding proteins
CEBPA	CCAAT/enhancer binding protein alpha
CEBPA^{dm}	single-mutated CEBPA
CEBPA^{sm}	double-mutated CEBPA
ChIP	Chromatin Immunoprecipitation
ChIP-seq	ChIP-sequencing
CLPs	common lymphoid progenitors
CMPs	common myeloid progenitors
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRi	CRISPR interference
crRNA	crispr RNA
Cys	cysteine rich region
DC	dendritic cell
DMEM	Dulbecco's Modified Eagle Media
DMSO	dimethyl sulfoxide
DNMTs	DNA-methyltransferases
dNTP	nucleoside triphosphate
DSβH	β-helix fold domain
EGTA	ethylene glycol-bis (β-aminoethyl ether)-N, N, N′, N′-tetraacetic acid
EP	erythrocyte progenitor
FBS	foetal bovine serum

GAPDH	glyceraldehyde 3-phosphate dehydrogenase
G-CSF	granulocyte colony-stimulating factor
GFP	green fluorescent protein
GMP	granulocyte/macrophage progenitor
GP	granulocyte progenitors
gRNA	guide RNA
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSCs	hematopoietic stem cell(s)
IL-3	Interleukin-3
KL cells	*CEBPA*$^{c\text{-}mut./p30}$cells
KRAB	Krüppel associated box
LB	Lysogeny broth
LL cells	CEBPA$^{p30/p30}$cells
LOF	loss-of-function
LSCs	Leukemic stem cell(s)
LT-HSCs	Long-term hematopoietic pluripotent stem cells
MAPK	mitogen-activated protein kinase
MacP	macrophage progenitor
MEPs	megakaryocyte/macrophage progenitor
MkP	megakaryocyte progenitor
MLL1	mixed-lineage leukemia 1
MPPs	multipotent progenitors
NK	natural killer
NP40	nonyl phenoxypolyethoxylethanol
p30	30 kDa protein
p42	42 kDa protein
PBS	phosphate buffered saline
PEI	polyethyleneimine
Psmd11	Proteasome 26S Subunit, Non-ATPase 11
qRT-PCR (qPCR)	quantitative real time PCR

Renilla	Renilla luciferase
Rpa3	replication protein a 3
RPMI	Roswell Park Memorial Institute
SDS	sodium dodecyl sulphate
sgRNA	single guide RNA
shRNA	small hairpin RNA
SpCas9	Streptococcus pyogenes Cas9
ST-HSCs	short-term hematopoietic pluripotent stem cells
TADs	transcriptional activation domains
TAE	TRIS-Acetate-EDTA-buffer
TE	TRIS-EDTA
TET2	Ten-Eleven-Translocation 2 factor
TFs	transcription factor(s)
tracrRNA	transactivating-crispr RNA
TRIS	tris(hydroxymethyl)aminomethane
VSV-G	Vesicular Stomatitis Virus Glycoprotein
WHO	World Health Organization
WT-1	Wilms' tumor protein 1

8. Literature

Bereshchenko, Oxana, Elena Mancini, Susan Moore, Daniel Bilbao, Robert Månsson, Sidinh Luc, Amit Grover, Sten Eirik W. Jacobsen, David Bryder, and Claus Nerlov. 2009. 'Hematopoietic Stem Cell Expansion Precedes the Generation of Committed Myeloid Leukemia-Initiating Cells in C/EBPα Mutant AML'. *Cancer Cell* 16 (5): 390–400.

Bonifer, Constanze, Maarten Hoogenkamp, Hanna Krysinska, and Hiromi Tagoh. 2008. 'How Transcription Factors Program Chromatin—Lessons from Studies of the Regulation of Myeloid-Specific Genes'. *Seminars in Immunology* 20 (4): 257–63.

Bresnick, Emery H., and Kirby D. Johnson. 2019. 'Blood Disease–Causing and –Suppressing Transcriptional Enhancers: General Principles and GATA2 Mechanisms'. *Blood Advances* 3 (13): 2045–56.

Brinkman, Eva K., Tao Chen, Mario Amendola, and Bas van Steensel. 2014. 'Easy Quantitative Assessment of Genome Editing by Sequence Trace Decomposition'. *Nucleic Acids Research* 42 (22): e168–e168.

Brocken, Daan J.W., Mariliis Tark-Dame, and Remus T. Dame. 2018. 'DCas9: A Versatile Tool for Epigenome Editing'. *Current Issues in Molecular Biology*, 15–32.

Das, Partha M., Kavitha Ramachandran, Jane vanWert, and Rakesh Singal. 2004. 'Chromatin Immunoprecipitation Assay'. *BioTechniques* 37 (6): 961–69.

Di Genua, C. et al. (2020) 'C/EBPα and GATA-2 Mutations Induce Bilineage Acute Erythroid Leukemia through Transformation of a Neomorphic Neutrophil-Erythroid Progenitor', Cancer Cell, 37(5), pp. 690-704.e8.

Dzierzak, E. and Bigas, A. (2018) 'Blood Development: Hematopoietic Stem Cell Dependence and Independence', Cell Stem Cell, 22(5), pp. 639–651.

Fasan, A, C Haferlach, T Alpermann, S Jeromin, V Grossmann, C Eder, S Weissmann, et al. 2014. 'The Role of Different Genetic Subtypes of CEBPA Mutated AML'. *Leukemia* 28 (4): 794–803.

Feng, Yimei, Xiaoping Li, Kaniel Cassady, Zhongmin Zou, and Xi Zhang. 2019. 'TET2 Function in Hematopoietic Malignancies, Immune Regulation, and DNA Repair'. *Frontiers in Oncology* 9 (April): 210.

Gore, A. V. and Weinstein, B. M. (2016) 'DNA methylation in hematopoietic development and disease', *Experimental Hematology*, 44(9), pp. 783–790.

Greif, Philipp A., Annika Dufour, Nikola P. Konstandin, Bianka Ksienzyk, Evelyn Zellmeier, Belay Tizazu, Jutta Sturm, et al. 2012. 'GATA2 Zinc Finger 1 Mutations Associated with Biallelic CEBPA Mutations Define a Unique Genetic Entity of Acute Myeloid Leukemia'. *Blood* 120 (2): 395–403.

Groner, Anna C., Sylvain Meylan, Angela Ciuffi, Nadine Zangger, Giovanna Ambrosini, Nicolas Dénervaud, Philipp Bucher, and Didier Trono. 2010. 'KRAB–Zinc Finger Proteins and KAP1 Can Mediate Long-Range Transcriptional Repression through Heterochromatin Spreading'. Edited by Hiten D. Madhani. *PLoS Genetics* 6 (3): e1000869.

Hwang, Sang Mee. 2020. 'Classification of Acute Myeloid Leukemia'. *BLOOD RESEARCH* 55 (S1): S1–4.

Ito, Shinsuke, Ana C D'Alessio, Olena V Taranova, Kwonho Hong, and Yi Zhang. 2012. 'Role of Tet Proteins in 5mC to 5hmC Conversion, ES Cell Self- Renewal, and ICM Specification', 15.

Jan, M., T. M. Snyder, M. R. Corces-Zimmerman, P. Vyas, I. L. Weissman, S. R. Quake, and R. Majeti. 2012. 'Clonal Evolution of Preleukemic Hematopoietic Stem Cells Precedes Human Acute Myeloid Leukemia'. *Science Translational Medicine* 4 (149): 149ra118-149ra118.

Johnson, K. D., G. Kong, X. Gao, Y.-I. Chang, K. J. Hewitt, R. Sanalkumar, R. Prathibha, et al. 2015. 'Cis-Regulatory Mechanisms Governing Stem and Progenitor Cell Transitions'. *Science Advances* 1 (8): e1500503–e1500503.

Juliusson, Gunnar, Petar Antunovic, Åsa Derolf, Sören Lehmann, Lars Möllgård, Dick Stockelberg, Ulf Tidefelt, Anders Wahlin, and Martin Höglund. 2009. 'Age and Acute Myeloid Leukemia: Real World Data on Decision to Treat and Outcomes from the Swedish Acute Leukemia Registry'. *Blood* 113 (18): 4179–87.

Kassim, Adetola A., and Bipib N. Savani. 2017. 'Hematopoietic Stem Cell Transplantation for Acute Myeloid Leukemia: A Review'. *Hematology/Oncology and Stem Cell Therapy* 10 (4): 245–51.

Kato, Naoko, Jiro Kitaura, Noriko Doki, Yukiko Komeno, Naoko Watanabe-Okochi, Katsuhiro Togami, Fumio Nakahara, et al. 2011. 'Two Types of C/EBP␣ Mutations Play Distinct but Collaborative Roles in Leukemogenesis: Lessons from Clinical Data and BMT Models' 117 (1): 13.

Khwaja, Asim, Magnus Bjorkholm, Rosemary E. Gale, Ross L. Levine, Craig T. Jordan, Gerhard Ehninger, Clara D. Bloomfield, et al. 2016. 'Acute Myeloid Leukaemia'. *Nature Reviews Disease Primers* 2 (1): 16010.

Kirstetter, Peggy, Mikkel B. Schuster, Oksana Bereshchenko, Susan Moore, Heidi Dvinge, Elke Kurz, Kim Theilgaard-Mönch, et al. 2008. 'Modeling of C/EBPα Mutant Acute Myeloid Leukemia Reveals a Common Expression Signature of Committed Myeloid Leukemia-Initiating Cells'. *Cancer Cell* 13 (4): 299–310.

Koga, Shinichiro, Nobuhiro Yamaguchi, Tomoko Abe, Masayoshi Minegishi, Shigeru Tsuchiya, Masayuki Yamamoto, and Naoko Minegishi. 2007. 'Cell-Cycle–Dependent Oscillation of GATA2 Expression in Hematopoietic Cells' 109 (10): 9.

Kouzarides, Tony. 2007. 'Chromatin Modifications and Their Function'. *Cell* 128 (4): 693–705.

Labun, Kornel, Tessa G Montague, Maximilian Krause, Yamila N Torres Cleuren, Håkon Tjeldnes, and Eivind Valen. 2019. 'CHOPCHOP v3: Expanding the CRISPR Web Toolbox beyond Genome Editing'. *Nucleic Acids Research* 47 (W1): W171–74.

Lara-Astiaso, D., A. Weiner, E. Lorenzo-Vivas, I. Zaretsky, D. A. Jaitin, E. David, H. Keren-Shaul, et al. 2014. 'Chromatin State Dynamics during Blood Formation'. *Science* 345 (6199): 943–49.

Lekstrom-Himes, Julie, and Kleanthis G. Xanthopoulos. 1998. 'Biological Role of the CCAAT/Enhancer-Binding Protein Family of Transcription Factors'. *Journal of Biological Chemistry* 273 (44): 28545–48.

Leon, Smadar Ben-Tabou de-, and Eric H. Davidson. 2007. 'Gene Regulation: Gene Control Network in Development'. *Annual Review of Biophysics and Biomolecular Structure* 36 (1): 191–212.

Leubolt, Georg, Enric Redondo Monte, and Philipp A. Greif. 2020. '*GATA2* Mutations in Myeloid Malignancies: Two Zinc Fingers in Many Pies'. *IUBMB Life* 72 (1): 151–58.

Mehta, Charu, Kirby D. Johnson, Xin Gao, Irene M. Ong, Koichi R. Katsumura, Skye C. McIver, Erik A. Ranheim, and Emery H. Bresnick. 2017. 'Integrating Enhancer Mechanisms to Establish a Hierarchical Blood Development Program'. *Cell Reports* 20 (12): 2966–79.

Menendez-Gonzalez, Juan Bautista, Samantha Sinnadurai, Alex Gibbs, Leigh-anne Thomas, Maria Konstantinou, Alfonso Garcia-Valverde, Magali Boyer, et al. 2019. 'Inhibition of GATA2 Restrains Cell Proliferation and Enhances Apoptosis and Chemotherapy Mediated Apoptosis in Human GATA2 Overexpressing AML Cells'. *Scientific Reports* 9 (1): 12212.

Menendez-Gonzalez, Juan Bautista, Milica Vukovic, Ali Abdelfattah, Lubaid Saleh, Alhomidi Almotiri, Leigh-anne Thomas, Aloña Agirre-Lizaso, et al. 2019. 'Gata2 as a Crucial Regulator of Stem Cells in Adult Hematopoiesis and Acute Myeloid Leukemia'. *Stem Cell Reports* 13 (2): 291–306.

Mundade, Rasika, Hatice Gulcin Ozer, Han Wei, Lakshmi Prabhu, and Tao Lu. 2014. 'Role of ChIP-Seq in the Discovery of Transcription Factor Binding Sites, Differential Gene Regulation Mechanism, Epigenetic Marks and Beyond'. *Cell Cycle* 13 (18): 2847–52.

Nakajima, Hideaki, and Hiroyoshi Kunimoto. 2014. '*TET2* as an Epigenetic Master Regulator for Normal and Malignant Hematopoiesis'. *Cancer Science* 105 (9): 1093–99.

Orkin, Stuart H., and Leonard I. Zon. 2008. 'Hematopoiesis: An Evolving Paradigm for Stem Cell Biology'. *Cell* 132 (4): 631–44.

Ostergaard, Pia, Michael A Simpson, Fiona C Connell, Colin G Steward, Glen Brice, Wesley J Woollard, Dimitra Dafou, et al. 2011. 'Mutations in GATA2 Cause Primary Lymphedema Associated with a Predisposition to Acute Myeloid Leukemia (Emberger Syndrome)'. *Nature Genetics* 43 (10): 929–31.

Pabst, Thomas, Beatrice U. Mueller, Pu Zhang, Hanna S. Radomska, Sailaja Narravula, Susanne Schnittger, Gerhard Behre, Wolfgang Hiddemann, and Daniel G. Tenen. 2001. 'Dominant-Negative Mutations of CEBPA, Encoding CCAAT/Enhancer Binding Protein-α (C/EBPα), in Acute Myeloid Leukemia'. *Nature Genetics* 27 (3): 263–70.

Park, Sun-Mi, Hyunwoo Cho, Angela M. Thornton, Trevor S. Barlowe, Timothy Chou, Sagar Chhangawala, Lauren Fairchild, et al. 2019. 'IKZF2 Drives Leukemia Stem Cell Self-Renewal and Inhibits Myeloid Differentiation'. *Cell Stem Cell* 24 (1): 153-165.e7.

Quivoron, Cyril, Lucile Couronné, Véronique Della Valle, Cécile K. Lopez, Isabelle Plo, Orianne Wagner-Ballon, Marcio Do Cruzeiro, et al. 2011. 'TET2 Inactivation Results in Pleiotropic Hematopoietic Abnormalities in Mouse and Is a Recurrent Event during Human Lymphomagenesis'. *Cancer Cell* 20 (1): 25–38.

Rasmussen, Kasper D., Guangshuai Jia, Jens V. Johansen, Marianne T. Pedersen, Nicolas Rapin, Frederik O. Bagger, Bo T. Porse, Olivier A. Bernard, Jesper Christensen, and Kristian

Helin. 2015. ‘Loss of *TET2* in Hematopoietic Cells Leads to DNA Hypermethylation of Active Enhancers and Induction of Leukemogenesis’. *Genes & Development* 29 (9): 910–22.

Rishi, V., P. Bhattacharya, R. Chatterjee, J. Rozenberg, J. Zhao, K. Glass, P. Fitzgerald, and C. Vinson. 2010. ‘CpG Methylation of Half-CRE Sequences Creates C/EBP Binding Sites That Activate Some Tissue-Specific Genes’. *Proceedings of the National Academy of Sciences* 107 (47): 20311–16.

Rosenbauer, Frank, and Daniel G. Tenen. 2007. ‘Transcription Factors in Myeloid Development: Balancing Differentiation with Transformation’. *Nature Reviews Immunology* 7 (2): 105–17.

Sardina, Jose Luis, Samuel Collombet, Tian V. Tian, Antonio Gómez, Bruno Di Stefano, Clara Berenguer, Justin Brumbaugh, et al. 2018. ‘Transcription Factors Drive Tet2-Mediated Enhancer Demethylation to Reprogram Cell Fate’. *Cell Stem Cell* 23 (5): 727-741.e9.

Schmidt, Luisa, Elizabeth Heyes, and Florian Grebien. 2020. ‘Gain-of-Function Effects of N-Terminal *CEBPA* Mutations in Acute Myeloid Leukemia’. *BioEssays* 42 (2): 1900178.

Schmidt, Luisa, Elizabeth Heyes, Lisa Scheiblecker, Thomas Eder, Giacomo Volpe, Jon Frampton, Claus Nerlov, Peter Valent, Jolanta Grembecka, and Florian Grebien. 2019. ‘CEBPA-Mutated Leukemia Is Sensitive to Genetic and Pharmacological Targeting of the MLL1 Complex’. *Leukemia* 33 (7): 1608–19.

Seita, Jun, and Irving L. Weissman. 2010. ‘Hematopoietic Stem Cell: Self-Renewal versus Differentiation: Hematopoietic Stem Cell’. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 2 (6): 640–53.

Shah, Anjali, Therese M.-L. Andersson, Bernard Rachet, Magnus Björkholm, and Paul C. Lambert. 2013. ‘Survival and Cure of Acute Myeloid Leukaemia in England, 1971-2006: A Population-Based Study’. *British Journal of Haematology* 162 (4): 509–16.

Tallman, Martin S., D. Gary Gilliland, and Jacob M. Rowe. 2005. ‘Drug Therapy for Acute Myeloid Leukemia’. *Blood* 106 (4): 1154–63.

Towatari, Masayuki, Gillian E. May, Richard Marais, Gordon R. Perkins, Christopher J. Marshall, Sally Cowley, and Tariq Enver. 1995. ‘Regulation of GATA-2 Phosphorylation by Mitogen-Activated Protein Kinase and Interleukin-3’. *Journal of Biological Chemistry* 270 (8): 4101–7.

Wang, Jean C.Y., and John E. Dick. 2005. ‘Cancer Stem Cells: Lessons from Leukemia’. *Trends in Cell Biology* 15 (9): 494–501.

Wang, Yiping, Mengtao Xiao, Xiufei Chen, Leilei Chen, Yanping Xu, Lei Lv, Pu Wang, et al. 2015. ‘WT1 Recruits TET2 to Regulate Its Target Gene Expression and Suppress Leukemia Cell Proliferation’. *Molecular Cell* 57 (4): 662–73.

Weissmann, S, T Alpermann, V Grossmann, A Kowarsch, N Nadarajah, C Eder, F Dicker, et al. 2012. ‘Landscape of TET2 Mutations in Acute Myeloid Leukemia’. *Leukemia* 26 (5): 934–42.

Wouters, Bas J., Bob Löwenberg, Claudia A. J. Erpelinck-Verschueren, Wim L. J. van Putten, Peter J. M. Valk, and Ruud Delwel. 2009. ‘Double CEBPA Mutations, but Not Single CEBPA Mutations, Define a Subgroup of Acute Myeloid Leukemia with a Distinctive Gene Expression Profile That Is Uniquely Associated with a Favorable Outcome’. *Blood* 113 (13): 3088–91.

Yamashita, Masayuki, Paul V. Dellorusso, Oakley C. Olson, and Emmanuelle Passegué. 2020. ‘Dysregulated Haematopoietic Stem Cell Behaviour in Myeloid Leukaemogenesis’. *Nature Reviews Cancer* 20 (7): 365–82.

Yin, Yimeng, Ekaterina Morgunova, Arttu Jolma, Eevi Kaasinen, Biswajyoti Sahu, Syed Khund-Sayeed, Pratyush K. Das, et al. 2017. ‘Impact of Cytosine Methylation on DNA Binding Specificities of Human Transcription Factors’. *Science* 356 (6337): eaaj2239.

Zhang, D.-E., P. Zhang, N.-d. Wang, C. J. Hetherington, G. J. Darlington, and D. G. Tenen. 1997. 'Absence of Granulocyte Colony-Stimulating Factor Signaling and Neutrophil Development in CCAAT Enhancer Binding Protein -Deficient Mice'. *Proceedings of the National Academy of Sciences* 94 (2): 569–74.

Zhang, Pu, Junko Iwasaki-Arai, Hiromi Iwasaki, Maris L Fenyus, Tajhal Dayaram, Bronwyn M Owens, Hirokazu Shigematsu, et al. 2004. 'Enhancement of Hematopoietic Stem Cell Repopulating Capacity and Self-Renewal in the Absence of the Transcription Factor C/EBP α'. *Immunity* 21 (2): 853-863

Zhou, Vicky W., Alon Goren, and Bradley E. Bernstein. 2011. 'Charting Histone Modifications and the Functional Organization of Mammalian Genomes'. *Nature Reviews Genetics* 12 (1): 7–18.

Zhu, Heng, Guohua Wang, and Jiang Qian. 2016. 'Transcription Factors as Readers and Effectors of DNA Methylation'. *Nature Reviews Genetics* 17 (9): 551–65.

9. List of figures

10. List of tables