CEBPA IN NORMAL BLOOD CELL DEVELOPMENT AND IN MYELOID MALIGNANCIES

ROBERTO AVELLINO

THESIS_Roberto_Avellino.indd 1 07-05-18 10:17

Copyright © R.Avellino 2018

All rights reserved.
No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission from the author. The copyright of articles that have been published or accepted for publication has been transferred to the respective journals.

ISBN:
Layout: Egied Simons
Cover: Roberto Avellino
Printing: Optima (https://www.ogc.nl)

The work described in this thesis was performed at the Department of Hematology at the Erasmus Medical Center Cancer Institute, Erasmus University Medical Center, Rotterdam, the Netherlands.

Printing of this thesis was financially supported by KWF and Erasmus University Rotterdam.

THESIS_Roberto_Avellino.indd 2 07-05-18 10:17

CEBPA IN NORMAL BLOOD CELL DEVELOPMENT AND IN MYELOID MALIGNANCIES

CEBPA IN NORMALE BLOEDCELONTWIKKELING EN MYELOIDE MALIGNITEITEN

Proefschrift

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board.
The public defence shall be held on

Thursday 27th June 2018 at 9.30am

ROBERTO AVELLINO

Born in
Pieta', Malta

Erasmus University Rotterdam

THESIS_Roberto_Avellino.indd 3 07-05-18 10:17

DOCTORAL COMMITTEE

Promotors: Prof.dr. H.R Delwel
Prof.dr. I.P Touw

Other members: Prof.dr. H.G.M.P Raaijmakers
Prof.dr. C.P. Verrijzer
Prof.dr. J.J. Schuringa

THESIS_Roberto_Avellino.indd 4 07-05-18 10:17

THESIS_Roberto_Avellino.indd 5 07-05-18 10:17

THESIS_Roberto_Avellino.indd 6 07-05-18 10:17

CONTENTS

Chapter 1:	Introduction	9
Chapter 2:	Expression and regulation of C/EBPα in normal myelopoiesis and in malignant transformation	37
Chapter 3:	An Autonomous *CEBPA*-enhancer for myeloid-lineage priming and neutrophilic differentiation	59
Chapter 4:	*CEBPA*-enhancer targeting mutations are infrequent in acute myeloid leukemia	75
Chapter 5:	AML1-ETO reverses the active chromatin state and conformation of a myeloid specific enhancer to inhibit *CEBPA* expression	97
Chapter 6:	Neutropenia caused by loss of *CEBPA* expression reduces hemato-poietic stem cell numbers driven by a non-cell autonomous mechanism	119
Chapter 7:	Summary of the thesis and general discussion	171
Addendum:	English Summary	201
	Dutch Summary	203
	List of abbreviations	205
	Curriculum Vitae and list of publications	207
	PhD portfolio	209
	Word of thanks	211

THESIS_Roberto_Avellino.indd 7 07-05-18 10:17

1

THESIS_Roberto_Avellino.indd 8 07-05-18 10:17

INTRODUCTION

THESIS_Roberto_Avellino.indd 9 07-05-18 10:17

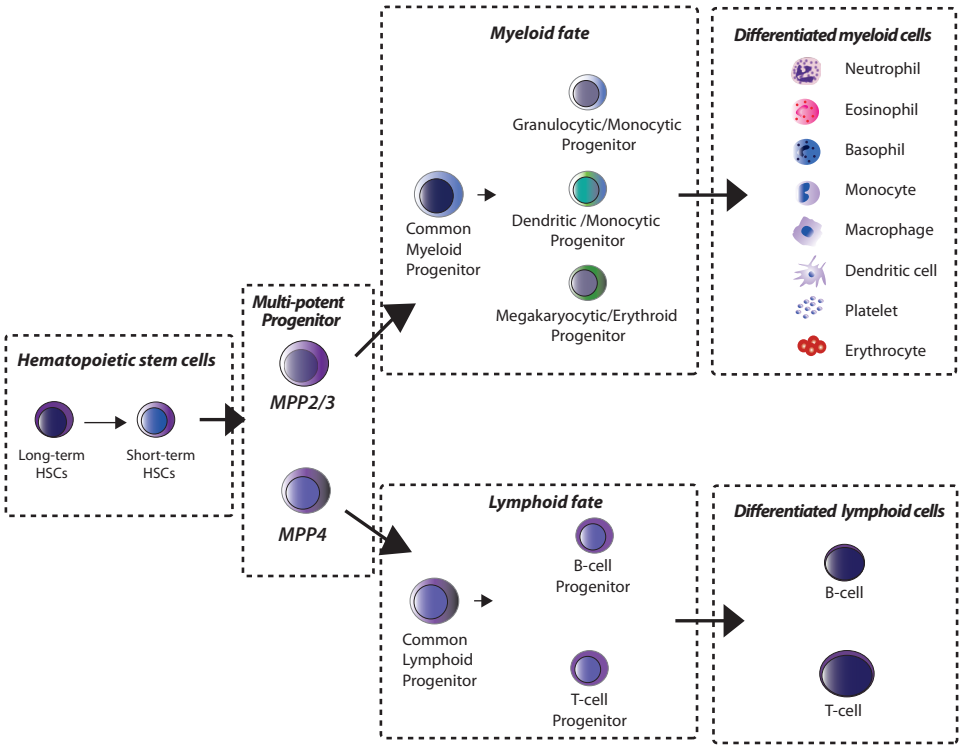


Figure 1. ***The hierarchy of hematopoietic development the bone marrow***. Long term hematopoietic stem cells (LT-HSCs) give rise to multipoint progenitors (MPP2/3 and MPP4) via a transient stage of short-term hematopoietic stem cells (ST-HSCs). MPP2/3 progenitors constitute the myeloid fate or common myeloid progenitors that give rise to the granulocytic/monocytic progenitor (GMP), dendritic/monocytic progenitor (DMP) and the megakaryocytic progenitor (MEP). These progenitors further mature to give rise to myeloid cells including the granulocytes (neutrophils, eosinophils, and basophils), mononuclear cells (monocytes, macrophages and dendritic cells), platelets and erythrocytes. The MPP4 progenitor is lymphoid primed and gives rise to the common lymphoid progenitor that constitutes B-cell and T-cell progenitors. These progenitors mature into naive B and T-cells that migrate to secondary lymphoid organs to complete mature stages via antigen encounter and engagement.

1.1 HEMATOPOIESIS

1.11 Introduction to hematopoiesis

The bone marrow is an active regenerative organ that supplies the adult human body on a daily basis with billions of blood cells to maintain the constant function of the immune system, the oxygen carrier system and the hemostatic system[1]. This ongoing process of blood cell formation is known as hematopoiesis, derived from the Greek word that means 'to make blood'. For long, hematopoiesis is believed to occur in hierarchy with the so-called long-term hematopoietic stem cells (LT-HSCs) sitting at the top of the apex. This infrequent LT-HSC population has high self-renewal capacity to maintain its rare pool (symmetric cell division) and further proliferate and differentiate to give rise to more committed and differentiated progenitors (asymmetric cell division) of the myeloid and lymphoid fate[2]. A representation of the hematopoietic hierarchy is shown in Figure 1.

The complex process of hematopoietic cell proliferation and differentiation involves extracellular communicatory processes between the hematopoietic stem/progenitor cells (HSPCs) and the bone marrow microenvironment[3]. The anatomical distribution of the bone marrow core and its environment allows this highly organized two-way/multi-way communication system to act in synchrony for maintaining the numbers and behavior of all bone marrow cells under a homeostatic and regenerative balance [4]. In response to extra-cellular regulatory processes, HSPCs undergo intracellular changes that prime lineage-specific epigenetic landscapes and activate gene expression programs in order to fully differentiate into functional blood cells [5, 6], which are later released into the systemic circulation to conduct their function at peripheral organs.

The priming and activation of gene expression programs during cell differentiation is fine-tuned by DNA binding transcription factors (TF) during early and adult life[7]. In collaboration with other nuclear proteins, these TFs form networks and together they bind sequence specific DNA elements, such as promoters and enhancers to modulate gene expression programs involved in differentiation. Altogether, these factors cooperate to shape the structure and function of the genome and promote cell type specific lineage commitment and differentiation[8]. Changes in the levels and activity of many of these factors, either induced experimentally or in disease, lead to aberrant bone marrow function and, ultimately, give rise to bone marrow failure syndromes and clonal hematological malignancies[9].

The following section gives a detailed description of hematopoiesis in early development in the fetus and in adult life. In addition it will be discussed how HSCs interact with the microenvironment to regulate their dynamics in quiescence, proliferation and differentiation.

1.12 The hematopoietic system: From the early fetus to adult hematopoiesis

The hematopoietic stem cells lie at the apex of the hematopoietic system and are considered to replenish the bone marrow with multipotent progenitors and subsequent lineage committed cells for the lifetime of an organism. These so-called long term hematopoietic stem cells are localized within specific bone marrow niches which modulate their kinetic activity upon physiological demand[10]. Blood cell production occurs through different stages of development and the major sources that are responsible for production and maintenance differ in fetal life from adult life.

In the early fetus, hematopoiesis occurs in two waves known as the primitive and definitive hematopoiesis. Primitive wave hematopoiesis occurs mainly in the yolk sac blood islands derived from a mesodermal cell type known as the hemangioblast that gives rise to hematopoietic and endothelial cells[11]. At day 7 (E7) of the murine gestation period, the majority of hematopoietic cells produced during the primitive wave of hematopoiesis are erythrocytes to supply oxygen for growing tissues and organs in the fetus[12, 13]. In addition, macrophages are also detected at this stage, most probably required for tissue resorption during organ development, which reach a quantitative abundance at E9.5 together with megakaryocytes [14, 15].

The second wave of blood cell production, known as the definitive wave, temporally overlaps with the first wave, and begins at E8.25 in the yolk sac. The majority of cells produced are c-Kit+ and CD41+ erythroid-myeloid progenitors, which are positive for the CD16/32 granulocytic marker at E9.5 and start also producing B-major adult hemoglobin erythrocytes, independent from the first wave derived erythrocytes[16-20]. These erythroid myeloid progenitors migrate and colonize the new forming liver, where predominant production of erythrocytes, granulocytes, monocytes, and macrophages occurs [16, 18, 21]. The lymphoid compartment is detected between E8.5/E9.5 in the second wave and is found in the yolk sac and aorta [22, 23].

The hematopoietic stem cells (HSCs) appear at E10.5, generated autonomously in different sites of the embryo. They first appear at the aorta-gonad-mesonephric (AGM) region and also in the placenta, the umbilical arteries and vitelline arteries. The cell derivatives of HSCs are the hemogenic endothelial cells that, under the regulation of Runx-1 and Gata-2, undergo an endothelial to hematopoietic cell transition or EHT. The fetal liver becomes the major niche for HSCs for a transient period of time before the permanent hematopoietic system is established in the bone marrow at birth [24-26]. After birth, HSCs proliferate autonomously for three weeks and then they enter in a dormant state unless an external insult challenges their quiescence into cell division [27].

Adult long-term hematopoiesis is majorly dependent on HSCs. By definition, long-term hematopoietic stem cells (LT-HSCs) are able to repopulate the bone marrow in transplantation settings, confirming their life-long property to sustain constant blood production [28]. LT-HSCs are known to reside in a quiescent niche where they divide rarely.

A proportion of these LT-HSCs in the mouse are estimated to divide every 30-50 days based on BrdU incorporation studies [29, 30]. A more dormant LT-HSC sub-population that constitutes 15% of the total HSC population is estimated to divide about five times in a mouse lifetime and do not contribute to daily blood cell production. However, this population is thought to participate in emergency cases such as inflammatory diseases, blood loss and tissue injury [30].

1.13 Hematopoietic stem cells and the bone marrow niche

The dynamics and kinetics of HSCs, including quiescence, self-renewal, proliferation, cell division and differentiation are modulated via interactions with the bone marrow microenvironment[3]. The bone marrow microenvironment or the bone marrow niche is composed of hematopoietic and non-hematopoietic cells. These non-hematopoietic cells make up the most significant compartment of the bone marrow stroma, subdivided into different niches but yet involved in one communication system to regulate HSC kinetics and maintenance [3, 31]. The bone marrow is highly vascularized with major arteries protruding into the venous sinusoids found close to the bone surface (endosteum). The endosteum is composed of osteoblast and osteoclasts. The majority of HSCs resides at perivascular niches and localize preferably to endosteal and sinusoidal regions [32, 33]. At these regions, HSCs are at close proximity of non-hematopoietic cells such as, bone-lining osteoblasts, endothelial cells of arterial and venal membranes, pericytes, mesenchymal stromal cells, also known as CXC-chemokine ligand 12 (CXCL12) abundant reticular cells (CAR-cells) and endothelial cells which constitute the biggest supply and production of stem cell factor (SCF), angiopoietin, transforming growth factor B (TGF-B) and CXCL12 to maintain HSCs under a quiescent physiological state[34-39]. In addition to non-hematopoietic cells, megakaryocytes also provide an important supply of TGF-B and CXC-chemokine ligand 4 to help maintaining HSC quiescence, whereas in emergency situations like radiation and myelo-ablation, megakaryocytes induce HSC expansion to increase HSC capacity and repopulate the bone marrow via the production and secretion of fibroblast growth factor 1 (FGF1)[40, 41]. Such mechanisms protect HSCs and keep them in a quiescent state to prevent HSC proliferation and HSC exhaustion that leads to bone marrow failure. **Figure 2** represents the bone marrow microenvironment and offers a schematic presentation of different niches.

1.14 Down the hematopoietic hierarchy: differentiation of LT-HSCs into specific lineages

Mechanisms how HSCs differentiate to produce their progeny is mainly based on in vitro colony assays and HSC transplantation in recipient mice. Specific growth factors, also called cytokines are responsible for driving the different progenitor cells into specific lineages by activating their corresponding membrane receptor [27]. For example, erythrocytes are generated under the stimulation of erythropoietin (Epo). Colony stimulating factor-3 (CSF-3),

G-CSF specifically stimulates neutrophil development, whereas thrombopoietin is a critical factor for platelet production[20].

LT-HSCs lie at the apex of the hematopoietic hierarchy and they give rise to progenitors with multi-lineage potential known asmulti-potent progenitors or MPPs[4]. The multipotent prognentor fraction is characterized by lineage makers (Lin-ve, Sca-1 pos, cKIT pos) CD48+CD150-, CD48+150+ and CD150+CD34+. Using functional assays, these MPPs were further subdivided based on their lineage potential; MPP2, MPP3 and MPP4 (Figure 1). These MPPs are driven by lineage-specific transcription factors to undergo myeloid or lymphoid differentiation. Lineage-specific TFs are under the control of a heptad of TFs that form a protein complex and bind genomic regions to activate the expression of lineage specific transcription factors[26, 42, 43]. Downstream of these MPPs, as defined by the hematopoietic hierarchy, are found myeloid and lymphoid fate progenitors that give rise to differentiated mature cells[44-46]. This thesis mainly focuses on myelopoiesis with a special interest in neutrophilic differentiation, thus introducing lymphopoiesis is beyond the scope of this thesis. A detailed description of the myeloid hierarchy is given based on previous findings reported in the last two decades, followed by new findings that are challenging previous conclusions and shaping in to new ideologies in adult myelopoiesis.

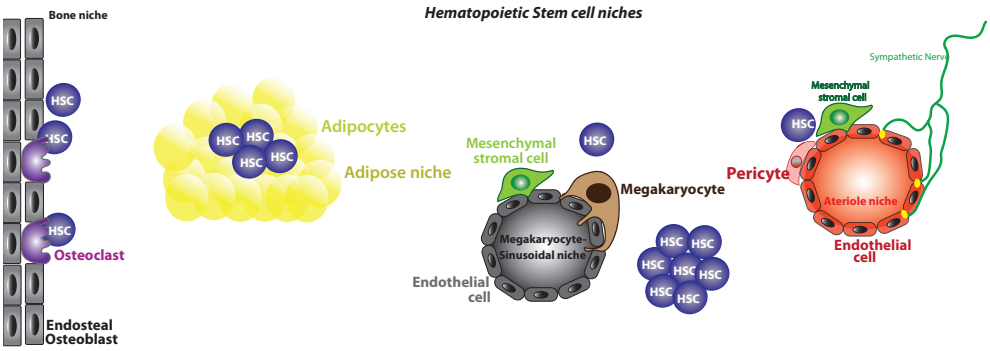


Figure 2. ***The hematopoietic stem cell (HSC) bone marrow nice***. The HSC niche is highly heterogeneous and composed of different types of stromal cells that modulate the kinetics of HSCs. Mostly identified niches include the bone niche, the sinusoidal and the megakaryocytic niche and the arteriole niche. These niches communicate with HSC by the secretion of factors that maintain HSCs under quiescence or stimulate proliferation in replicative stress, tissue injury, blood loss or severe infections.

1.15 Myelopoiesis: Myeloid progenitor cell populations and differentiation

The term myelopoiesis refers to the development of bone marrow cells towards the myeloid lineages. The myeloid compartment comprises a highly heterogeneous cell progenitor population that gives rise to mature neutrophils, eosinophils, basophils, mast cells, monocytes, macrophages, dendritic cells and megakaryocytes (Fig.1). The granulocytes,

monocytes, macrophages and the dendritic cells are the first line of innate immune cells that act as sentinels in the blood stream and in solid organs to protect against non-self-antigens and pathogens. Basophils, mast cells and eosinophils are less abundant in the circulatory system but are increased upon allergies or parasitic infections.

Irving Weismann and colleagues were the first to report the hierarchical view of myeloid progenitors based on clonogenic and transplantation assays[45]. The myeloid cell progenitors in murine bone marrows are characterized by cell surface markers and are subdivided into three Lineage^{-ve} cKIT^{+ve} Sca-1^{-ve} (LK) subsets: namely the CD16/32^{Low} CD34^{+ve} common myeloid progenitors (CMPs) which give rise to CD16/32^{High} CD34^{Low} granulyocytic/monocytic progenitors (GMPs) and CD16/32^{-ve} CD34^{-ve} megakaryocytic/erythroid progenitors (MEPs) [45]. They defined the CMP as the earliest myeloid progenitor since it has the potential to give rise to all myeloid cells[45]. The MEP population is more restricted to megakaryocytes and erythrocytes, whereas the GMP is more heterogeneous and mainly diverges to granulocytic and monocytic lineages [47, 48]. Dendritic cells are more of an independent myeloid cell entity, although dendritic myeloid progenitors also overlap with monocytic differentiation [49-53]. This constructed hierarchy is based on cell surface markers that define highly heterogeneous bone marrow progenitor cell populations. The impact of these marker-defined subsets on myelopoiesis *in vivo* will be discussed in the next sections.

1.16 Transcription factors in myelopoiesis: the concept of synergism and cross-antagonism in cell lineage fate decisions

Bone marrow transcription factors belong to a wide spectrum of transcription factor families that interact with accessible DNA sequences in regulatory elements to control gene expression. HSC and MPPs (or HSPCs) share a common TF complex composed of a so-called heptad of TFs including RUNX1, GATA-2, ERG, TAL1, LMO2, LYL1, FLI1[26, 42, 43]. This heptad of TF complexes binds to loci of genes encoding transcription factors essential for myeloid development, such as C/EBPα, PU.1, IRF8, GATA1 and GATA2, which drive lineage specific differentiation.

Many cell-lineage specific TFs are expressed at low levels in bone marrow myeloid progenitors. TFs compete or act synergistically in so-called transcription factor complexes to drive differentiation of any specific lineage [54]. In GMPs, the major TFs responsible for neutrophilic, monocytic and dendritic cell differentiation are C/EBPα, PU.1 and IRF8, respectively. These three represent classical examples where synergistic regulation takes place between PU.1 and IRF8 to support monocytic and/or dendritic cell differentiation[55], whereas C/EBPα antagonizes both PU.1 and IRF8 (and vice versa) to induce neutrophil differentiation[56-59]. In downstream progenitors of the GMPs, C/EBPα also activates the basophil gene expression program and antagonizes mast cell differentiation in the basophil/mast-cell bi-potent progenitor [60]. Moreover, lineage fate decisions between eosinophilic and basophilic progenitors are driven by cross-antagonism between C/EBPα and GATA2 [60,

61]. In addition, the GATA transcription factors GATA1 and GATA2 counteract each other in MEPs to differentiate into erythroid and megakaryocytic lineages, respectively [62, 63]. Based on ectopic expression of any of these factors in bone marrow progenitors or in cell line models, it is concluded that lineage fate determination is based on TF dosage that need to exceed a certain expression threshold to favor one lineage over the other[64-66]. However, this concept has largely been questioned and can possibly be replaced by recent findings using single cell technologies showing that a single TF might be responsible to drive differentiation of one cell lineage. In fact, single-cell technologies are revolutionizing and contradicting previous reports stating that cell fate decisions are based on syngergistic and antagonistic actions of transcription factors.

1.17 Challenging hierarchical vs unilineage hematopoiesis

In the last two decades, efforts has been made to understand the hematopoietic hierarchy based on the instructive force of lineage specific transcription factors in differentiation [67]. As stated in the previous section, TFs act together in synergy to bind enhancers and promoters and consequently activate genes required for lineage specific differentiation. However, the biology and dynamics of transcription factor complexes was majorly studied in immortalized cell line models, which, although they overcome limitations associated with cell numbers, may not reflect the function of TFs during differentiation *in vivo*. The advent of new technologies to study lineage tracing combined with single-cell RNA technology has challenged the concept of hierarchical hematopoiesis and the concept stating that TFs act in synergy or compete with each other to drive cell differentiation in the bone marrow [68, 69].

Dissecting the heterogeneous pool of myeloid progenitors using single-cell transcriptomics and lineage tracing, Ido Amit and colleagues showed that immune-phenotypic CMPs, GMPs and MEPs can be clustered into 19 different sub-populations. Each sub-population is driven by a lineage specific TF with no evidence that cell lineages exhibit transcription factors that compete for lineage differentiation as was previously thought [70]. This study opened up new frontiers in understanding how lineage specification takes place. It is now recognized that cell lineage commitment within the myeloid fate occurs at earlier stages than previously recognized [71-73], and lineage specific transcription factors are responsible to drive the differentiation program along with downstream TFs to support terminal differentiation [70]. The single cell approach used in the study by Paul et al [70], revolutionized the concept of myeloid hierarchy and shifted the line of thought towards unilineage myelopoiesis (See Figure 3).

The next section describes in detail how C/EBPα, the transcription factor under investigation in this thesis, plays a major role in neutrophilic differentiation and how its function is required to activate the neutrophilic gene expression program in the bone marrow.

1.18 C/EBPα in myelopoiesis and differentiation in the bone marrow

The CEBP leucine zipper family consists of six DNA binding proteins that share a highly conserved bZIP DNA-binding domain at the C-terminus. The family members are named alphabetically in C/EBPα, C/EBPβ, C/EBPδ, C/EBPε, C/EBPγ and C/EBPζ, [74-76]. Most of the CEBP proteins share a transactivation domain at the N-terminal site and its main function is to recruit general transcription factors to modulate gene expression of downstream target genes. C/EBPα was the first transcription factor of the CEBP family to be identified. The other CEBPs, including C/EBPβ, C/EBPδ, C/EBPε and C/EBPγ have all been implicated in myelopoiesis as well [77, 78]. Here I will focus on the role of C/EBPα in myelopoiesis, which is discussed in more detail in Chapter 2.

The function of *CEBPA* has been studied in different tissues including the liver, adipose, lung and the bone marrow[79]. Its function in the bone marrow is clearly associated myelopoiesis and with defects in this lineage upon germ-line knockout of the *Cebpa* gene in a murine model[80]. Using flow-cytometric analysis, it was found that the granulocytic program was halted and differentiation was blocked between the CMP to GMP stage, thus concluding that C/EBPα is indispensable for neutrophilic differentiation at a stage beyond the generation of CMPs [80-83].

Ido Amit and colleagues challenged these previous findings and concepts[70]. They investigated the transcriptional continuum of the neutrophilic differentiation program, from initiation to termination, by using C/EBPα (initiation) and C/EBPε (termination) knockout mouse models. Single cell transcriptomic analysis of bone marrow cells from C/EBPα knockout mice demonstrated the absence of neutrophil lineage priming activation and a complete abrogation of the neutrophilic cluster. In C/EBPε knockout bone marrow, a block of the neutrophilic differentiation was observed at a pre-terminally differentiated stage, which did not affect the priming and initiation of the neutrophilic program [70]. This suggests that C/EBPα initiates neutrophilic differentiation at earlier stages than the CMP, in contrast to what has been concluded by Tenen and colleagues[80, 84] and the accumulation of CMP progenitors observed are most probably myeloid fate progenitors that are primed to differentiate into myeloid cells other than neutrophils.

C/EBPα activates genes that make up the neutrophil gene expression program. C/EBPα binds to many genes that encode important functional components for neutrophils such as GCSF receptor; growth factors and interleukins including, GCSF, IL-1b, IL-6, IL-8, IL-12; intracellular enzymes required to fight pathogens such as lactoferrin, lysozyme, COX-2, MIP-1a, MIP-1b, myeloperoxidase, and neutrophil elastase[85-91].

Like any other transcription factor, C/EBPα utilizes mechanisms to regulate the expression of genes such as accessibility of chromatin[92], recruitment of ATP-dependent chromatin remodelers[93, 94] and changes in chromatin topology and chromatin states with regards to histone modifications[95, 96].

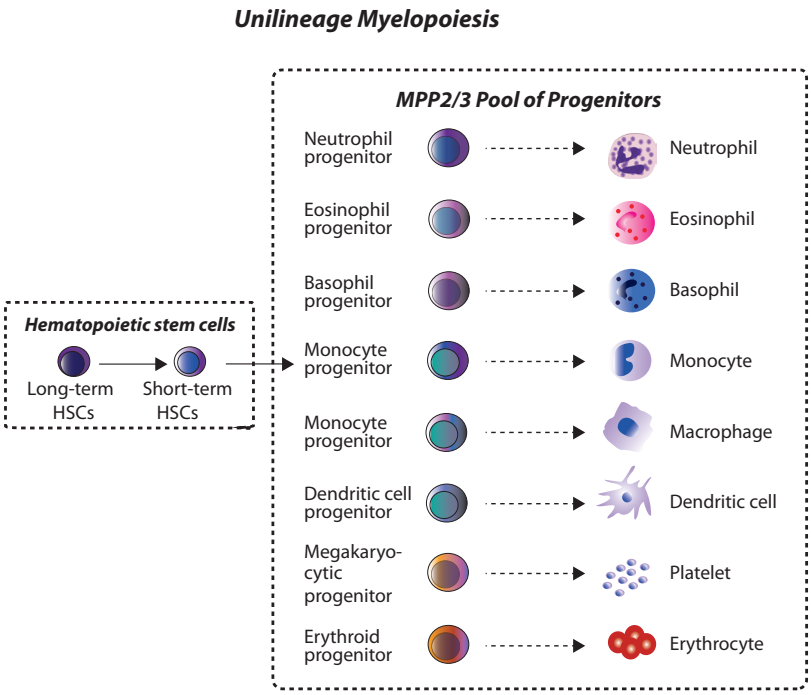


Figure 3. ***Emerging concept of uni-lineage myelopoiesis***. In this model adult bone marrow establishes myeloid biased cell lineages, where every single single cell is derived from a specific pre-programmed progenitor cell.

1.2 CHROMATIN AND TRANSCRIPTIONAL REGULATION

1.21 Introduction

Extracellular cues drive signaling pathways inside cells and orchestrate nuclear factors to modulate chromatin structure and function [97-99]. Nuclear factors fall into a wide spectrum of functional quaternary structural proteins [7, 100]. In the following sections I will focus on (1) the chromatin architecture and how it relates to function in terms of gene expression regulation, (2) how nuclear factors, which are sub-classified into chromatin factors and transcription factors, orchestrate the genome structure to drive differentiation in hematopoiesis and (3) how mutations in any of these factors perturb their function with consequences for phenotypic changes and the onset of diseases.

1.22 The fundamental components of chromatin

The histone octamer is composed of two copies of H2A, H2B, H3 and H4 histone proteins, which are wrapped around 147bps of DNA. Heterodimerization of histones is first initiated between H3 and H4, followed by H2A and H2B heterodimers causing strong interactions between the octamers that are further stabilized by the H1 linker histone to form a bigger histone-DNA complex structure called the nucleosome [101](See Figure 4a).

Change in nucleosomal structure is under the influence of post-translational modifications occurring at specific amino acids[102], either within the core (histone globule)[103] or at the positively charged N-terminal tails that protrude from the outside of the histone octamers [104](See Figure 4b). These modifications include acetylation, methylation, phosphorylation, ubiquitination and sumoylation. In particular, the N-terminal tails of histone 3 and histone 4 are highly susceptible for histone modifications [105-109]. Previous studies investigated profoundly the possibility of lysine residues to become modified at the N-terminal tail of histone 3 which include Lysine 4 (H3K4), H3K9, H3K27 and in core or globular domain histone 4 including H4K5, H4K8, H4K12, H4K16 [110]. Lysines are modified by the addition of small chemical moieties such as acetyl residues from acetyl CoA and catalyzed by active histone acetyl-transferases (HATs), known as writers. This process is called lysine acylation. HATs fall into a large family, which based on cellular localization are classified as type A and type B. The type A HATs are located in the nucleus and they play a major role in different DNA based mechanisms such as transcription, DNA replication and DNA repair. The type A HATs are subdivided into five sub-classes based to their homology. For instance, the MYST family members include Tip60, HBO1, HMOF, MOZ and MORF; the GNAT (GCN5-related N-acetyltransferases) subgroup that includes PCAF, GNC5 and ELP3; and the CBP/p300 family [111-116]. The type B HATs are localized in the cytoplasm and their major role is to modify nascent histones [117, 118]. The large family of histone deacetylases (HDACs) acts as erasers of acetylated lysine residues from modified histones.

In addition, lysines are also susceptible for methylation by histone methyl-transferases. Like HATs, lysine histone methyltransferases (KMTs) are subdivided into two major subgroups, but based on the presence or absence of a so-called a SET domain. KMTs with the SET domain include MLL, EZH2, NSD1 and G9a. These KMTs methylate a wide-array of lysine residues on histone H3 including K4, K9, K27 and K36, and on H4K20. Like acetylation, histone methylation may be erased by demethylases, including LSD1 and UTX [118, 119].

These modifications eventually change the nucleosomal dynamics and function i.e. it may result in either open chromatin or closed chromatin. Open chromatin allows for nuclear proteins such as ATP-dependent chromatin remodelers and transcription factors to bind and interact with DNA[120]. This accessibility and histone modifications define genes and their regulatory elements such as promoters, enhancers, silencers and insulators, and have a direct impact on gene regulation and transcriptional control by orchestrating the interaction of nuclear factors with chromatin[121-123].

1.23 Chromatin and transcriptional control: active gene loci engaging into active chromatin loops

The modifications present on histones determine whether a gene or a regulatory element is accessible, activated or deactivated. Many histone modifications have been associated with transcriptional activity [120, 124]. Figure 5 prov ides a schematic representation of how histone marks at nucleosomes shape the genome and relate to gene regulation. The MLL protein complex deposits H3K4me3 histone marks at active gene promoters [125, 126]. Active gene promoters signal the recruitment of the general transcription components such as RNA polymerase and the basal transcription pre-initiation complex (TFIID, TFIIB, TFIIH, TFIIF, TFIIE, and TFIIA) [127-129]. The TATA binding proteins (TBP), which specifically recognizes elements in promoters known as TATA boxes, recruit all the transcription components to initiate transcription of the corresponding genes found at the closest proximity of the protein-loaded promoter [130]. However, gene promoters require additional regulatory elements such as enhancers for sufficient transcriptional output.

Enhancers are defined as DNA conserved sequences of 200-500bps in length, which contains clustered recognition sites for transcription factor binding[131]. Enhancers define tissue specificity based on their activity in gene regulation (See Fig.4c). The most studied active histone marks for enhancers are the H3K4me1, H3K27ac and H3K9ac deposited by MLL, CBP/p300 and PCAF/GCN5, respectively [132]. H3K4me1 is associated with molecular priming of genomic loci[133] whereas H3K27ac and H3K9ac are associated with long-term enhancer activity. H3K27ac is a mark for potentially active enhancers to which transcription factors may bind via specific DNA motifs and form complexes with RNA polymerases, transcriptional co-activators and chromatin readers. Chromatin readers (e.g. BRD4, BRG1, CHD7, and TIP60) bind to modified histones and enhance long-term activity of the enhancer [134].

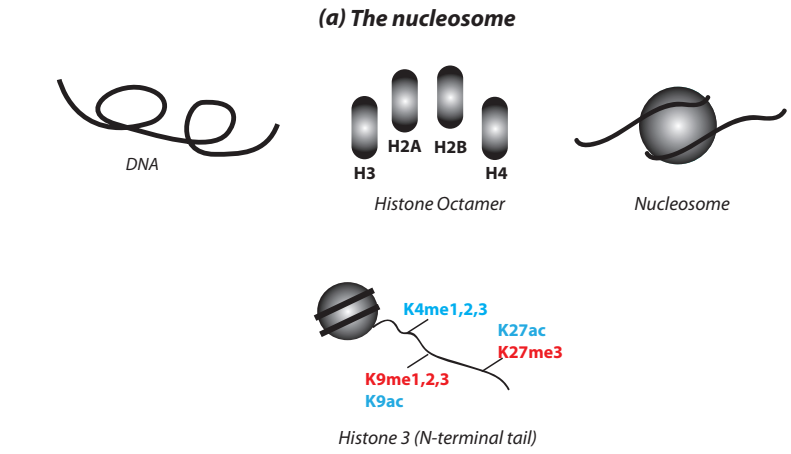


Figure 4. ***Nucleosomes and histone modifications near enhancers***. (a) The nucleosome is composed of ~147bps of DNA wrapped around an octamer of histones. All histone proteins, i.e. H2A, H2B, H3 and H4 exists as a dimer to then form together an octamer. Nucleosomes are susceptible to histone modifications that mainly occur at the N-terminal tails. Few examples of active histone modifications are illustrated in blue whereas repressive histone marks are illustrated in red. (b) Histone modifications dictate the compaction and re-compaction of nucleosomes that occur at enhancers. H3K27me3 is a mark for closed chromatin. Open chromatin shows H3K27 acetylation (H3K27ac). (c) Nucleosome-free and histone active non-coding regions such as enhancers, allow accessibility and recruitment of nuclear proteins to bind and regulate gene expression in a tissue-specific fashion. Different examples showing that tissue specific enhancers bind lineage-specific transcription factors, as illustrated with corresponding colors.

An additional feature of active enhancers is the production on-site non-coding RNA transcripts known as eRNAs that occur as bi-directional or unidirectional and non-polyadenylated. Emerging roles for eRNA in transcriptional control include recruiting and enhancing transcriptional co-activators[135], exosome-dependent RNA turnover to modulate enhancer activity[136] and involvement in higher-order chromatin regulation and organization of chromatin loops[137].

Enhancers are usually located distantly from their corresponding genes. However, in the nuclear three-dimensional space, active enhancers engage with genes and their promoters and relocate at close proximity to modulate transcriptional output by forming chromatin loops [138, 139]. This chromatin loop formation occurs at active gene loci and each loop is anchored by architectural proteins such as CTCF or cohesin[140-142]. Chromatin loop formation occurs when two CTCF-bound DNA regions (motifs) on DNA are in a converging position and ready to move towards each other while extruding out a chromatin loop, which engages promoters, their genes and their corresponding enhancers[143, 144]. The cohesion complex forms a ring structure around the anchors to secure its conformation[145-147]. These chromatin loops do not occur randomly but are contained into structural domains called topological associated domains (TADs). Each TAD size varies between 200kb to 1Mb of genomic DNA and one TAD is separated from adjacent ones by TAD-insulated borders, also bound by CTCF and cohesion (~85% of the genome)[148-150]. CTCF and cohesin also bind inside TADs, dividing genes into independent chromatin loops or sub-TADs [151-153].

Multiple adjacent TADs form high order chromatin structures highly organized into chromosomes[138]. During interphase, the TAD-structure per chromosome maintains its boundaries to remain intact by forming chromosome territories [154](See Figure 6). This high-order organization of intact chromatin folding per chromosome is based on the biophysical properties of the 30nm DNA fiber[138], which allows chromatin to fold into a nuclear diameter of 5 to 10um and promote mechanisms such as DNA repair, replication and transcription to take place in the nucleus[155-157].

The organization, structure and function of the genome is under tight regulation during development and in adult-life[142]. Any changes in function caused by mutations occurring in genes encoding major components that safe-guard the genome, such as transcription factors and chromatin regulators, or mutations in the non-coding elements where most of the regulatory part of the genome resides, predispose cells to the onset of malignancy[158-160].

1.3 C/EBPα in acute myeloid leukemia (AML)

AML is a highly heterogeneous disease of the bone marrow involving immature committed myeloid progenitors, called blast cells[161]. The two main characteristics that features AML include, bone marrow progenitors blocked in myeloid differentiation and outgrowth of these myeloid progenitors that suppress normal hematopoiesis[162]. Many driver mutations have been reported in the last decade in AML patients [163-166]. These mutations generate oncoproteins that halt the myeloid gene expression program by interfering with transcription of downstream target genes involved in cell-cycle and differentiation, or by deregulating post--transcriptional events [163]. Commonly mutated genes encode the following proteins include NPM1, FLT3, DNMT3a, K/N-RAS, IDH1/2, C/EBPα, RUNX1, KIT, fusion proteins generated by translocations such as RUNX1-ETO, PML-RARA, CBFb-MYH11, and MLL gene fusions [163

Oncoproteins in AML are generated by recurrent mutations either in coding sequences of genes [167], or in the non-coding genome [168] that leads to gene deregulation. The most common transcription factor involved in recurrent chromosomal abnormalities is RUNX1 or its counter-part CBFb [169-172]. Fusion oncoproteins generated by translocations, such as AML1-ETO in AML with a translocation t(8;21) or AML1-EVI1 in patients with a translocation t(3;21), deregulate the differentiation-mediated function of RUNX1 and disturb the transcriptional program of normal myeloid differentiation[134]. Point mutations in RUNX1 gene also lead to an abnormal RUNX1 protein function, which is normally detected in very immature AMLs with an undifferentiated phenotype[173-176]. Mutations involving the MLL methyl-transferase generates fusion genes with more than 80 partner genes involved[177-179]. These fusion oncoproteins consists of an abnormal MLL protein with a defective myeloid transcriptional program and acquires a self-renewal signature, which results in a full-blown leukemia in mice[160, 180, 181].

In AML C/EBPα is mutated in around 10% of the cases [182]. However, several reports have shown how several oncoproteins (mentioned above and in more detail in Chapter 2) in AML target the *CEBPA* gene or protein and abrogate its function[183]. Mechanisms of action of how C/EBPα is deregulated in AML at different levels of gene expression are explained in more detail in Chapter 2 of this thesis. In chapter 5 of this thesis we will discuss how the translocation t(8;21) in AML deregulates *CEBPA* expression.

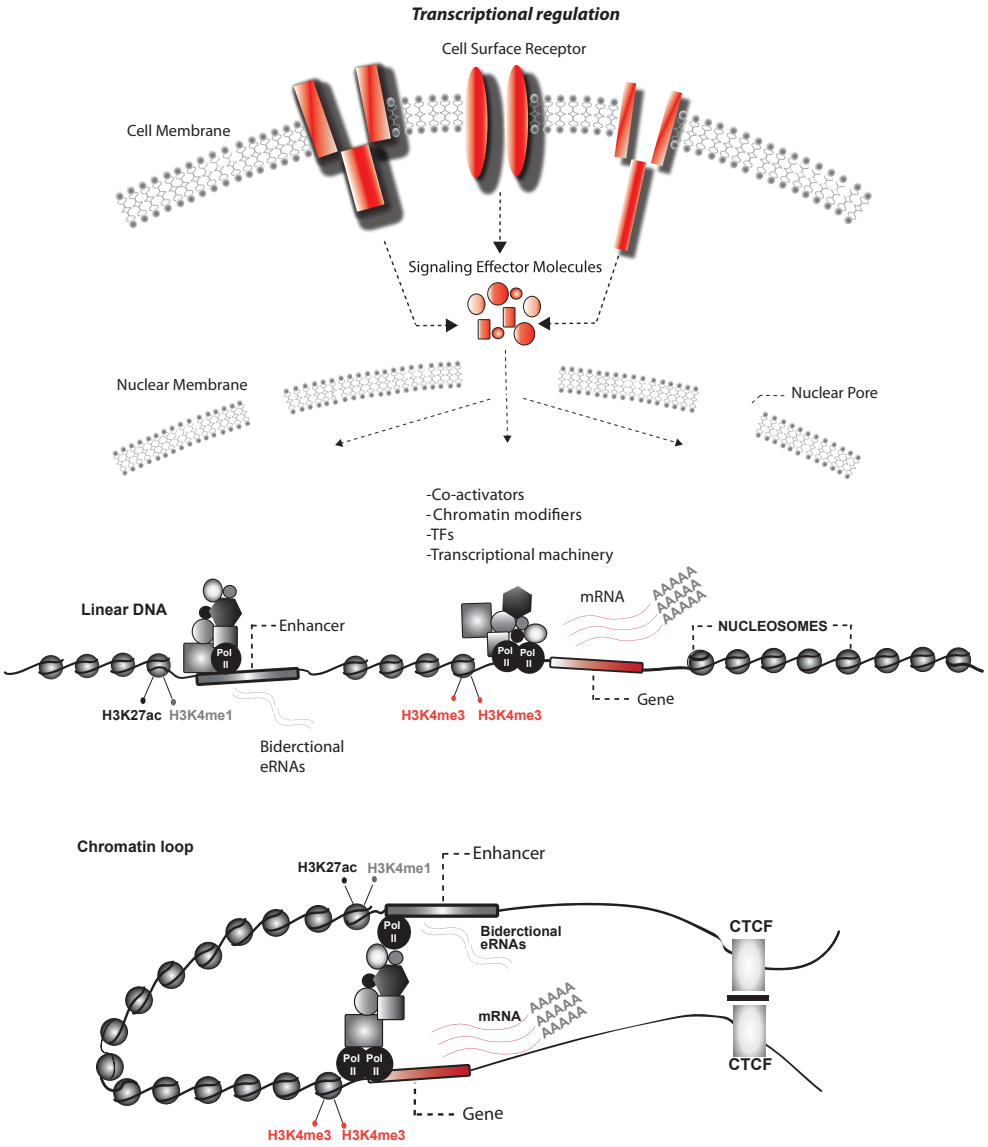


Figure 5. ***Chromatin components of transcriptional regulation***. Extracellular factors engage with cell surface receptors to initiate a cascade of signaling through the nucleus via cytoplasmic signal effector molecules. Once inside the nucleus, signaling effector molecules activate nuclear proteins such as chromatin modifiers and transcription factors to regulate gene expression. Chromatin modifiers and transcription factors free nucleosomes at gene promoters and enhancers to allow proteins of the transcription machinery to bind and activate gene expression. Chromatin modifiers deposit and read post-translational modifications such as H3K4me3 at gene promoters and H3K27ac and H3K4me1 at enhancers. Together with nuclear proteins, enhancers and genes come into close proximity upon chromatin looping a feature of transcription activation.

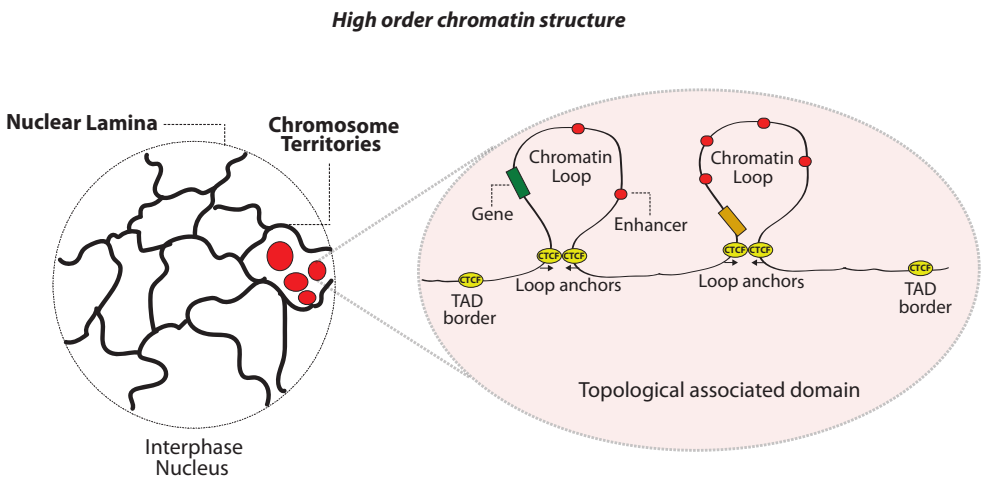


Figure 6. Nuclear sub-division into high-order chromatin structure. These TADs are separated from each other by insulated borders bound by the architectural
protein CTCF.

SCOPE OF THIS THESIS

The main objective of this thesis is to dissect the mechanisms involved in the gene regulation of *CEBPA* during normal and malignant hematopoiesis. Using a diverse range of molecular technologies combined with genome editing, in vitro and in vivo models were generated to answer questions and test for different hypotheses.

In Chapter 2, a perspective review discusses current issues concerning transcriptional mechanisms on how *CEBPA* is (1) regulated in different myeloid cells and in long-term hematopoietic stem cells and (2) deregulated in myeloid malignancies. These questions were supported by experimental approaches in models generated to study gene regulation in hematopoiesis.

In Chapter 3 we focused on the transcriptional control of *CEBPA* in normal hematopoiesis. Using chromatin profiling in human and mouse cell types, a potential myeloid *CEBPA*-enhancer was identified and its function was investigated in cell line and in vivo models to understand its role in transcription regulation of *CEBPA* and how its loss of activity influences the whole hematopoietic system.

In Chapter 4 and 5 we investigated for mechanisms related to low *CEBPA* expression levels in AML. In Chapter 4 we focused on enhancer deregulation by (1) investigating for the presence of potential DNA mutations in the *CEBPA* locus in Chapter 4, and (2) investigating for epigenetic and chromatin state deregulation by the AML oncoprotein AML1-ETO in Chapter 5.

In Chapter 6 we experimentally investigated whether loosing long-term hematopoietic stem cells in our neutropenic *Cebpa*-enhancer knock out model is a cell autonomous event caused by loss of *Cebpa* expression in LT-HSCs or whether it involves an extrinsic mechanism involving neutropenic bone marrow progenitors, HSC exhaustion and the microenvironment

In the final chapter we summarize the findings of this thesis followed by a discussion on how to further investigate and unveil more transcriptional control layers of *CEBPA* in health and disease.

REFERENCES

1. Doulatov, S., et al., *Hematopoiesis: a human perspective.* Cell Stem Cell, 2012. **10**(2): p. 120-36.
2. Jordan, C.T. and I.R. Lemischka, *Clonal and systemic analysis of long-term hematopoiesis in the mouse.* Genes Dev, 1990. **4**(2): p. 220-32.
3. Morrison, S.J. and D.T. Scadden, *The bone marrow niche for haematopoietic stem cells.* Nature, 2014. **505**(7483): p. 327-34.
4. Pietras, E.M., et al., *Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions.* Cell Stem Cell, 2015. **17**(1): p. 35-46.
5. Arzate-Mejia, R.G., D. Valle-Garcia, and F. Recillas-Targa, *Signaling epigenetics: novel insights on cell signaling and epigenetic regulation.* IUBMB Life, 2011. **63**(10): p. 881-95.
6. Atlasi, Y. and H.G. Stunnenberg, *The interplay of epigenetic marks during stem cell differentiation and development.* Nat Rev Genet, 2017. **18**(11): p. 643-658.
7. Heinz, S. and C.K. Glass, *Roles of lineage-determining transcription factors in establishing open chromatin: lessons from high-throughput studies.* Curr Top Microbiol Immunol, 2012. **356**: p. 1-15.
8. Bottardi, S., et al., *Lineage-specific transcription factors in multipotent hematopoietic progenitors: a little bit goes a long way.* Cell Cycle, 2007. **6**(9): p. 1035-9.
9. Rosenbauer, F. and D.G. Tenen, *Transcription factors in myeloid development: balancing differentiation with transformation.* Nat Rev Immunol, 2007. **7**(2): p. 105-17.
10. Crane, G.M., E. Jeffery, and S.J. Morrison, *Adult haematopoietic stem cell niches.* Nat Rev Immunol, 2017. **17**(9): p. 573-590.
11. Palis, J. and A. Koniski, *Analysis of hematopoietic progenitors in the mouse embryo.* Methods Mol Med, 2005. **105**: p. 289-302.
12. Palis, J., et al., *Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse.* Development, 1999. **126**(22): p. 5073-84.
13. Palis, J. and M.C. Yoder, *Yolk-sac hematopoiesis: the first blood cells of mouse and man.* Exp Hematol, 2001. **29**(8): p. 927-36.
14. McGrath, K.E., J.M. Frame, and J. Palis, *Early hematopoiesis and macrophage development.* Semin Immunol, 2015. **27**(6): p. 379-87.
15. Tober, J., et al., *The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis.* Blood, 2007. **109**(4): p. 1433-41.
16. Frame, J.M., K.E. McGrath, and J. Palis, *Erythro-myeloid progenitors: "definitive" hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells.* Blood Cells Mol Dis, 2013. **51**(4): p. 220-5.
17. McGrath, K. and J. Palis, *Ontogeny of erythropoiesis in the mammalian embryo.* Curr Top Dev Biol, 2008. **82**: p. 1-22.
18. McGrath, K.E., et al., *Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo.* Cell Rep, 2015. **11**(12): p. 1892-904.
19. Palis, J., et al., *Primitive erythropoiesis in the mammalian embryo.* Int J Dev Biol, 2010. **54**(6-7): p. 1011-8.
20. Tober, J., K.E. McGrath, and J. Palis, *Primitive erythropoiesis and megakaryopoiesis in the yolk sac are independent of c-myb.* Blood, 2008. **111**(5): p. 2636-9.

21. Palis, J., *Hematopoietic stem cell-independent hematopoiesis: emergence of erythroid, megakaryocyte, and myeloid potential in the mammalian embryo.* FEBS Lett, 2016. **590**(22): p. 3965-3974.
22. Yokota, T., et al., *Tracing the first waves of lymphopoiesis in mice.* Development, 2006. **133**(10): p. 2041-51.
23. Bertrand, J.Y., et al., *Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo.* Development, 2007. **134**(23): p. 4147-56.
24. Lugus, J.J., et al., *GATA2 functions at multiple steps in hemangioblast development and differentiation.* Development, 2007. **134**(2): p. 393-405.
25. Mukouyama, Y., et al., *The AML1 transcription factor functions to develop and maintain hematogenic precursor cells in the embryonic aorta-gonad-mesonephros region.* Dev Biol, 2000. **220**(1): p. 27-36.
26. Wilson, N.K., et al., *Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators.* Cell Stem Cell, 2010. **7**(4): p. 532-44.
27. Bowie, M.B., et al., *Hematopoietic stem cells proliferate until after birth and show a reversible phase-specific engraftment defect.* J Clin Invest, 2006. **116**(10): p. 2808-16.
28. Spangrude, G.J., D.M. Brooks, and D.B. Tumas, *Long-term repopulation of irradiated mice with limiting numbers of purified hematopoietic stem cells: in vivo expansion of stem cell phenotype but not function.* Blood, 1995. **85**(4): p. 1006-16.
29. Cheshier, S.H., et al., *In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells.* Proc Natl Acad Sci U S A, 1999. **96**(6): p. 3120-5.
30. Kiel, M.J., et al., *Haematopoietic stem cells do not asymmetrically segregate chromosomes or retain BrdU.* Nature, 2007. **449**(7159): p. 238-42.
31. Kiel, M.J. and S.J. Morrison, *Maintaining hematopoietic stem cells in the vascular niche.* Immunity, 2006. **25**(6): p. 862-4.
32. Kunisaki, Y., et al., *Arteriolar niches maintain haematopoietic stem cell quiescence.* Nature, 2013. **502**(7473): p. 637-43.
33. Silberstein, L.E. and C.P. Lin, *A new image of the hematopoietic stem cell vascular niche.* Cell Stem Cell, 2013. **13**(5): p. 514-6.
34. Nombela-Arrieta, C., et al., *Quantitative imaging of haematopoietic stem and progenitor cell localization and hypoxic status in the bone marrow microenvironment.* Nat Cell Biol, 2013. **15**(5): p. 533-43.
35. Horio, E., et al., *Role of endothelial cell-derived angptl2 in vascular inflammation leading to endothelial dysfunction and atherosclerosis progression.* Arterioscler Thromb Vasc Biol, 2014. **34**(4): p. 790-800.
36. Ding, L. and S.J. Morrison, *Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches.* Nature, 2013. **495**(7440): p. 231-5.
37. Hanoun, M. and P.S. Frenette, *This niche is a maze; an amazing niche.* Cell Stem Cell, 2013. **12**(4): p. 391-2.
38. Greenbaum, A., et al., *CXCL12 in early mesenchymal progenitors is required for haematopoietic stem-cell maintenance.* Nature, 2013. **495**(7440): p. 227-30.
39. Thoren, L.A., et al., *Kit regulates maintenance of quiescent hematopoietic stem cells.* J Immunol, 2008. **180**(4): p. 2045-53.
40. Bruns, I., et al., *Megakaryocytes regulate hematopoietic stem cell quiescence through CXCL4 secretion.* Nat Med, 2014. **20**(11): p. 1315-20.

41. Zhao, M., et al., *Megakaryocytes maintain homeostatic quiescence and promote post-injury regeneration of hematopoietic stem cells.* Nat Med, 2014. **20**(11): p. 1321-6.
42. Beck, D., et al., *Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes.* Blood, 2013. **122**(14): p. e12-22.
43. Diffner, E., et al., *Activity of a heptad of transcription factors is associated with stem cell programs and clinical outcome in acute myeloid leukemia.* Blood, 2013. **121**(12): p. 2289-300.
44. Akashi, K., et al., *Lymphoid development from stem cells and the common lymphocyte progenitors.* Cold Spring Harb Symp Quant Biol, 1999. **64**: p. 1-12.
45. Akashi, K., et al., *A clonogenic common myeloid progenitor that gives rise to all myeloid lineages.* Nature, 2000. **404**(6774): p. 193-7.
46. Kondo, M., et al., *Cell-fate conversion of lymphoid-committed progenitors by instructive actions of cytokines.* Nature, 2000. **407**(6802): p. 383-6.
47. Manz, M.G., et al., *Prospective isolation of human clonogenic common myeloid progenitors.* Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11872-7.
48. Traver, D., et al., *Fetal liver myelopoiesis occurs through distinct, prospectively isolatable progenitor subsets.* Blood, 2001. **98**(3): p. 627-35.
49. Kawamura, S., et al., *Identification of a Human Clonogenic Progenitor with Strict Monocyte Differentiation Potential: A Counterpart of Mouse cMoPs.* Immunity, 2017. **46**(5): p. 835-848 e4.
50. Manz, M.G., et al., *Dendritic cell development from common myeloid progenitors.* Ann N Y Acad Sci, 2001. **938**: p. 167-73; discussion 173-4.
51. Lee, J., et al., *Lineage specification of human dendritic cells is marked by IRF8 expression in hematopoietic stem cells and multipotent progenitors.* Nat Immunol, 2017. **18**(8): p. 877-888.
52. See, P., et al., *Mapping the human DC lineage through the integration of high-dimensional techniques.* Science, 2017. **356**(6342).
53. van Beek, J.J., et al., *Human blood myeloid and plasmacytoid dendritic cells cross activate each other and synergize in inducing NK cell cytotoxicity.* Oncoimmunology, 2016. **5**(10): p. e1227902.
54. Wilson, N.K., et al., *Integrated genome-scale analysis of the transcriptional regulatory landscape in a blood stem/progenitor cell model.* Blood, 2016. **127**(13): p. e12-23.
55. Kanno, Y., et al., *Immune cell-specific amplification of interferon signaling by the IRF-4/8-PU.1 complex.* J Interferon Cytokine Res, 2005. **25**(12): p. 770-9.
56. Kurotaki, D., et al., *IRF8 inhibits C/EBPalpha activity to restrain mononuclear phagocyte progenitors from differentiating into neutrophils.* Nat Commun, 2014. **5**: p. 4978.
57. Yanez, A. and H.S. Goodridge, *Interferon regulatory factor 8 and the regulation of neutrophil, monocyte, and dendritic cell production.* Curr Opin Hematol, 2016. **23**(1): p. 11-7.
58. Reddy, V.A., et al., *Granulocyte inducer C/EBPalpha inactivates the myeloid master regulator PU.1: possible role in lineage commitment decisions.* Blood, 2002. **100**(2): p. 483-90.
59. Friedman, A.D., et al., *Regulation of granulocyte and monocyte differentiation by CCAAT/enhancer binding protein alpha.* Blood Cells Mol Dis, 2003. **31**(3): p. 338-41.
60. Iwasaki, H., et al., *The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages.* Genes Dev, 2006. **20**(21): p. 3010-21.

61. Arinobu, Y., et al., *Developmental checkpoints of the basophil/mast cell lineages in adult murine hematopoiesis.* Proc Natl Acad Sci U S A, 2005. **102**(50): p. 18105-10.
62. Dore, L.C. and J.D. Crispino, *Transcription factor networks in erythroid cell and megakaryocyte development.* Blood, 2011. **118**(2): p. 231-9.
63. Hasemann, M.S., et al., *Phosphorylation of serine 248 of C/EBPalpha is dispensable for myelopoiesis but its disruption leads to a low penetrant myeloid disorder with long latency.* PLoS One, 2012. **7**(6): p. e38841.
64. Xie, H., et al., *Stepwise reprogramming of B cells into macrophages.* Cell, 2004. **117**(5): p. 663-76.
65. Feng, R., et al., *PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells.* Proc Natl Acad Sci U S A, 2008. **105**(16): p. 6057-62.
66. Graf, T. and T. Enver, *Forcing cells to change lineages.* Nature, 2009. **462**(7273): p. 587-94.
67. Laiosa, C.V., M. Stadtfeld, and T. Graf, *Determinants of lymphoid-myeloid lineage diversification.* Annu Rev Immunol, 2006. **24**: p. 705-38.
68. Jaitin, D.A., et al., *Each cell counts: Hematopoiesis and immunity research in the era of single cell genomics.* Semin Immunol, 2015. **27**(1): p. 67-71.
69. Giladi, A. and I. Amit, *Single-Cell Genomics: A Stepping Stone for Future Immunology Discoveries.* Cell, 2018. **172**(1-2): p. 14-21.
70. Paul, F., et al., *Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors.* Cell, 2015. **163**(7): p. 1663-77.
71. Carrelha, J., et al., *Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells.* Nature, 2018.
72. Laurenti, E. and B. Gottgens, *From haematopoietic stem cells to complex differentiation landscapes.* Nature, 2018. **553**(7689): p. 418-426.
73. Notta, F., et al., *Distinct routes of lineage development reshape the human blood hierarchy across ontogeny.* Science, 2016. **351**(6269): p. aab2116.
74. Jenkins, N.A., et al., *Mouse chromosomal location of the CCAAT/enhancer binding proteins C/EBP beta (Cebpb), C/EBP delta (Cebpd), and CRP1 (Cebpe).* Genomics, 1995. **28**(2): p. 333-6.
75. Landschulz, W.H., et al., *Isolation of a recombinant copy of the gene encoding C/EBP.* Genes Dev, 1988. **2**(7): p. 786-800.
76. Tsukada, J., et al., *The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation.* Cytokine, 2011. **54**(1): p. 6-19.
77. Nerlov, C., *The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control.* Trends Cell Biol, 2007. **17**(7): p. 318-24.
78. Nerlov, C., *C/EBPs: recipients of extracellular signals through proteome modulation.* Curr Opin Cell Biol, 2008. **20**(2): p. 180-5.
79. Lourenco, A.R. and P.J. Coffer, *A tumor suppressor role for C/EBPalpha in solid tumors: more than fat and blood.* Oncogene, 2017. **36**(37): p. 5221-5230.
80. Zhang, D.E., et al., *Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice.* Proc Natl Acad Sci U S A, 1997. **94**(2): p. 569-74.
81. Begay, V., J. Smink, and A. Leutz, *Essential requirement of CCAAT/enhancer binding proteins in embryogenesis.* Mol Cell Biol, 2004. **24**(22): p. 9744-51.

82. Heath, V., et al., *C/EBPalpha deficiency results in hyperproliferation of hematopoietic progenitor cells and disrupts macrophage development in vitro and in vivo.* Blood, 2004. **104**(6): p. 1639-47.
83. Ye, M., et al., *C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence.* Nat Cell Biol, 2013. **15**(4): p. 385-94.
84. Zhang, P., et al., *Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha.* Immunity, 2004. **21**(6): p. 853-63.
85. Zhang, P., et al., *Induction of granulocytic differentiation by 2 pathways.* Blood, 2002. **99**(12): p. 4406-12.
86. Wang, W., et al., *C/EBPalpha and G-CSF receptor signals cooperate to induce the myeloperoxidase and neutrophil elastase genes.* Leukemia, 2001. **15**(5): p. 779-86.
87. Khanna-Gupta, A., et al., *Chromatin immunoprecipitation (ChIP) studies indicate a role for CCAAT enhancer binding proteins alpha and epsilon (C/EBP alpha and C/EBP epsilon) and CDP/cut in myeloid maturation-induced lactoferrin gene expression.* Blood, 2003. **101**(9): p. 3460-8.
88. Gombart, A.F., et al., *Regulation of neutrophil and eosinophil secondary granule gene expression by transcription factors C/EBP epsilon and PU.1.* Blood, 2003. **101**(8): p. 3265-73.
89. Zhang, P., et al., *Upregulation of interleukin 6 and granulocyte colony-stimulating factor receptors by transcription factor CCAAT enhancer binding protein alpha (C/EBP alpha) is critical for granulopoiesis.* J Exp Med, 1998. **188**(6): p. 1173-84.
90. Saunders, M.A., et al., *Selective suppression of CCAAT/enhancer-binding protein beta binding and cyclooxygenase-2 promoter activity by sodium salicylate in quiescent human fibroblasts.* J Biol Chem, 2001. **276**(22): p. 18897-904.
91. Wang, D., I. Paz-Priel, and A.D. Friedman, *NF-kappa B p50 regulates C/EBP alpha expression and inflammatory cytokine-induced neutrophil production.* J Immunol, 2009. **182**(9): p. 5757-62.
92. van Oevelen, C., et al., *C/EBPalpha Activates Pre-existing and De Novo Macrophage Enhancers during Induced Pre-B Cell Transdifferentiation and Myelopoiesis.* Stem Cell Reports, 2015. **5**(2): p. 232-47.
93. Pedersen, T.A., et al., *Cooperation between C/EBPalpha TBP/TFIIB and SWI/SNF recruiting domains is required for adipocyte differentiation.* Genes Dev, 2001. **15**(23): p. 3208-16.
94. Muller, C., et al., *The CCAAT enhancer-binding protein alpha (C/EBPalpha) requires a SWI/SNF complex for proliferation arrest.* J Biol Chem, 2004. **279**(8): p. 7353-8.
95. Xu, J., Y. Kawai, and I.J. Arinze, *Dual role of C/EBPalpha as an activator and repressor of Galphai2 gene transcription.* Genes Cells, 2013. **18**(12): p. 1082-94.
96. Savickiene, J., et al., *Epigenetic and molecular mechanisms underlying the antileukemic activity of the histone deacetylase inhibitor belinostat in human acute promyelocytic leukemia cells.* Anticancer Drugs, 2014. **25**(8): p. 938-49.
97. Johnson, K., et al., *Regulatory events in early and late B-cell differentiation.* Mol Immunol, 2005. **42**(7): p. 749-61.
98. Seydel, F., et al., *GM-CSF induces STAT5 binding at epigenetic regulatory sites within the Csf2 promoter of non-obese diabetic (NOD) mouse myeloid cells.* J Autoimmun, 2008. **31**(4): p. 377-84.
99. Josefowicz, S.Z., et al., *Chromatin Kinases Act on Transcription Factors and Histone Tails in Regulation of Inducible Transcription.* Mol Cell, 2016. **64**(2): p. 347-361.

100. Lin, Y.C., et al., *Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate.* Nat Immunol, 2012. **13**(12): p. 1196-204.
101. Ramaswamy, A. and I. Ioshikhes, *Dynamics of modeled oligonucleosomes and the role of histone variant proteins in nucleosome organization.* Adv Protein Chem Struct Biol, 2013. **90**: p. 119-49.
102. Neff, T. and S.A. Armstrong, *Chromatin maps, histone modifications and leukemia.* Leukemia, 2009. **23**(7): p. 1243-51.
103. Pradeepa, M.M., et al., *Histone H3 globular domain acetylation identifies a new class of enhancers.* Nat Genet, 2016. **48**(6): p. 681-6.
104. Lennartsson, A. and K. Ekwall, *Histone modification patterns and epigenetic codes.* Biochim Biophys Acta, 2009. **1790**(9): p. 863-8.
105. Hong, L., et al., *Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA.* J Biol Chem, 1993. **268**(1): p. 305-14.
106. Jorgensen, S., G. Schotta, and C.S. Sorensen, *Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity.* Nucleic Acids Res, 2013. **41**(5): p. 2797-806.
107. Megee, P.C., B.A. Morgan, and M.M. Smith, *Histone H4 and the maintenance of genome integrity.* Genes Dev, 1995. **9**(14): p. 1716-27.
108. Dion, M.F., et al., *Genomic characterization reveals a simple histone H4 acetylation code.* Proc Natl Acad Sci U S A, 2005. **102**(15): p. 5501-6.
109. Martin, A.M., et al., *Redundant roles for histone H3 N-terminal lysine residues in subtelomeric gene repression in Saccharomyces cerevisiae.* Genetics, 2004. **167**(3): p. 1123-32.
110. Zentner, G.E. and S. Henikoff, *Regulation of nucleosome dynamics by histone modifications.* Nat Struct Mol Biol, 2013. **20**(3): p. 259-66.
111. Bowers, E.M., et al., *Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor.* Chem Biol, 2010. **17**(5): p. 471-82.
112. Marmorstein, R. and S.Y. Roth, *Histone acetyltransferases: function, structure, and catalysis.* Curr Opin Genet Dev, 2001. **11**(2): p. 155-61.
113. Roth, S.Y., J.M. Denu, and C.D. Allis, *Histone acetyltransferases.* Annu Rev Biochem, 2001. **70**: p. 81-120.
114. Voss, A.K. and T. Thomas, *MYST family histone acetyltransferases take center stage in stem cells and development.* Bioessays, 2009. **31**(10): p. 1050-61.
115. Chen, H., et al., *Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300.* Cell, 1997. **90**(3): p. 569-80.
116. Marmorstein, R., *Structure of histone acetyltransferases.* J Mol Biol, 2001. **311**(3): p. 433-44.
117. Sun, X.J., et al., *The Role of Histone Acetyltransferases in Normal and Malignant Hematopoiesis.* Front Oncol, 2015. **5**: p. 108.
118. Vu, L.P., L. Luciani, and S.D. Nimer, *Histone-modifying enzymes: their role in the pathogenesis of acute leukemia and their therapeutic potential.* Int J Hematol, 2013. **97**(2): p. 198-209.
119. Lenstra, T.L., et al., *The specificity and topology of chromatin interaction pathways in yeast.* Mol Cell, 2011. **42**(4): p. 536-49.
120. Munshi, A., et al., *Histone modifications dictate specific biological readouts.* J Genet Genomics, 2009. **36**(2): p. 75-88.

121. Glass, C.K., *Genetic and genomic approaches to understanding macrophage identity and function.* Arterioscler Thromb Vasc Biol, 2015. **35**(4): p. 755-62.
122. Heinz, S., et al., *The selection and function of cell type-specific enhancers.* Nat Rev Mol Cell Biol, 2015. **16**(3): p. 144-54.
123. Link, V.M., D. Gosselin, and C.K. Glass, *Mechanisms Underlying the Selection and Function of Macrophage-Specific Enhancers.* Cold Spring Harb Symp Quant Biol, 2015. **80**: p. 213-21.
124. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications.* Cell Res, 2011. **21**(3): p. 381-95.
125. Chang, P.Y., et al., *Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription.* J Mol Biol, 2010. **400**(2): p. 137-44.
126. Senisterra, G., et al., *Small-molecule inhibition of MLL activity by disruption of its interaction with WDR5.* Biochem J, 2013. **449**(1): p. 151-9.
127. Kaiser, K. and M. Meisterernst, *The human general co-factors.* Trends Biochem Sci, 1996. **21**(9): p. 342-5.
128. Reines, D., J.W. Conaway, and R.C. Conaway, *The RNA polymerase II general elongation factors.* Trends Biochem Sci, 1996. **21**(9): p. 351-5.
129. Roeder, R.G., *The role of general initiation factors in transcription by RNA polymerase II.* Trends Biochem Sci, 1996. **21**(9): p. 327-35.
130. Burley, S.K., *The TATA box binding protein.* Curr Opin Struct Biol, 1996. **6**(1): p. 69-75.
131. Spitz, F. and E.E. Furlong, *Transcription factors: from enhancer binding to developmental control.* Nat Rev Genet, 2012. **13**(9): p. 613-26.
132. Jin, Q., et al., *Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation.* EMBO J, 2011. **30**(2): p. 249-62.
133. Mercer, E.M., et al., *Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors.* Immunity, 2011. **35**(3): p. 413-25.
134. Loke, J., et al., *RUNX1-ETO and RUNX1-EVI1 Differentially Reprogram the Chromatin Landscape in t(8;21) and t(3;21) AML.* Cell Rep, 2017. **19**(8): p. 1654-1668.
135. Bose, D.A., et al., *RNA Binding to CBP Stimulates Histone Acetylation and Transcription.* Cell, 2017. **168**(1-2): p. 135-149 e22.
136. Pefanis, E., et al., *RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity.* Cell, 2015. **161**(4): p. 774-89.
137. Xu, J.W., S. Ling, and J. Liu, *Higher-Order Chromatin Regulation of Inflammatory Gene Expression.* Mediators Inflamm, 2017. **2017**: p. 7848591.
138. Sewitz, S.A., Z. Fahmi, and K. Lipkow, *Higher order assembly: folding the chromosome.* Curr Opin Struct Biol, 2017. **42**: p. 162-168.
139. Doyle, B., et al., *Chromatin loops as allosteric modulators of enhancer-promoter interactions.* PLoS Comput Biol, 2014. **10**(10): p. e1003867.
140. Oti, M., et al., *CTCF-mediated chromatin loops enclose inducible gene regulatory domains.* BMC Genomics, 2016. **17**: p. 252.
141. Tark-Dame, M., et al., *Depletion of the chromatin looping proteins CTCF and cohesin causes chromatin compaction: insight into chromatin folding by polymer modelling.* PLoS Comput Biol, 2014. **10**(10): p. e1003877.

142. Noordermeer, D. and D. Duboule, *Chromatin looping and organization at developmentally regulated gene loci.* Wiley Interdiscip Rev Dev Biol, 2013. **2**(5): p. 615-30.
143. Goloborodko, A., J.F. Marko, and L.A. Mirny, *Chromosome Compaction by Active Loop Extrusion.* Biophys J, 2016. **110**(10): p. 2162-8.
144. Sanborn, A.L., et al., *Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes.* Proc Natl Acad Sci U S A, 2015. **112**(47): p. E6456-65.
145. Rao, S.S.P., et al., *Cohesin Loss Eliminates All Loop Domains.* Cell, 2017. **171**(2): p. 305-320 e24.
146. Haarhuis, J.H.I., et al., *The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension.* Cell, 2017. **169**(4): p. 693-707 e14.
147. Barrington, C., R. Finn, and S. Hadjur, *Cohesin biology meets the loop extrusion model.* Chromosome Res, 2017. **25**(1): p. 51-60.
148. Dixon, J.R., et al., *Topological domains in mammalian genomes identified by analysis of chromatin interactions.* Nature, 2012. **485**(7398): p. 376-80.
149. Hansen, A.S., et al., *Recent evidence that TADs and chromatin loops are dynamic structures.* Nucleus, 2018. **9**(1): p. 20-32.
150. Nora, E.P., et al., *Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization.* Cell, 2017. **169**(5): p. 930-944 e22.
151. Berlivet, S., et al., *Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs.* PLoS Genet, 2013. **9**(12): p. e1004018.
152. Deng, W. and G.A. Blobel, *Detecting Long-Range Enhancer-Promoter Interactions by Quantitative Chromosome Conformation Capture.* Methods Mol Biol, 2017. **1468**: p. 51-62.
153. Matharu, N. and N. Ahituv, *Minor Loops in Major Folds: Enhancer-Promoter Looping, Chromatin Restructuring, and Their Association with Transcriptional Regulation and Disease.* PLoS Genet, 2015. **11**(12): p. e1005640.
154. Dixon, J.R., D.U. Gorkin, and B. Ren, *Chromatin Domains: The Unit of Chromosome Organization.* Mol Cell, 2016. **62**(5): p. 668-80.
155. Zhao, P.A., J.C. Rivera-Mulia, and D.M. Gilbert, *Replication Domains: Genome Compartmentalization into Functional Replication Units.* Adv Exp Med Biol, 2017. **1042**: p. 229-257.
156. Petryk, N., et al., *Replication landscape of the human genome.* Nat Commun, 2016. **7**: p. 10208.
157. Pope, B.D., et al., *Topologically associating domains are stable units of replication-timing regulation.* Nature, 2014. **515**(7527): p. 402-5.
158. Jia, R., et al., *Novel insights into chromosomal conformations in cancer.* Mol Cancer, 2017. **16**(1): p. 173.
159. Fowler, F. and J.K. Tyler, *Anchoring Chromatin Loops to Cancer.* Dev Cell, 2017. **42**(3): p. 209-211.
160. Folle, G.A., *Nuclear architecture, chromosome domains and genetic damage.* Mutat Res, 2008. **658**(3): p. 172-83.
161. Aziz, H., et al., *Gene Mutations as Emerging Biomarkers and Therapeutic Targets for Relapsed Acute Myeloid Leukemia.* Front Pharmacol, 2017. **8**: p. 897.
162. Thomas, D. and R. Majeti, *Biology and relevance of human acute myeloid leukemia stem cells.* Blood, 2017. **129**(12): p. 1577-1585.
163. Cancer Genome Atlas Research, N., et al., *Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia.* N Engl J Med, 2013. **368**(22): p. 2059-74.

164. Welch, J.S., et al., *The origin and evolution of mutations in acute myeloid leukemia.* Cell, 2012. **150**(2): p. 264-78.

165. Walter, M.J., et al., *Clonal architecture of secondary acute myeloid leukemia.* N Engl J Med, 2012. **366**(12): p. 1090-8.

166. Ley, T.J., et al., *DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome.* Nature, 2008. **456**(7218): p. 66-72.

167. Medinger, M. and J.R. Passweg, *Acute myeloid leukaemia genomics.* Br J Haematol, 2017. **179**(4): p. 530-542.

168. Groschel, S., et al., *A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia.* Cell, 2014. **157**(2): p. 369-381.

169. Yeh, J.R., et al., *AML1-ETO reprograms hematopoietic cell fate by downregulating scl expression.* Development, 2008. **135**(2): p. 401-10.

170. Paschka, P., *Core binding factor acute myeloid leukemia.* Semin Oncol, 2008. **35**(4): p. 410-7.

171. Niebuhr, B., et al., *Gatekeeper function of the RUNX1 transcription factor in acute leukemia.* Blood Cells Mol Dis, 2008. **40**(2): p. 211-8.

172. Maki, K., T. Yamagata, and K. Mitani, *Role of the RUNX1-EVI1 fusion gene in leukemogenesis.* Cancer Sci, 2008. **99**(10): p. 1878-83.

173. Sood, R., Y. Kamikubo, and P. Liu, *Role of RUNX1 in hematological malignancies.* Blood, 2017. **129**(15): p. 2070-2082.

174. Koeffler, H.P. and G. Leong, *Preleukemia: one name, many meanings.* Leukemia, 2017. **31**(3): p. 534-542.

175. Deltcheva, E. and R. Nimmo, *RUNX transcription factors at the interface of stem cells and cancer.* Biochem J, 2017. **474**(11): p. 1755-1768.

176. Bellissimo, D.C. and N.A. Speck, *RUNX1 Mutations in Inherited and Sporadic Leukemia.* Front Cell Dev Biol, 2017. **5**: p. 111.

177. Winters, A.C. and K.M. Bernt, *MLL-Rearranged Leukemias-An Update on Science and Clinical Approaches.* Front Pediatr, 2017. **5**: p. 4.

178. Yokoyama, A., *Transcriptional activation by MLL fusion proteins in leukemogenesis.* Exp Hematol, 2017. **46**: p. 21-30.

179. Brien, G.L., D.G. Valerio, and S.A. Armstrong, *Exploiting the Epigenome to Control Cancer-Promoting Gene-Expression Programs.* Cancer Cell, 2016. **29**(4): p. 464-476.

180. Deshpande, A.J., J. Bradner, and S.A. Armstrong, *Chromatin modifications as therapeutic targets in MLL-rearranged leukemia.* Trends Immunol, 2012. **33**(11): p. 563-70.

181. Krivtsov, A.V. and S.A. Armstrong, *MLL translocations, histone modifications and leukaemia stem-cell development.* Nat Rev Cancer, 2007. **7**(11): p. 823-33.

182. Wouters, B.J., et al., *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, 2009. **113**(13): p. 3088-91.

183. Avellino, R. and R. Delwel, *Expression and regulation of C/EBPalpha in normal myelopoiesis and in malignant transformation.* Blood, 2017. **129**(15): p. 2083-2091.

2

THESIS_Roberto_Avellino.indd 36 07-05-18 10:17

EXPRESSION AND REGULATION OF C/EBPα IN NORMAL MYELOPOIESIS AND IN MALIGNANT TRANSFORMATION

Roberto Avellino and Ruud Delwel

Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands

Blood 2017

Running title: C/EBPα in hematopoiesis and myeloid malignancy

THESIS_Roberto_Avellino.indd 37 07-05-18 10:17

ABSTRACT

One of the most studied transcription factors in hematopoiesis is the leucine zipper CCAAT enhancer binding protein alpha (C/EBPα), which is mainly involved in cell fate decisions for myeloid differentiation. Its involvement in acute myeloid leukemia (AML) is diverse, with patients frequently exhibiting mutations, deregulation of gene expression or alterations in the function of C/EBPα. In this review, we emphasize the importance of C/EBPα for neutrophil maturation, its role in myeloid priming of hematopoietic stem and progenitor cells and its indispensable requirement for AML development. We discuss that mutations in the ORF of *CEBPA* lead to an altered C/EBPα function, affecting expression of downstream genes and consequently deregulating myelopoiesis. The emerging transcriptional mechanisms of *CEBPA* are discussed based on recent studies. Novel insights on how these mechanisms may be deregulated by oncoproteins or mutations/variants in *CEBPA*-enhancers are suggested in principal to reveal novel mechanisms of how *CEBPA* is deregulated at the transcriptional level.

INTRODUCTION

Lineage specific transcription factors (LTFs) prime chromatin states of hematopoietic stem/progenitor cells (HSPCs) to drive commitment and differentiation of specific cell types in the bone marrow [1]. One of the most studied lineage specific TFs involved in hematopoietic development is CCAAT enhancer binding protein alpha (C/EBPα), a leucine zipper transcription factor mainly involved in myeloid development. C/EBPα has been reported to be involved in monopoieisis and granulopoiesis. The mechanism of action of C/EBPα and the interaction with other leucine zipper proteins driving monopoiesis has been extensively reviewed previously [2]. Here we particularly focus on the role of C/EBPα in early hematopoietic development, granulopoiesis and malignant transformation of myeloid progenitor cells. C/EBPα is the founder of the C/EBP family of transcription factors, consisting of C/EBPα, C/EBPβ, C/EBPδ, C/EBPε, C/EBPγ and C/EBPζ, which are named according to their order of discovery[3-12]. All CEBP members share a similar C-terminal domain for DNA binding and dimerization but differ in the N-terminal domain, with *CEBPA* possessing two transactivation domains for transcription control and protein interactions. The CEBP family of TFs are involved in many different biological pathways, which has been discussed elsewhere[13].

Here we focus on the importance of C/EBPα as a major TF of the neutrophilic differentiation program, and how it is recognized as an indispensable factor for the initiation of acute myeloid leukemia (AML). At the same time, frequent aberrations deregulating C/EBPα function or expression are observed in different AML subtypes. Based on recent findings, we will discuss the transcriptional control driven by a specific enhancer regulating *CEBPA* in the bone marrow, followed by the potential role of this enhancer to be hijacked by different AML-related onco-proteins to deregulate *CEBPA* expression. By placing C/EBPα at the center of the myeloid lineage hierarchy, this review offers a perspective on C/EBPα as a target of diverse physiological and oncogenic events, which ultimately contribute to the onset or development of AML.

C/EBPα IN NORMAL AND MALIGNANT MYELOID PRIMING AND DEVELOPMENT

The non-redundant role of **C/***EBPα for neutrophil development***.**

CEBPA is an intron-less gene located on chromosome 19q in humans and on chromosome 7 in mice, which encodes a 42KD and a 30KD DNA binding protein, both derived from the same gene but translated from two distinct AUG translational start sites (Figure 1).

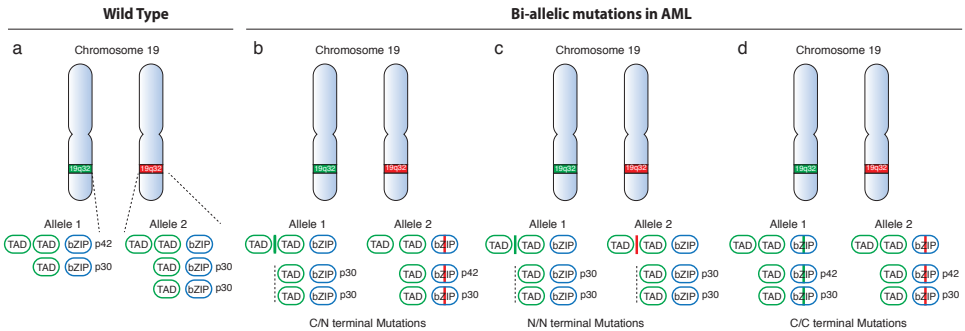


Figure 1: *CEBPA* is located on chromosome 19q.32. From the *CEBPA* mRNA two major proteins are generated from two distinct AUG start sites, i.e. a p42 and a p30 isoform (1a). *CEBPA* may be transcribed from the two alleles (Allele A. Bi-allelic mutations in *CEBPA* may occur at the N-terminus or at the C-terminus (C/N mutation). Patients with a C/N double mutation generate a p30 isoform only from one allele (termed allele 1) and bZIP mutated p42 and p30 isoforms (red) from the other allele (1b). Bi-allelic mutations at the N-terminal (N/N) generates only p30 isoforms but not p42 isoforms (1c). Bi-allelic mutations at the C-terminus (C/C) generate p42/p30 C-terminally mutated isoforms from both alleles, all defective in the bZIP domain (1d). *CEBPA* double mutant leukemias never express a wild type p42 protein. TAD=Transcription Activating Domain. bZIP= Basic Leucine Zipper Domain. **I**=N-terminal mutation; **I**= C-terminal mutation.

Besides its role in the bone marrow, C/EBPα is also essential for the development of other organs, such as lung, liver, intestine and female reproductive organs[14-21]. In the hematopoietic system, C/EBPα is primarily expressed in cells of the myeloid lineage. Zhang and colleagues were the first to show that germ-line deletion of *Cebpa* in mice causes a block in neutrophil differentiation in bone marrow[19]. This model limited the study of C/EBPα in adult hematopoiesis, since mice died of lung and liver complications shortly after birth. By generating an Mx1-Cre driven conditional knock-out mice, it was demonstrated that the excision of *Cebpa* in the bone marrow of adult mice failed to generate granulocyte/monocyte progenitors and resulted in a complete block of neutrophilic development at the CMP (common myeloid progenitor) stage[22].

During cell fate decisions, C/EBPα primes and activates the myeloid gene expression program by binding promoters or enhancers of myeloid-related genes such as *CSF3R*, *IL-6R, CEBPE, GFI-1* or *KLF5*[23-26], in mouse models as well as in human CD34+ HSPCs of either cord blood or leukemic origin [27, 28]. C/EBPα competes with other transcription factors

to attenuate the expression of non-myeloid lineage genes in progenitors of multi-lineage potential [29-31]. In support with this latter observation, murine HSPCs isolated shortly after *Cebpa* deletion lose the expression of certain myeloid genes and recapitulate the expression of T-cell genes such as *Cd7* or *Lck*, suggesting a switch towards a myeloid/T-lymphoid phenotype. This mixed myeloid/T-lymphoid phenotype is also observed in a rare $CD34^+$ leukemia subtype in humans, in which *CEBPA* is silenced by DNA hyper-methylation[32].

This early phenotype observed upon *CEBPA* deactivation in murine models and human leukemia, indicates that (1) C/EBPα has a critical role in regulation of myeloid gene expression at an early hematopoietic stage, and (2) it acts as a repressor of non-myeloid genes. Both observations are discussed further in the coming sections and supported by *in vitro* and *in vivo* models.

Myeloid priming of hematopoietic stem cells by C/EBPα.

Given its probable function as a pioneer transcription factor[33, 34], at which stage of differentiation does C/EBPα prime the myeloid genome? Porse and colleagues conducted chromatin studies in murine $cKit^+$ $CD150^+$ long term HSCs (LT-HSCS) to show that C/ebpα binds to chromatin at loci of myeloid associated genes before they are marked by active chromatin modifications [34]. This suggests that C/ebpα binds and primes genes for myeloid commitment. A number of studies suggest that C/ebpα may act as a pioneer transcription factor in cooperation with other transcription factors, such as Pu.1 or Runx1, to prime the myeloid gene expression program at very early stages of hematopoiesis [33].

Wolfler and colleagues generated a *Cebpa*-Cre/YFP reporter mouse and reported that C/ebpα is expressed in less than 10% of the $cKIT^+CD150^+$ defined LT-HSCs population in the bone marrow [35]. These findings are in line with recent published data from the groups of Ido Amit and from Paul Lee Grimes, who conducted single cell RNA sequencing and found that a small proportion of bone marrow HSPCs expresses *Cebpa [36, 37]*.

In addition to the block in myeloid differentiation, the conditional *Cebpa* knockout mouse model by Tenen and colleagues shows that LT-HSCs exit from quiescence, increase in cycling and expansion [22, 38]. On the other hand, the conditional *Cebpa* knockout mouse model by Porse and colleagues showed that LT-HSCs undergo exhaustion and increase in apoptosis [34]. These studies may seem conflicting, but these differences are most probably due to the dissimilarity in timing of the analysis of these mice. In parallel with these findings, a severe quantitative loss of LT-HSCs has also been reported in two separate studies in which a myeloid specific (+37kb) *CEBPA*-enhancer was deleted [39, 40]. Therefore, phenotypically, it is very clear that reduced *Cebpa* expression, either by deleting the *Cebpa* gene or the *Cebpa*-enhancer, exerts detrimental effects on the frequency of

LT-HSC population. Interestingly, to investigate whether reduced *Cebpa* levels also influence the function of HSCs, transplantation experiments of *CEBPA*-enhancer deleted bone marrow cells in recipient mice, showed no difference in the frequency of LT-HSCs when compared to controls 19 weeks post-transplantation, which can be explained by residual *Cebpa* levels present after enhancer deletion. In fact, transplantation of bone marrow cells exhibiting complete ablation of *Cebpa* expression from *Cebpa* knockout mice showed 20-fold reduction in LT-HSC numbers at 16 weeks post-transplantation [34]. Altogether these findings suggest that C/EBPα is required to maintain the integrity of HSCs and hence the hematopoietic system.

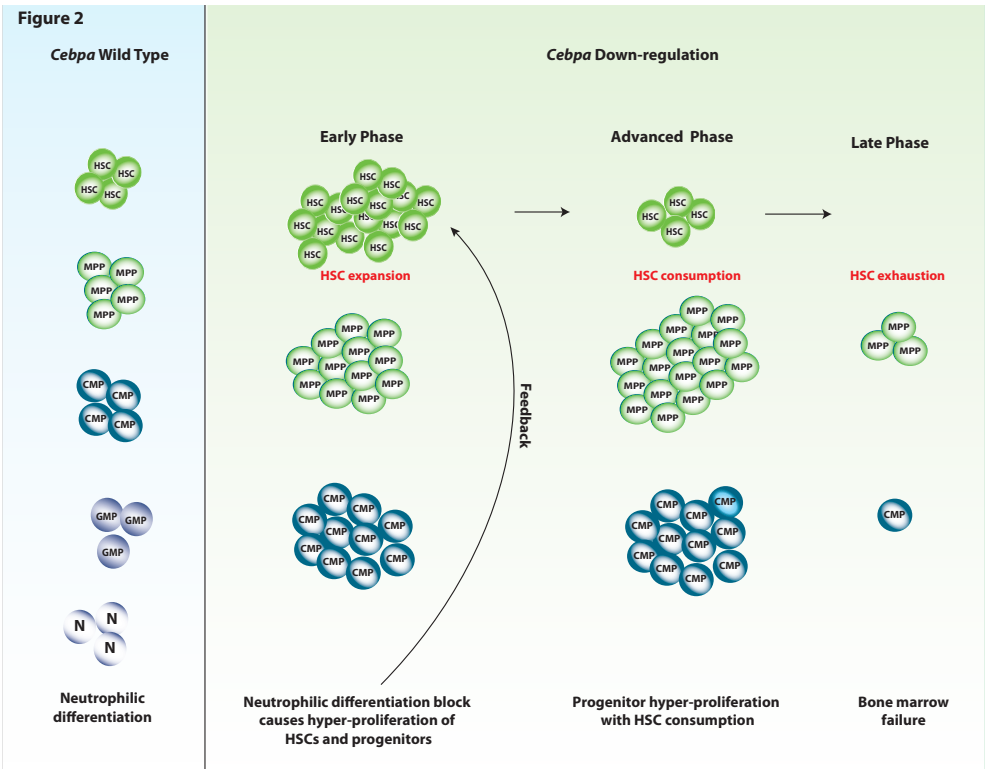


Figure 2: C/EBPα when expressed in early hematopoietic stem and progenitor cells primes and drives myeloid differentiation leading to the production of neutrophils (N). We hypothesize that downregulation of C/EBPα in the bone marrow in mice results in the following phases. **The early phase**: block of neutrophilic differentiation causes a feedback mechanism on hematopoietic stem cells (HSCs) and progenitors to stimulate more differentiation. This results in increased in cycling of (HSCs) and expansion of downstream progenitors (MPP: Multi-Potent Progenitors; CMP: Common Myeloid Progenitors). **Advanced phase**: HSCs are consumed, while progenitors still show expansion. **Late phase**: HSC and progenitor exhaustion leading to bone marrow failure. GMP: Granulocyte Macrophage Progenitor.

Together these results all point towards the importance of *C/*ebpα for the maintenance of the bone marrow HSC integrity, and raise further an important question: why does the loss of *Cebpa* in mice in only a subset of stem/progenitor cells, influences the behavior of the HSC/progenitor compartment, including *Cebpa*-HSC/progenitor cells? Possibly, a feedback mechanism is activated to stimulate the generation of additional myeloid primed progenitors and compensate for the loss of the granulocytic lineage. This would lead to HSC exhaustion caused by an induced HSC response to exit from the quiescence state and undergo cell division to generate sufficient myeloid progenitors (Figure 2). It could also be that the loss of such a critical transcription factor might alter the cell to cell communication between the *Cebpa*$^{+}$ HSPC compartment and the bone marrow microenvironment, which further imposes a global effect on the HSC population[41]. To understand the mechanism of action of the above-proposed ideas, further experimental investigations are required. Knocking out *Cebpa* at a later stage of myeloid commitment, such as in CMPs, would provide a sufficient time-frame to monitor the biological response of the HSCs to compensate for the loss of granulopoiesis in a time dependent manner. Transplantation of *Cebpa* knockout stem/progenitor cells into sub-lethally irradiated wild type recipient mice provides another possibility to study whether these *Cebpa* knockout cells stimulate HSC exhaustion resulting in bone marrow failure. These models should also address the involvement of the microenvironment, and whether stromal and endothelial cells from the transplanted host become susceptible to alterations that ultimately influence the HSC population. Whether the proposed events are also a hallmark in human disease, is yet to be investigated.

C/EBPα in myeloid reprograming and cell fate decisions.

Trans-differentiation studies by Graf and colleagues previously showed that when an estrogen-inducible *C/*EBPα-ER is expressed in a non-myeloid bone marrow cell, acts in synergy with other transcription factors to induce myeloid differentiation. They showed that *C/*EBPα, in collaboration with PU.1 or C/EBPβ, reverses the lymphoid phenotype of B-cell or T-cell progenitors into myeloid progenitors to eventually differentiate into monocytes/granulocytes *in vitro*[42, 43]. Whether the cell commits to one cell type or another, depends on the transcription factor network that is available within the cell under investigation. In the presence of the Yamanaka factors (Oct4, cMYC, KLF4 and SOX2)[44], C/EBPα poises B cells rapidly into induced pluripotent stem cells by first generating GMPs [33, 45-48]. Other studies showed that C/EBPα is also capable to transdifferentiate non-hematopoietic cells into various cell types. For example, fibroblasts can be transformed into myeloid progenitors by C/EBPα in the presence of PU.1 or into adipocytes in the presence of PPARgamma and SREBP-1[49-52].

The repressive role of C/EBPα in hematopoiesis is less understood. Previous studies have shown that C/EBPα induces cell cycle exit coupled with differentiation, by repressing cell cycle related transcription factors including Myc [53] and c-Jun[54, 55]. A recent study

showed that C/EBPα repress the B-cell lineage genes by forming a complex with chromatin modifying proteins including LSD1 and HDAC1 [46]. Such mechanism might also be applied in other cell types since progenitors from *Cebpa* knockout mice exhibit upregulation of T-cell related genes, such as Cd7 and Lck, while C/EBPα retroviral re-introduction in these progenitors attenuated the expression of these genes to drive myeloid differentiation[32].

During cell fate decisions, *Cebpa* is negatively regulated by other non-myeloid factors to exclude the myeloid differentiation program. De Obaldia et. al reported that in mouse bone marrow, C/ebpα expression is repressed by the Notch1-target Hes1. Interestingly, they showed that T-cell development in Hes1 deficient progenitors was restored upon C/ebpα deletion , indicating that C/ebpα acts as a main repressor of T-cell development[56]. Moreover, Rothenberg and colleagues have shown that *Cebpa* expression in the thymus is lost during T-cell commitment by an extensive increase of a repressive histone mark deposition in the *Cebpa* locus[57].

In summary, C/ebpα is a potent differentiation mediator in different cell types, especially for the myeloid lineage. Its importance in activating the myeloid program is indispensable to generate GMPs from any cell type, which occurs only in the presence of other pioneer myeloid factors such as Pu.1. In addition, *Cebpa* is one of the primary targets to be shut down by other lineage specific transcription factors in order to exclude myelopoiesis, emphasizing its important role as a granulocytic differentiation transcription factor.

C/EBPα in AML initiation and development.

It has long been observed that AML is initiated in myeloid committed progenitors. [58, 59] Cozzio and colleagues, demonstrated that murine CMPs and GMPs that were transduced with MLL-ENL fusion gene acquired leukemogenic potential *in vitro* and *in vivo*[60]. Based on the myeloid progenitor potential of generating *in vitro* serial replating and serial transplantation with minimal cell numbers *in vivo*, these cells were termed leukemic stem cells (LSCs). Such LSCs with a strong leukemogenic potential were also later confirmed when the MLL-AF9 fusion oncoprotein was retrovirally transduced in mouse GMPs[61]. Interestingly, such leukemic GMPs shared identical gene expression profiles with wild type GMPs, but acquired a self-renewal gene expression program to propagate *in vitro* and *in vivo*[61]. Leukemic blasts from *Cebpa* biallelic mutant knock-in mice share an identical gene expression program as these *MLL-AF9* transduced leukemic GMPs [62]. Collectively, these findings indicate that a myeloid differentiation program is required to initiate myeloid leukemia.

Moreover, such findings raise paradoxical questions whether the expression of C/EBPα, which may act as a tumor suppressor protein, is also required for the initiation of leukemia. Gene expression profiles of AML patient cells showed that *CEBPA* is sufficiently expressed at the mRNA level in all AML subtypes, except for one subgroup with a myeloid/T-lymphoid immunophenotype [32]. Although this needs confirmation at the protein level, such observations indicate that AML requires an adequate degree of differentiation by involving

C/EBPα, at least at early stages of leukemic initiation. Bo Porse and colleagues used the MLL-ENL and MLL-AF9 mouse model to test this hypothesis. Elegantly, they compared the degree of leukemogenesis exhibited by these two fusion oncoproteins in a *Cebpa* wild type and *Cebpa* knockout background. The latter failed to generate any leukemia both *in vitro* and also *in vivo*[34]. They confirmed that once the leukemia was propagated in the *Cebpa* wild type mice, it did not matter anymore whether *Cebpa* was expressed or not, thus confirming the role of C/EBPα in AML initiation as a differentiation mediator, but not for AML maintenance.

Using similar mouse models, a recent study by Ye et al. confirms these findings. In addition, they show that by administering the growth factors IL-3 or GM-CSF, the MLL fusion oncoproteins induce myeloid leukaemia even in the absence of C/ebpα [63]. Thus, a divergent pathway may take over the C/ebpα pathway when enhanced signalling is introduced by the ectopic addition of myeloid growth factors, even when C/ebpα is not expressed. Possibly, this divergence is via a redundant pathway that takes place in cases of 'emergency', and the main factor known to be responsible for similar events in granulopoiesis is the other CEBP family member C/ebpβ [64]. Thus, these findings wrap up the whole concept that leukaemia develops only if sufficient degree of differentiation is present, with C/EBPα being one of the important factors for AML initiation.

MUTANT *CEBPA* AS A DRIVER OF AML

CEBPA mutations in AML: Biological function and clinical implications.

The strong myeloid phenotype observed in *Cebpa* knockout mice prompted the Tenen group to investigate AML patients for mutations in the open reading frame (ORF) of *CEBPA* [34, 65-70]. Mutations in *CEBPA* ORF in human AML occur in approximately 7-15% of the cases [34, 65-70]. Of these *CEBPA* mutant AMLs, 30% exhibit mutations on one allele, termed here as *CEBPA* single mutants (*CEBPA*^{sm}), with the majority occurring at the N-terminus as out-of frame mutations. The other 70% of *CEBPA* mutant AMLs have both alleles affected and are usually termed as *CEBPA* double mutant (*CEBPA*$^{dm)}$ AMLs (Figure 1). Usually, one allele carries an out of frame N-terminal mutation and the other carries an in-frame C-terminal mutation (referred to as N/C mutant) [65]. Very rare combinations of biallelic mutations such as N/N or C/C also occur in human AMLs, and may either be derived from same C- or N- mutations in the different alleles, or they resulted from mitotic recombination of the q arm of chromosome 19 [71, 72].

The N-terminal out-of-frame mutations induce a stop codon after the first ATG site leading to translation initiation from the 3rd ATG site, generating the shorter p30 isoform in excess, which has been reported to act as a dominant negative of the p42 isoform [34, 65-70]. An imbalance in the p42:p30 ratio caused by N-terminal mutations is associated with an increased proliferation and minimal differentiation of myeloid progenitor cells [73, 74].

The C-terminal in-frame mutations on the other allele cause a defect in the bZIP domain. However, the C-terminal mutations still generate a balanced ratio of p42:p30 isoforms. Depending on the position of the mutation at the bZIP domain, the C/EBPα p42 is defective either in (1) the DNA binding domain, which disables the DNA binding property of C/EBPα or (2) at the leucine zipper domain, which interferes with the formation of homo- and heterodimers [75-77]. This implies that *CEBPA^{dm}* cases lack the wild type p42 isoform, thus translating only defective C/EBPα isoforms which support the onset of leukemia (see next section).

CEBPA^{dm} AMLs are among the patients with a favourable prognosis [34, 65-70]. *CEBPA* mutation analysis is now routinely carried out in AML at diagnosis in many academic centres and applied as a prognostic marker. Single *CEBPA* mutation has no predictive value on treatment response or survival in AML. Treatment of choice and outcome for *CEBPA^{sm}* AML is determined by mutations in other genes, such as in NPM1 and/or FLT3 [65, 78, 79].

CEBPA^{dm} but not CEBPA^{sm} drives the onset of AML in human and in mouse.

All *CEBPA^{dm}* AMLs carry a very similar gene expression profile that distinguishes them from other AML subtypes, including *CEBPA^{sm}* leukemias [65, 67, 80]. In fact, a gene expression signature of approximately 20 genes only has been defined which predicts the presence of a *CEBPA^{dm}* in AML. These results demonstrate that *CEBPA^{dm}* AML should be considered as a unique AML subtype, which is biologically different from *CEBPA^{sm}* leukemias.

Mutations in the ORF deregulate the function of C/EBPα, but are they sufficient to drive leukemogenesis? Nerlov and colleagues addressed this question by generating mutant mice that carried either an N-terminal or a C-terminal mutation in *Cebpa* [73, 81]. Single mutant animals did not develop AML, suggesting that other mutations are indeed required. As mentioned, *CEBPA^{sm}* AMLs frequently harbour mutations in genes such as *NPM1, FLT3* or *ASXL1*[78, 79]. Thus these mutant genes seem to collaborate with single mutations in *CEBPA* (mostly N-terminal) to drive AML. An additional line of evidence to this statement is observed in families with germline N-terminal *CEBPA^{sm}*. These individuals only develop AML if a second mutation is acquired on the other allele (usually as a C-terminal mutant), or else an additional mutation occurs in other AML-related genes [82]. This is in line with the mouse reports by Nerlov and colleagues, demonstrating that AML arises upon transplantation of foetal liver cells carrying combinations of knock-in *CEBPA^{dm}* mutations, i.e. N/N, C/N or C/C mutations, with the C/N combination driving the most aggressive form of AML[73, 81]. A latency of 9-14 months for these mice to develop AML suggests that other cooperative mutations are acquired in time, which might decrease the latency of AML in these mice when introduced.

Gene mutations of high interest, frequently observed in human *CEBPA^{dm}* AML, are the ones that occur in the transcription factor GATA2. In approximately 30-40% of *CEBPA^{dm}* AMLs mutations occur in *GATA2*, most frequently in the DNA binding zinc-finger domain [83, 84].

In addition, two recent studies also presented for the first time mutations co-occurring in the *CSF3R* gene with *CEBPA*^{dm} AML[85, 86]. The generation of models by cross breeding *Cebpa* C/N with *Gata2* zinc finger mutant mice or with *Csf3R* mutant mice, would aid in our understanding of the biological role of these mutations in *CEBPA*^{dm} driven human leukemia development.

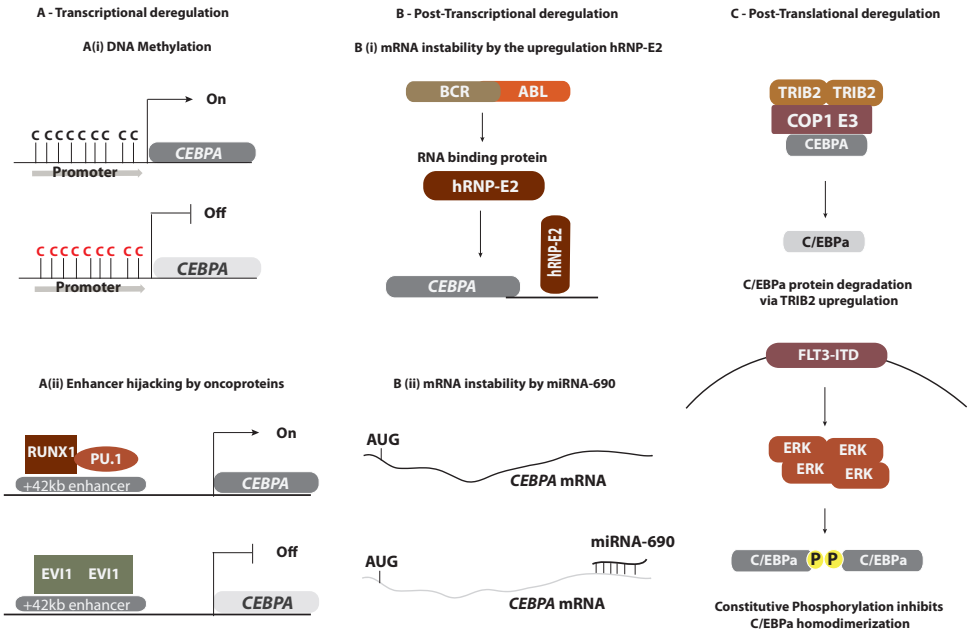


Figure 3: Different mechanisms of C/EBPα deregulation in myeloid malignancies. (A) Transcriptional deregulation: (i) Promoter DNA methylation silences *CEBPA* ; (ii) Oncogenic transcription regulators (e.g. AML1-ETO or EVI1) bind to *CEBPA*-enhancer and downregulate mRNA expression; (B) Post-transcriptional deregulation : (i) The BCR-ABL oncoprotein upregulates an RNA binding protein that destabilizes *CEBPA* transcripts; (ii) upregulation of miRNA-690 destabilizes *CEBPA* transcripts and alters granulocytic development; (C) Post-translational deregulation : (i) Increase in tribbles pseudokinase 2 (TRIB2) levels degrades *CEBPA* protein by the recruitment of COP1 ubiquitin ligase[118]; (ii) Internal tandem duplication in the FLT3 receptor constitutively activates the signalling molecule ERK which inhibits homo-dimerization of C/EBPα and hence interferes with its function by phosphorylation on serine 21.

Transcriptional control of *CEBPA*

Aberrant mechanisms deregulating *CEBPA* function or expression occur at the transcriptional, post-transcriptional, translational and post-translational level in myeloid malignancies, as shown in Figure 3 [87-95]. Studies reported transcriptional deregulation of *CEBPA* either by DNA methylation or by the recruitment of transcriptional repressor complexes to regulatory elements in the *CEBPA* locus [96]. Post-transcriptional deregulation of *CEBPA* was shown by the BCR-ABL fusion oncoprotein and by micro-RNA-690 [93, 97]. In both scenarios, the *CEBPA* transcripts are more susceptible to instability. Protein degradation mediated by Tribbles2 and altered phosphorylation through increased FLT3-ITD signaling,

are examples of aberrant post-translational control in leukemia [98, 99]. These mechanisms have previously been reviewed in detail by others [2, 100], thus, in this review, we focus more on the emerging roles of chromatin configuration and enhancers of *CEBPA* in myeloid development and disease.

The topology of the CEBPA locus in transcription regulation.

Studies are revealing the role of enhancers on *CEBPA* transcriptional output during neutrophilic maturation in the bone marrow[40, 101] and how they are susceptible for deregulation by oncogenic mechanisms in AML. *CEBPA* is located in a topological associated domain (TAD) of 170kb on the long arm of chromosome 19, with borders demarcated by the architectural protein CTCF[39, 102, 103]. These borders restrict interactions to occur between genes and their corresponding enhancers within the TAD. Genes that occur within the same TAD are usually co-regulated [104, 105]. In fact, the *CEBPA*-TAD includes also another CEBP family member *CEBPG,* which is located 5' of *CEBPA*. Both genes are located at close proximity in the nuclear three dimensional space and form loop interactions with other 14 potential enhancers located 5'and 3' of *CEBPA*. The enhancers within the *CEBPA* TAD are active in different *CEBPA* expressing tissues in combinatorial patterns and in a cell type specific manner. In the bone marrow, myeloid primed progenitors and mature myeloid cells exhibit a combinatorial pattern of 8 active enhancers from the 14 found in the whole locus. Of these eight enhancers, one located at +42kb has been studied extensively [39, 40, 101, 106, 107]. This enhancer is highly conserved between human and mice (located at +37kb in mice) and is active in the fetal liver and the dorsal aorta during early fetal development [108]. At later developmental stages and adult life, it's main function is to prime the myeloid gene expression program for neutrophilic differentiation by forming chromatin complexes, involving hematopoietic stem cell transcription factors such as PU.1, ERG, RUNX1 and C/EBPα itself to modulate *CEBPA* expression. In two separate studies, deletion of this enhancer in mice downregulated the expression of *Cebpa,* which ultimately reduced GMP formation and neutrophilic differentiation. The bone marrow cells lacking the +37 kb enhancer showed characteristics of myeloid transformation, i.e. they could indefinitely be re-plated with no evidence of neutrophil development in vitro, similar to C/ebpα knock out bone marrow cells. Other enhancers seem to be important for neutrophil development as well, but at later stages of maturation. Multiple enhancer knock-outs of these enhancers, either alone or in combination should shed light on their function in regulation of *Cebpa* expression and myeloid development.

Hijacking of the locus by potential once-proteins

Given that the +37kb enhancer acts autonomously to regulate *Cebpa* expression in early HSPCs and later in CMPs and GMPs, it is possible that changes in enhancer function, possibly in combination with the other enhancers, might have potential leukemiogenic implications in humans. Genome wide investigations of oncogenic transcription factor binding show that

leukemia associated onco-proteins physically interact with several genomic loci to deregulate the expression of proto-oncogenes and tumor suppressor genes [109]. The +42kb enhancer (human +37 homologue) is a common target for the fusion onco-proteins known as *RUNX1-RUNX1T1* (previous called AML1-ETO) and CBFB-MYH11 (unpublished data). Similarly, EVI1 also called MECOM or PRDM3, binds the +42Kb enhancer in human EVI1 transformed leukemias (personal observations). Interestingly, *CEBPA* expression in these subtypes of AMLs is often downregulated, and the direct binding of these onco-proteins to the enhancer could explain this effect. In fact, Perkins and colleagues demonstrated that Evi1 binding to the murine +37Kb enhancer is associated with a reduction of *Cebpa* expression in Evi1 expressing mouse myeloid leukemia lines[94]. Potentially, the onco-proteins may alter active chromatin states at the enhancer(s) leading to inactivation of *CEBPA*. Whether this mechanism is also the case in human AML, needs further investigations.

The onco-protein RUNX1-RUNX1T1 has been for long associated with negative regulation of *CEBPA* expression in AML [110-113], but the mechanism underlying this interesting finding is poorly understood. RUNX1-RUNX1T1 is known to recruit histone deacetylases (HDACs) via the nervy homologues of RUNX1T1 and binds to DNA on promoters and enhancers via the DNA-binding domain of RUNX1. Since RUNX1-RUNX1T1 binds the +42kb enhancer at RUNX1 motifs as demonstrated in cell lines as well as in patient samples[112], we hypothesize that RUNX1-RUNX1T1 recruits HDACs causing deacetylation and disturbs the enhancer-promoter interaction leading to the downregulation of *CEBPA* expression.

Unexplained low CEBPA levels in AML: do mutations or nucleotide variants in the regulatory domains play a role?

Low *CEBPA* expression without any evidence of known underlying AML-related abnormalities accounts for 10-30% of AMLs. Unknown onco-proteins that may deregulate *CEBPA* expression cannot be excluded. However, other possible oncogenic mechanisms such as deletions or mutations causing alterations in transcription factor binding sites should be further investigated. The presence of single nucleotide variants occurring in transcription factor consensus sequences within enhancers might be another reasonable explanation for decreased *CEBPA* levels. Genome-wide association studies showed several events of SNVs (single nucleotide variants) occurring at enhancers of disease-causing genes [114, 115]. Using the knowledge revealed so far about the role of SNVs or mutations in gene deregulation and disease susceptibility, it could be hypothesized that such alterations may be causative for low *CEBPA* expression in these *CEBPA*^{low} AML subgroups. The application of DNA custom captured sequencing in large AML cohorts would be a potential tool to reveal mutations or SNVs that may be occurring in the *CEBPA* locus. Such studies could aid to our understanding of deregulated expression of *CEBPA* in myeloid malignancies. These potential mutations or SNVs could be used as susceptibility markers for myeloid-related disorders in the normal population as well as in pre-leukemic conditions.

Concluding remarks and future perspectives

This review provides a comprehensive perspective about the indispensable role of C/EBPα in neutrophilic differentiation. The emerging dual function of C/EBPα in AML received particular attention in this review; adequate *CEBPA* levels are required for the initiation of AML, whereas mutations in its ORF or deregulated expression, are key for the leukemic state. The role of C/EBPα in hematopoietic stem cell and progenitor cell biology is still not well understood. This is largely masked by the heterogeneity of bone marrow samples obtained from mouse models or patients, which makes it a challenge to define clearly its function within the HSPC population. This hurdle can be overcome by single cell RNA-seq, (in combination with *in vitro* and *in vivo* functional studies) a widely used application that offers the advantage to study the transcriptome of single HSPCs that express C/EBPα, and hence be able to start answering these questions. In the review we also discussed the potential of bypassing the C/EBPα-differentiation pathway by the administration of growth factors [63, 64] Otherwise, it has been proposed that C/EBPβ could take over and drive neutrophil development in the absence of functional C/EBPα via emergency hematopoiesis [101]. In cases where C/EBPα is expressed but functionally defective, such approach can open a new window for drug administration to induce differentiation in a C/EBPα-independent pathway. Another alternative is to reactivate C/EBPα expression by small molecules [116]. These approaches offer a therapeutic alternative[117] for a wide-range of patients with low *CEBPA* expression that might be predisposed to pre-leukemic conditions and hence, overcome any differentiation-related aberrations in combination with other currently available therapy.

ACKNOWLEDGEMENTS

This work was supported, in whole or in part, by the Dutch cancer foundation KWF and TATA memorial trust foundation for funding this project. The authors thank Dr. Bas Wouters for critically reading the manuscript.

REFERENCES

1. Corces, M.R., et al., *Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution.* Nat Genet, 2016.
2. Friedman, A.D., *C/EBPalpha in normal and malignant myelopoiesis.* Int J Hematol, 2015. **101**(4): p. 330-41.
3. Antonson, P., et al., *A novel human CCAAT/enhancer binding protein gene, C/EBPepsilon, is expressed in cells of lymphoid and myeloid lineages and is localized on chromosome 14q11.2 close to the T-cell receptor alpha/delta locus.* Genomics, 1996. **35**(1): p. 30-8.
4. Wood, S., M. Schertzer, and M.L. Yaremko, *Sequence identity locates CEBPD and FGFR1 to mapped human loci within proximal 8p.* Cytogenet Cell Genet, 1995. **70**(3-4): p. 188-91.
5. Jenkins, N.A., et al., *Mouse chromosomal location of the CCAAT/enhancer binding proteins C/EBP beta (Cebpb), C/EBP delta (Cebpd), and CRP1 (Cebpe).* Genomics, 1995. **28**(2): p. 333-6.
6. Ramji, D.P., et al., *The two C/EBP isoforms, IL-6DBP/NF-IL6 and C/EBP delta/NF-IL6 beta, are induced by IL-6 to promote acute phase gene transcription via different mechanisms.* Nucleic Acids Res, 1993. **21**(2): p. 289-94.
7. Hendricks-Taylor, L.R., et al., *The CCAAT/enhancer binding protein (C/EBP alpha) gene (CEBPA) maps to human chromosome 19q13.1 and the related nuclear factor NF-IL6 (C/EBP beta) gene (CEBPB) maps to human chromosome 20q13.1.* Genomics, 1992. **14**(1): p. 12-7.
8. Williams, S.C., C.A. Cantwell, and P.F. Johnson, *A family of C/EBP-related proteins capable of forming covalently linked leucine zipper dimers in vitro.* Genes Dev, 1991. **5**(9): p. 1553-67.
9. Nishizawa, M., N. Wakabayashi-Ito, and S. Nagata, *Molecular cloning of cDNA and a chromosomal gene encoding GPE1-BP, a nuclear protein which binds to granulocyte colony-stimulating factor promoter element 1.* FEBS Lett, 1991. **282**(1): p. 95-7.
10. Lum, L.S., et al., *A cloned human CCAAT-box-binding factor stimulates transcription from the human hsp70 promoter.* Mol Cell Biol, 1990. **10**(12): p. 6709-17.
11. Landschulz, W.H., et al., *Isolation of a recombinant copy of the gene encoding C/EBP.* Genes Dev, 1988. **2**(7): p. 786-800.
12. Cao, Z., R.M. Umek, and S.L. McKnight, *Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells.* Genes Dev, 1991. **5**(9): p. 1538-52.
13. Tsukada, J., et al., *The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation.* Cytokine, 2011. **54**(1): p. 6-19.
14. Fan, H.Y., et al., *CCAAT/enhancer-binding proteins (C/EBP)-alpha and -beta are essential for ovulation, luteinization, and the expression of key target genes.* Mol Endocrinol, 2011. **25**(2): p. 253-68.
15. Sugahara, K., et al., *Mice lacking CCAAt/enhancer-binding protein-alpha show hyperproliferation of alveolar type II cells and increased surfactant protein mRNAs.* Cell Tissue Res, 2001. **306**(1): p. 57-63.
16. Tomizawa, M., et al., *Hepatocytes deficient in CCAAT/enhancer binding protein alpha (C/EBP alpha) exhibit both hepatocyte and biliary epithelial cell character.* Biochem Biophys Res Commun, 1998. **249**(1): p. 1-5.
17. Soriano, H.E., et al., *Lack of C/EBP alpha gene expression results in increased DNA synthesis and an increased frequency of immortalization of freshly isolated mice [correction of rat] hepatocytes.* Hepatology, 1998. **27**(2): p. 392-401.

18. Oesterreicher, T.J., et al., *Intestinal maturation in mice lacking CCAAT/enhancer-binding protein alpha (C/EPBalpha).* Biochem J, 1998. **330 (Pt 3)**: p. 1165-71.
19. Zhang, D.E., et al., *Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice.* Proc Natl Acad Sci U S A, 1997. **94**(2): p. 569-74.
20. Flodby, P., et al., *Increased hepatic cell proliferation and lung abnormalities in mice deficient in CCAAT/enhancer binding protein alpha.* J Biol Chem, 1996. **271**(40): p. 24753-60.
21. Wang, N.D., et al., *Impaired energy homeostasis in C/EBP alpha knockout mice.* Science, 1995. **269**(5227): p. 1108-12.
22. Zhang, P., et al., *Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha.* Immunity, 2004. **21**(6): p. 853-63.
23. Friedman, A.D., *Transcriptional control of granulocyte and monocyte development.* Oncogene, 2007. **26**(47): p. 6816-28.
24. Ma, O., et al., *Granulopoiesis requires increased C/EBPalpha compared to monopoiesis, correlated with elevated Cebpa in immature G-CSF receptor versus M-CSF receptor expressing cells.* PLoS One, 2014. **9**(4): p. e95784.
25. Federzoni, E.A., et al., *CEBPA-dependent HK3 and KLF5 expression in primary AML and during AML differentiation.* Sci Rep, 2014. **4**: p. 4261.
26. Zhang, P., et al., *Upregulation of interleukin 6 and granulocyte colony-stimulating factor receptors by transcription factor CCAAT enhancer binding protein alpha (C/EBP alpha) is critical for granulopoiesis.* J Exp Med, 1998. **188**(6): p. 1173-84.
27. Schepers, H., et al., *Reintroduction of C/EBPalpha in leukemic CD34+ stem/progenitor cells impairs self-renewal and partially restores myelopoiesis.* Blood, 2007. **110**(4): p. 1317-25.
28. Cammenga, J., et al., *Induction of C/EBPalpha activity alters gene expression and differentiation of human CD34+ cells.* Blood, 2003. **101**(6): p. 2206-14.
29. Iwasaki, H., et al., *The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages.* Genes Dev, 2006. **20**(21): p. 3010-21.
30. Suh, H.C., et al., *C/EBPalpha determines hematopoietic cell fate in multipotential progenitor cells by inhibiting erythroid differentiation and inducing myeloid differentiation.* Blood, 2006. **107**(11): p. 4308-16.
31. Fukuchi, Y., et al., *Comprehensive analysis of myeloid lineage conversion using mice expressing an inducible form of C/EBP alpha.* EMBO J, 2006. **25**(14): p. 3398-410.
32. Wouters, B.J., et al., *Distinct gene expression profiles of acute myeloid/T-lymphoid leukemia with silenced CEBPA and mutations in NOTCH1.* Blood, 2007. **110**(10): p. 3706-14.
33. van Oevelen, C., et al., *C/EBPalpha Activates Pre-existing and De Novo Macrophage Enhancers during Induced Pre-B Cell Transdifferentiation and Myelopoiesis.* Stem Cell Reports, 2015. **5**(2): p. 232-47.
34. Hasemann, M.S., et al., *C/EBPalpha is required for long-term self-renewal and lineage priming of hematopoietic stem cells and for the maintenance of epigenetic configurations in multipotent progenitors.* PLoS Genet, 2014. **10**(1): p. e1004079.
35. Wolfler, A., et al., *Lineage-instructive function of C/EBPalpha in multipotent hematopoietic cells and early thymic progenitors.* Blood, 2010. **116**(20): p. 4116-25.

36. Olsson, A., et al., *Single-cell analysis of mixed-lineage states leading to a binary cell fate choice.* Nature, 2016. **537**(7622): p. 698-702.
37. Paul, F., et al., *Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors.* Cell, 2015. **163**(7): p. 1663-77.
38. Ye, M., et al., *C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence.* Nat Cell Biol, 2013. **15**(4): p. 385-94.
39. Avellino, R., et al., *An autonomous CEBPA-enhancer specific for myeloid-lineage priming and neutrophilic differentiation.* Blood, 2016. **127**(24): p. 2991-3003.
40. Guo, H., S. Cooper, and A.D. Friedman, *In Vivo Deletion of the Cebpa +37 kb Enhancer Markedly Reduces Cebpa mRNA in Myeloid Progenitors but Not in Non-Hematopoietic Tissues to Impair Granulopoiesis.* PLoS One, 2016. **11**(3): p. e0150809.
41. Hornick, N.I., et al., *AML suppresses hematopoiesis by releasing exosomes that contain microRNAs targeting c-MYB.* Sci Signal, 2016. **9**(444): p. ra88.
42. Di Tullio, A., et al., *CCAAT/enhancer binding protein alpha (C/EBP(alpha))-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation.* Proc Natl Acad Sci U S A, 2011. **108**(41): p. 17016-21.
43. Laiosa, C.V., et al., *Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBP alpha and PU.1 transcription factors.* Immunity, 2006. **25**(5): p. 731-44.
44. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.
45. Di Stefano, B., et al., *C/EBPalpha poises B cells for rapid reprogramming into induced pluripotent stem cells.* Nature, 2014. **506**(7487): p. 235-9.
46. Di Stefano, B., et al., *C/EBPalpha creates elite cells for iPSC reprogramming by upregulating Klf4 and increasing the levels of Lsd1 and Brd4.* Nat Cell Biol, 2016. **18**(4): p. 371-81.
47. Bueno, C., et al., *Reprogramming human B cells into induced pluripotent stem cells and its enhancement by C/EBPalpha.* Leukemia, 2016. **30**(3): p. 674-82.
48. Kallin, E.M., et al., *Tet2 facilitates the derepression of myeloid target genes during CEBPalpha-induced transdifferentiation of pre-B cells.* Mol Cell, 2012. **48**(2): p. 266-76.
49. Di Tullio, A. and T. Graf, *C/EBPalpha bypasses cell cycle-dependency during immune cell transdifferentiation.* Cell Cycle, 2012. **11**(14): p. 2739-46.
50. Feng, R., et al., *PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells.* Proc Natl Acad Sci U S A, 2008. **105**(16): p. 6057-62.
51. Liu, S., et al., *Transdifferentiation of fibroblasts into adipocyte-like cells by chicken adipogenic transcription factors.* Comp Biochem Physiol A Mol Integr Physiol, 2010. **156**(4): p. 502-8.
52. Liu, S., et al., *Oleate induces transdifferentiation of chicken fibroblasts into adipocyte-like cells.* Comp Biochem Physiol A Mol Integr Physiol, 2009. **154**(1): p. 135-41.
53. Johansen, L.M., et al., *c-Myc is a critical target for c/EBPalpha in granulopoiesis.* Mol Cell Biol, 2001. **21**(11): p. 3789-806.
54. Rangatia, J., et al., *Elevated c-Jun expression in acute myeloid leukemias inhibits C/EBPalpha DNA binding via leucine zipper domain interaction.* Oncogene, 2003. **22**(30): p. 4760-4.

55. Rangatia, J., et al., *Downregulation of c-Jun expression by transcription factor C/EBPalpha is critical for granulocytic lineage commitment.* Mol Cell Biol, 2002. **22**(24): p. 8681-94.
56. De Obaldia, M.E., et al., *T cell development requires constraint of the myeloid regulator C/EBP-alpha by the Notch target and transcriptional repressor Hes1.* Nat Immunol, 2013. **14**(12): p. 1277-84.
57. Zhang, J.A., et al., *Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity.* Cell, 2012. **149**(2): p. 467-82.
58. Griffin, J.D. and B. Lowenberg, *Clonogenic cells in acute myeloblastic leukemia.* Blood, 1986. **68**(6): p. 1185-95.
59. Fialkow, P.J., et al., *Acute nonlymphocytic leukemia: heterogeneity of stem cell origin.* Blood, 1981. **57**(6): p. 1068-73.
60. Cozzio, A., et al., *Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors.* Genes Dev, 2003. **17**(24): p. 3029-35.
61. Huntly, B.J., et al., *MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors.* Cancer Cell, 2004. **6**(6): p. 587-96.
62. Hasemann, M.S., et al., *Mutation of C/EBPalpha predisposes to the development of myeloid leukemia in a retroviral insertional mutagenesis screen.* Blood, 2008. **111**(8): p. 4309-21.
63. Ye, M., et al., *Hematopoietic Differentiation Is Required for Initiation of Acute Myeloid Leukemia.* Cell Stem Cell, 2015. **17**(5): p. 611-23.
64. Hirai, H., et al., *C/EBPbeta is required for 'emergency' granulopoiesis.* Nat Immunol, 2006. **7**(7): p. 732-9.
65. Wouters, B.J., et al., *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, 2009. **113**(13): p. 3088-91.
66. Snaddon, J., et al., *Mutations of CEBPA in acute myeloid leukemia FAB types M1 and M2.* Genes Chromosomes Cancer, 2003. **37**(1): p. 72-8.
67. Barjesteh van Waalwijk van Doorn-Khosrovani, S., et al., *Biallelic mutations in the CEBPA gene and low CEBPA expression levels as prognostic markers in intermediate-risk AML.* Hematol J, 2003. **4**(1): p. 31-40.
68. Preudhomme, C., et al., *Favorable prognostic significance of CEBPA mutations in patients with de novo acute myeloid leukemia: a study from the Acute Leukemia French Association (ALFA).* Blood, 2002. **100**(8): p. 2717-23.
69. Gombart, A.F., et al., *Mutations in the gene encoding the transcription factor CCAAT/enhancer binding protein alpha in myelodysplastic syndromes and acute myeloid leukemias.* Blood, 2002. **99**(4): p. 1332-40.
70. Pabst, T., et al., *Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia.* Nat Genet, 2001. **27**(3): p. 263-70.
71. Serrano, E., et al., *Uniparental disomy may be associated with microsatellite instability in acute myeloid leukemia (AML) with a normal karyotype.* Leuk Lymphoma, 2008. **49**(6): p. 1178-83.
72. Raghavan, M., et al., *Segmental uniparental disomy is a commonly acquired genetic event in relapsed acute myeloid leukemia.* Blood, 2008. **112**(3): p. 814-21.
73. Kirstetter, P., et al., *Modeling of C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells.* Cancer Cell, 2008. **13**(4): p. 299-310.
74. Keeshan, K., et al., *Transcription activation function of C/EBPalpha is required for induction of granulocytic differentiation.* Blood, 2003. **102**(4): p. 1267-75.
75. Leroy, H., et al., *CEBPA point mutations in hematological malignancies.* Leukemia, 2005. **19**(3): p. 329-34.

76. Porse, B.T., et al., *Loss of C/EBP alpha cell cycle control increases myeloid progenitor proliferation and transforms the neutrophil granulocyte lineage.* J Exp Med, 2005. **202**(1): p. 85-96.
77. Nerlov, C., *C/EBPalpha mutations in acute myeloid leukaemias.* Nat Rev Cancer, 2004. **4**(5): p. 394-400.
78. Sarojam, S., et al., *Characterization of CEBPA Mutations and Polymorphisms and their Prognostic Relevance in De Novo Acute Myeloid Leukemia Patients.* Asian Pac J Cancer Prev, 2015. **16**(9): p. 3785-92.
79. Fasan, A., et al., *The role of different genetic subtypes of CEBPA mutated AML.* Leukemia, 2014. **28**(4): p. 794-803.
80. Taskesen, E., et al., *Prognostic impact, concurrent genetic mutations, and gene expression features of AML with CEBPA mutations in a cohort of 1182 cytogenetically normal AML patients: further evidence for CEBPA double mutant AML as a distinctive disease entity.* Blood, 2011. **117**(8): p. 2469-75.
81. Bereshchenko, O., et al., *Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in C/EBPalpha mutant AML.* Cancer Cell, 2009. **16**(5): p. 390-400.
82. Tawana, K., et al., *Disease evolution and outcomes in familial AML with germline CEBPA mutations.* Blood, 2015. **126**(10): p. 1214-23.
83. Fasan, A., et al., *GATA2 mutations are frequent in intermediate-risk karyotype AML with biallelic CEBPA mutations and are associated with favorable prognosis.* Leukemia, 2013. **27**(2): p. 482-5.
84. Greif, P.A., et al., *GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia.* Blood, 2012. **120**(2): p. 395-403.
85. Maxson, J.E., et al., *CSF3R mutations have a high degree of overlap with CEBPA mutations in pediatric AML.* Blood, 2016. **127**(24): p. 3094-8.
86. Lavallee, V.P., et al., *Chemo-genomic interrogation of CEBPA mutated AML reveals recurrent CSF3R mutations and subgroup sensitivity to JAK inhibitors.* Blood, 2016. **127**(24): p. 3054-61.
87. Lin, T.C., et al., *CEBPA methylation as a prognostic biomarker in patients with de novo acute myeloid leukemia.* Leukemia, 2011. **25**(1): p. 32-40.
88. Lu, Y., et al., *C/EBPA gene mutation and C/EBPA promoter hypermethylation in acute myeloid leukemia with normal cytogenetics.* Am J Hematol, 2010. **85**(6): p. 426-30.
89. Jost, E., et al., *Aberrant DNA methylation of the transcription factor C/EBPalpha in acute myelogenous leukemia.* Leuk Res, 2009. **33**(3): p. 443-9.
90. Chim, C.S., A.S. Wong, and Y.L. Kwong, *Infrequent hypermethylation of CEBPA promotor in acute myeloid leukaemia.* Br J Haematol, 2002. **119**(4): p. 988-90.
91. Radomska, H.S., et al., *Block of C/EBP alpha function by phosphorylation in acute myeloid leukemia with FLT3 activating mutations.* J Exp Med, 2006. **203**(2): p. 371-81.
92. Tavor, S., et al., *Restoration of C/EBPalpha expression in a BCR-ABL+ cell line induces terminal granulocytic differentiation.* J Biol Chem, 2003. **278**(52): p. 52651-9.
93. Perrotti, D., et al., *BCR-ABL suppresses C/EBPalpha expression through inhibitory action of hnRNP E2.* Nat Genet, 2002. **30**(1): p. 48-58.
94. Wilson, M., et al., *EVI1 Interferes with Myeloid Maturation via Transcriptional Repression of Cebpa, via Binding to Two Far Downstream Regulatory Elements.* J Biol Chem, 2016. **291**(26): p. 13591-607.
95. Helbling, D., et al., *The leukemic fusion gene AML1-MDS1-EVI1 suppresses CEBPA in acute myeloid leukemia by activation of Calreticulin.* Proc Natl Acad Sci U S A, 2004. **101**(36): p. 13312-7.

96. Figueroa, M.E., et al., *Genome-wide epigenetic analysis delineates a biologically distinct immature acute leukemia with myeloid/T-lymphoid features.* Blood, 2009. **113**(12): p. 2795-804.
97. Hegde, V.L., et al., *Distinct microRNA expression profile and targeted biological pathways in functional myeloid-derived suppressor cells induced by Delta9-tetrahydrocannabinol in vivo: regulation of CCAAT/enhancer-binding protein alpha by microRNA-690.* J Biol Chem, 2013. **288**(52): p. 36810-26.
98. Fragliasso, V., et al., *Phosphorylation of serine 21 modulates the proliferation inhibitory more than the differentiation inducing effects of C/EBPalpha in K562 cells.* J Cell Biochem, 2012. **113**(5): p. 1704-13.
99. Zheng, R., et al., *Internal tandem duplication mutation of FLT3 blocks myeloid differentiation through suppression of C/EBPalpha expression.* Blood, 2004. **103**(5): p. 1883-90.
100. Reckzeh, K. and J. Cammenga, *Molecular mechanisms underlying deregulation of C/EBPalpha in acute myeloid leukemia.* Int J Hematol, 2010. **91**(4): p. 557-68.
101. Cooper, S., H. Guo, and A.D. Friedman, *The +37 kb CEBPA-enhancer Is Critical for Cebpa Myeloid Gene Expression and Contains Functional Sites that Bind SCL, GATA2, C/EBPalpha, PU.1, and Additional Ets Factors.* PLoS One, 2015. **10**(5): p. e0126385.
102. Dixon, J.R., et al., *Topological domains in mammalian genomes identified by analysis of chromatin interactions.* Nature, 2012. **485**(7398): p. 376-80.
103. Zuin, J., et al., *Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells.* Proc Natl Acad Sci U S A, 2014. **111**(3): p. 996-1001.
104. Symmons, O., et al., *Functional and topological characteristics of mammalian regulatory domains.* Genome Res, 2014. **24**(3): p. 390-400.
105. Jin, F., et al., *A high-resolution map of the three-dimensional chromatin interactome in human cells.* Nature, 2013. **503**(7475): p. 290-4.
106. Guo, H., O. Ma, and A.D. Friedman, *The Cebpa +37-kb enhancer directs transgene expression to myeloid progenitors and to long-term hematopoietic stem cells.* J Leukoc Biol, 2014. **96**(3): p. 419-26.
107. Guo, H., et al., *Runx1 deletion or dominant inhibition reduces Cebpa transcription via conserved promoter and distal enhancer sites to favor monopoiesis over granulopoiesis.* Blood, 2012. **119**(19): p. 4408-18.
108. Wilson, N.K., et al., *Integrated genome-scale analysis of the transcriptional regulatory landscape in a blood stem/progenitor cell model.* Blood, 2016. **127**(13): p. e12-23.
109. Pabst, T. and B.U. Mueller, *Transcriptional dysregulation during myeloid transformation in AML.* Oncogene, 2007. **26**(47): p. 6829-37.
110. Pabst, T., et al., *AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia.* Nat Med, 2001. **7**(4): p. 444-51.
111. Burel, S.A., et al., *Dichotomy of AML1-ETO functions: growth arrest versus block of differentiation.* Mol Cell Biol, 2001. **21**(16): p. 5577-90.
112. Ptasinska, A., et al., *Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal.* Cell Rep, 2014. **8**(6): p. 1974-88.
113. Ptasinska, A., et al., *Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding.* Leukemia, 2012. **26**(8): p. 1829-41.

114. Waszak, S.M., et al., *Population Variation and Genetic Control of Modular Chromatin Architecture in Humans.* Cell, 2015. **162**(5): p. 1039-50.
115. Grubert, F., et al., *Genetic Control of Chromatin States in Humans Involves Local and Distal Chromosomal Interactions.* Cell, 2015. **162**(5): p. 1051-65.
116. Koschmieder, S., et al., *CDDO induces granulocytic differentiation of myeloid leukemic blasts through translational up-regulation of p42 CCAAT enhancer binding protein alpha.* Blood, 2007. **110**(10): p. 3695-705.
117. Reebye, V., et al., *Gene activation of CEBPA using saRNA: preclinical studies of the first in human saRNA drug candidate for liver cancer.* Oncogene, 2018.
118. Keeshan, K., et al., *Transformation by Tribbles homolog 2 (Trib2) requires both the Trib2 kinase domain and COP1 binding.* Blood, 2010. **116**(23): p. 4948-57.

3

THESIS_Roberto_Avellino.indd 58 07-05-18 10:17

AN AUTONOMOUS *CEBPA*-ENHANCER FOR MYELOID-LINEAGE PRIMING AND NEUTROPHILIC DIFFERENTIATION

Roberto Avellino[1], Marije Havermans[1], Claudia Erpelinck[1], Mathijs A. Sanders[1], Remco Hoogenboezem[1], Harmen J. G. van de Werken[2], Elwin Rombouts[1], Kirsten van Lom[1], Paulina M.H., van Strien[1], Claudia Gebhard[3,4], Michael Rehli[3,4], John Pimanda[5], Dominik Beck[5], Stefan Erkeland[1,6], Thijs Kuiken[7], Hans de Looper[1], Stefan Gröschel[1,8], Ivo Touw[1], Eric Bindels[1], and Ruud Delwel[1]

[1] Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands.
[2] Cancer Computational Biology Center, Erasmus University Medical Center, Rotterdam, The Netherlands.
[3] Department of Internal Medicine III, University Hospital Regensburg, Regensburg, Germany.
[4] Regensburg Centre for Interventional Immunology (RCI), Regensburg, Germany
[5] Prince of Wales Clinical Schools and Lowy Cancer Research Centre, UNSW, Sydney, Australia.
[6] Department of Immunology, Erasmus University Medical Center, Rotterdam, The Netherlands.
[7] Department of Viroscience, Erasmus University Medical Center, Rotterdam, The Netherlands.
[8] Department of Translational Oncology, National Center for Tumor Diseases (NCT), Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany.
[9] Department of Cell Biology, Erasmus University Medical Center, Rotterdam, The Netherlands.
[10] Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands.[2]

Blood 2016

Running title: Myeloid specific enhancer for CEBPA

THESIS_Roberto_Avellino.indd 59 07-05-18 10:17

ABSTRACT

Neutrophilic differentiation is dependent on *CCAAT* enhancer binding protein alpha (C/EBPα), a transcription factor (TF) expressed in multiple organs including the bone marrow. Using functional genomic technologies in combination with CRISPR/Cas9 genome editing and in vivo mouse modeling, we show that *CEBPA* is located in a 170kb topological associated domain that contains 14 potential enhancers. Of these, one enhancer located +42kb from *CEBPA* is active and engages with the *CEBPA*-promoter in myeloid cells only. Germ-line deletion of the homologous enhancer in mice in vivo reduces *Cebpa* levels exclusively in hematopoietic stem cells and myeloid-primed progenitor cells leading to severe defects in the granulocytic lineage, without affecting any other *Cebpa*-expressing organ. The enhancer-deleted progenitors cells lose their myeloid transcription program and are blocked in differentiation. Deletion of the enhancer also causes loss of hematopoietic stem cell maintenance. We conclude that a single +42kb enhancer is essential for *CEBPA* expression in myeloid cells only.

INTRODUCTION

All cell types in the bone marrow are derived from a pool of hematopoietic stem and progenitor cells (HSPCs) that sustain blood cell development throughout the life of an organism. Prior to lineage commitment and differentiation, HSPCs undergo chromatin changes brought about by lineage-specific transcription factors (LTFs) to prime and activate lineage-specific gene expression programs[1]. Priming of cell lineages involves the accessibility and activity of cell type-specific enhancers by LTFs to regulate the expression of genes responsible for any given cell lineage[2,3,4].

Cell lineage priming occurs during cell fate decisions which is mainly dependent on the concentration or dosage of LTFs[5,30,49]. For instance, lymphoid-primed progenitors (LMPPs) have high concentrations of lymphoid related LTFs such as IKAROS, E47 and EBF that bind and activate lymphoid specific enhancers to induce lymphoid development[53]. To skew differentiation towards myelopoiesis, these factors become negatively regulated upon increased dosage of the inhibitors of differentiation (ID) TFs, in order to favour increased PU.1 levels and promote myeloid commitment[50]. The leucine zipper transcription factor CCAAT enhancer binding protein alpha, C/EBPα, instructs myeloid differentiation via the priming and activation of myeloid-associated genes in HSPCs[43] and competes for genomic occupancy with other TFs, such as PU.1 and GATA2 in the myeloid-erythroid progenitor compartment, to favour neutrophilic differentiation over monocytic, erythroid and megakaryocytic cell differentiation[6,47]. The important role of C/EBPα in myelopoiesis is substantiated by the diverse oncogenic mechanisms that target C/EBPα levels and function in various subsets of acute myeloid leukemia (AML)[7,8,9,10,11,12]. Moreover, *Cebpa* knock-out mouse models show severe myeloid defects in the bone marrow[13] as well as in several other organs including the liver[14], lung[15], bone tissue[16] as well as in epithelium of the gut[17], implying its broad role as a differentiation TF in several organs. The broad role of C/EBPα as a differentiation mediator in multiple tissues suggests that *CEBPA* is under the control of tissue-specific transcriptional regulatory mechanisms[26]. Transcription regulation occurs in a hierarchical order of multi-step processes that involve the structural organization of the genome to regulate gene expression programs via tissue specific enhancers[18,25,51].

In this study, we investigated how *CEBPA* transcription is regulated during myelopoiesis. We show that the *CEBPA* locus harbours, in total, 14 enhancers and we asked whether *CEBPA* contacts tissue specific enhancers in different *CEBPA*-expressing cell types. Using a combination of functional genomics and CRISPR/Cas9 genome editing in human cell lines and mouse models, we show that the +42kb enhancer acts autonomously in myeloid primed hematopoietic stem cells (HSCs) in the bone marrow to drive adequate *CEBPA* expression levels necessary for full neutrophilic maturation.

MATERIALS AND METHODS

Cell lines and Cell culture

Cell lines were cultured as follows: U937, THP-1, Raji, Jurkat in 90% RPMI and 10% Fetal Calf Serum (FCS); Hep3B, H292 and HepG2 80% RPMI and 20% FCS; HEK293T and HeLa in 90% DMEM and 10% FCS. All cell lines were supplemented with 50 U/ml penicillin and 50 µg/ml streptomycin.

High resolution circularized chromatin conformation capture sequencing (4C-seq)

High resolution 4C-seq was conducted as previously described[55]. In brief, $10x10^6$ cells were crosslinked with 2% formaldehyde for 10 minutes at room temperature. Glycine (0.125M) was added to quench the crosslinking reaction and cells were centrifuged and suspended in lysis buffer to disrupt membranes and isolate chromatin. A primary four-base cutter, either DPNII or NLAIII, was used for digestion, followed by diluted ligations. After precipitation, chromatin was further subjected to a second round of digestions with a different 4 base cutter (either Csp6I or BFAII) and ligated to small-circularized plasmids. Primers for CEBPA viewpoint (Forward: TACTGCTTCTTTACTGCGATC; Reverse; CAAGCAGAAGACGGCATACGA) and for the 21Kkb contact domain viewpoint (Forward: GCCCAGGAGCCTGTGAGATC; Reverse: ACTCTGAGTGCAGAGAGGAG) were designed as previously reported[55]. Primers for 4C-Seq taking the viewpoints at the 5' border of the 170kb topological associated domain (TAD) near *CEBPG* (Forward: TTTTACAAGTCACAGGGATC; Reverse: ACGTCCTCTGTATTGCCTAG) and the 3'border of the TAD, near the promoter of *SLC7A10* (Forward : CCAGCACACACTGCAAGATC Reverse : GGAGGGAGTTCTGTGTGG). Inverse PCR was carried out to amplify sample libraries that were pooled and spiked with 40% PhiX viral genome sequencing library to increase sample diversity. Multiplexed sequencing was performed on the HiSeq2500 platform. 4C-seq data analysis is explained in the Supplementary Methods.

ChIP-seq

ChIP experiments were performed as previously described[58]. Cells were crosslinked at room temperature for 10 minutes with 1% formaldehyde and sonicated to shear the chromatin. Immunoprecipitation of crosslinked chromatin was performed overnight at 4^0C with antibodies directed against the histone mark H3K27ac, the co-activator p300, and TFs including RUNX1, LMO2, PU.1, ERG, TAL1 and SCL, or an equal amount of isotype IgG as background control (http://149.171.101.136/python/BloodChIP/search.html). Descriptions detailing the preparation of library preparation, genome alignment and peak calling are included in the Supplementary Methods.

Flowcytometry and sorting

Flowcytometry and sorting were carried out on the LSRII and the FACSAria IIIU (Becton Dickinson) respectively using the following fluorescent antibodies: CD11B- APC/GR1-FITC/B220-PE/CD45 PerCP CY5/LIN bio cocktail streptavidin-pacific orange/cKIT-APC/SCA1-PB/CD48-FITC/CD150-PE-CY7/CD16-32APC-CY7/CD34-PE. All antibodies were purchased from BD Biosciences and Biologend. Sorted fractions were collected in 500µl PBS with 5% FCS, spun down and re-suspended in 800µl of Trizol and used for RNA-seq.

RNA-seq

Total sample RNA was extracted using Trizol with Genelute LPA (Sigma) as a carrier and SMARTer Ultra Low RNA kit for Illumina Sequencing (Clontech) was used for cDNA synthesis according to the manufacturer's protocol. The cDNA was sheared with the Covaris device and further processed according to the TruSeq RNA Sample Preparation v2 Guide (Illumina). The amplified sample libraries were subjected to paired-end sequencing (2 x 75 bp) and aligned against mm10 using TopHat v2[60,61]. Alignment and processing of RNA-seq data are documented in the Supplementary Methods.

Luciferase reporter assays

The full canonical CEBPA promoter was PCR amplified from gDNA and cloned into the pGL4.11 (Luc2CP) (EcoRV/HindIII) (Promega) luciferase construct. The +9kb or +42kb enhancers were PCR amplified and cloned into pGL4.11 (luc2CP) (Sal1/BamH1) 3' of the luciferase gene in the same construct where the full canonical *CEBPA* promoter was cloned (Supplementary Materials for primer sequences). HEK293T cells were transfected with Lipofectamine 2000 (Invitrogen), U937 electroporated with Lonza KIT (Kit-C), Jurkat, K562, THP-1, HepG2 and H292 cells with X-tremeGENE HP DNA Transfection Reagent (Roche). Cells were harvested after 48 hours, and luciferase activity was measured with the Dual-Luciferase Reporter Assay System (Promega) on a Victor X3 plate reader (Perkin Elmer). All assays were measured in duplicates and performed minimally three times.

CRISPR in human cell lines

CRISPR in human cell lines and in mouse fertilized eggs was carried out as explained in the Supplementary.

RESULTS

CEBPA interacts with multiple intergenic regions, one of which is only prominent in *CEBPA*^{pos} myeloid cells

Promoter-enhancer interactions occur within topological associated domains (TADs) and gene promoters confined within such a domain, might share the same set of enhancers for transcriptional regulation[22,27,28,52]. HI-C sequencing data[24], a comprehensive technique to capture the conformation of genomes[22], shows that *CEBPA* is located in a 170kb conserved TAD (labeled **2** in Fig.1a and Fig.S1a) on chromosome 19. This TAD also contains *CEBPG* and the promoter of *SLC7A10*, located 5' and 3' of *CEBPA* respectively (Fig.1a). To determine the interacting regions within the 170kb TAD (*CEBPA*-TAD) we applied high-resolution 4C-seq[55] taking the *CEBPA*-promoter as a viewpoint (Fig.1b). Eight contact regions, located 5' and 3' of *CEBPA* were found in all cell lines investigated[36], i.e. in *CEBPA*-expressing (*CEBPA*^{pos}) myeloid cell lines HL-60, MOLM-1, U937 and THP-1, the lymphoid *CEBPA*^{neg} cell lines Jurkat and Raji, the lung *CEBPA*^{pos} cell line H292 and the cervical *CEBPA*^{neg} cell line HeLa (Fig.1b and Fig.S1b). These contact domains varied in size between 10 and 22kb (median = 11.65kb) (Fig.S1d). Taking the borders of the *CEBPA*-TAD as viewpoints (Fig.S1c), 4C-seq revealed that the interactions were confined to this TAD. No significant interactions with the adjacent TADs 1 and 3 were found, in line with the binding of the architectural protein CTCF[23,24] to the borders of the *CEBPA*-TAD (labeled **1** and **3** – Fig.1a to 1c).

A region of 21kb prominently contacts the *CEBPA* promoter

We next investigated whether any of the 8 contact regions identified by 4C-seq showed differential promoter interactions in *CEBPA*^{pos} myeloid cell lines compared to *CEBPA*^{neg} cells. A semi-quantitative analysis of 4C-seq data was conducted, by comparing three myeloid *CEBPA*^{pos} (MOLM-1, U937 and HL-60) with three non-myeloid *CEBPA*^{neg} (Jurkat, Raji and HeLa) cell lines. The contact region of approximately 21kb in size located 3' of *CEBPA* (Fig. S1d), showed a more significant interaction (FDR<0.05), in the *CEBPA*^{pos} cell lines compared to *CEBPA*^{neg} cells (Fig.1d). In contrast, no major interaction differences were observed for the other *CEBPA* promoter interacting regions. A reciprocal 4C-Seq experiment using the 21kb contact region as a viewpoint, confirmed that the interaction with the *CEBPA* gene occurred at a higher frequency in *CEBPA*^{pos} myeloid cell lines (FDR<0.05) (Fig.S1e). These findings show that a distant region of 21kb interacts with *CEBPA* prominently in *CEBPA*^{pos} myeloid cells, suggesting a myeloid-specific chromatin conformation at this region.

Two potential enhancers with myeloid preference located within the 21kb-contact region

To identify regions of active chromatin in peripheral blood neutrophils and monocytes we conducted H3K27ac ChIP-seq and compared the determined H3K27ac profiles to those

obtained from public available data derived from other human primary *CEBPA*^{pos} (n=12) and *CEBPA*^{neg} (n=4) tissues (http://www.roadmapepigenomics.org/). The H3K27ac ChIP-seq profiles revealed 14 potential enhancers within the *CEBPA*-TAD, located upstream and downstream of *CEBPA* (Fig.2; Green profile). Each tissue investigated harbored a distinct combination of H3K27ac marked regions, suggesting tissue specificity and differential regulation of *CEBPA* expression(Fig.2; Blue profiles). Tissues which do not express *CEBPA* were all devoid of H3K27ac marked sites within the *CEBPA*-TAD (Fig.2; Red profiles), except for the *CEBPG* promoter[19,20]. Of these 14 potential enhancers, ten located 5' (-9kb, -4kb, -25kb, -47kb and -64kb) and 3' (+9kb, +29kb, +34kb, +42kb, +50kb) of *CEBPA* are found within the eight contact regions identified by 4C-seq (Fig.S2). The +34kb and the +42kb regions were exclusively H3K27ac marked in neutrophils and monocytes. These two regions are located within the 21kb large contact region that showed increased interaction in myeloid cells (Fig.1d). The +9kb region is H3K27ac marked in all the *CEBPA*-expressing tissues investigated, suggesting a tissue-broad role in *CEBPA* regulation (Fig.2a). These findings show that from a total of 14 potential enhancers located within the *CEBPA*-TAD, the +34kb and +42kb regions appear to be myeloid-specific, suggesting the presence of an enhancer-rich chromatin site important for *CEBPA* transcriptional regulation.

The +42kb enhancer is occupied by hematopoietic specific transcription factors in HSPCs

CEBPA is expressed at low levels in CD34+ progenitor cells and increases upon neutrophilic maturation (Fig.3a). The low *CEBPA* expression levels in CD34+ progenitor cells (Fig.3a) correlate with the number of potential enhancers found by H3K27ac ChIP-seq (ENCODE)[29] i.e., only the +42kb and the +9kb potential enhancers are active at this stage of hematopoiesis (Fig.3b). Motif analysis revealed that the +42kb enhancer contains DNA binding motifs (a CTCF-interacting zinc finger transcription factor[40]) and multiple HSPC-related (TFs) (Fig.S3a). In contrast, the +9kb enhancer contains DNA binding motifs corresponding to a universal set of TFs (Fig.S3a). Furthermore, the HSPC-related LTFs and other TFs including LYL11, RUNX1, GATA2, FLI1, ERG and LMO2[21,30] bind the +42kb region in CD34+ cells (Fig.3c). Recruitment of p300 to enhancers is highly suggestive of enhancer activity[31]. ChIP-Seq in the *CEBPA*^{pos} MOLM-1 cell line demonstrated strong binding of the histone acetyltransferase p300 to the +42kb enhancer (Fig.3d). This binding was also demonstrated by ChIP-qPCR in the *CEBPA*^{pos} myeloid cells HL-60, U937 and THP-1 (Fig.3e). No binding was found in the lymphoid cell lines Jurkat and Raji or the *CEBPA*^{pos} non-hematopoietic cell lines H292 (lung) and HeLa (cervical). Together, our data show that the +42kb enhancer is a critical region highly occupied by a HSPC-related TF complex that potentially initiates *CEBPA* expression in CD34+ progenitor cells.

THESIS_Roberto_Avellino.indd 65 07-05-18 10:17

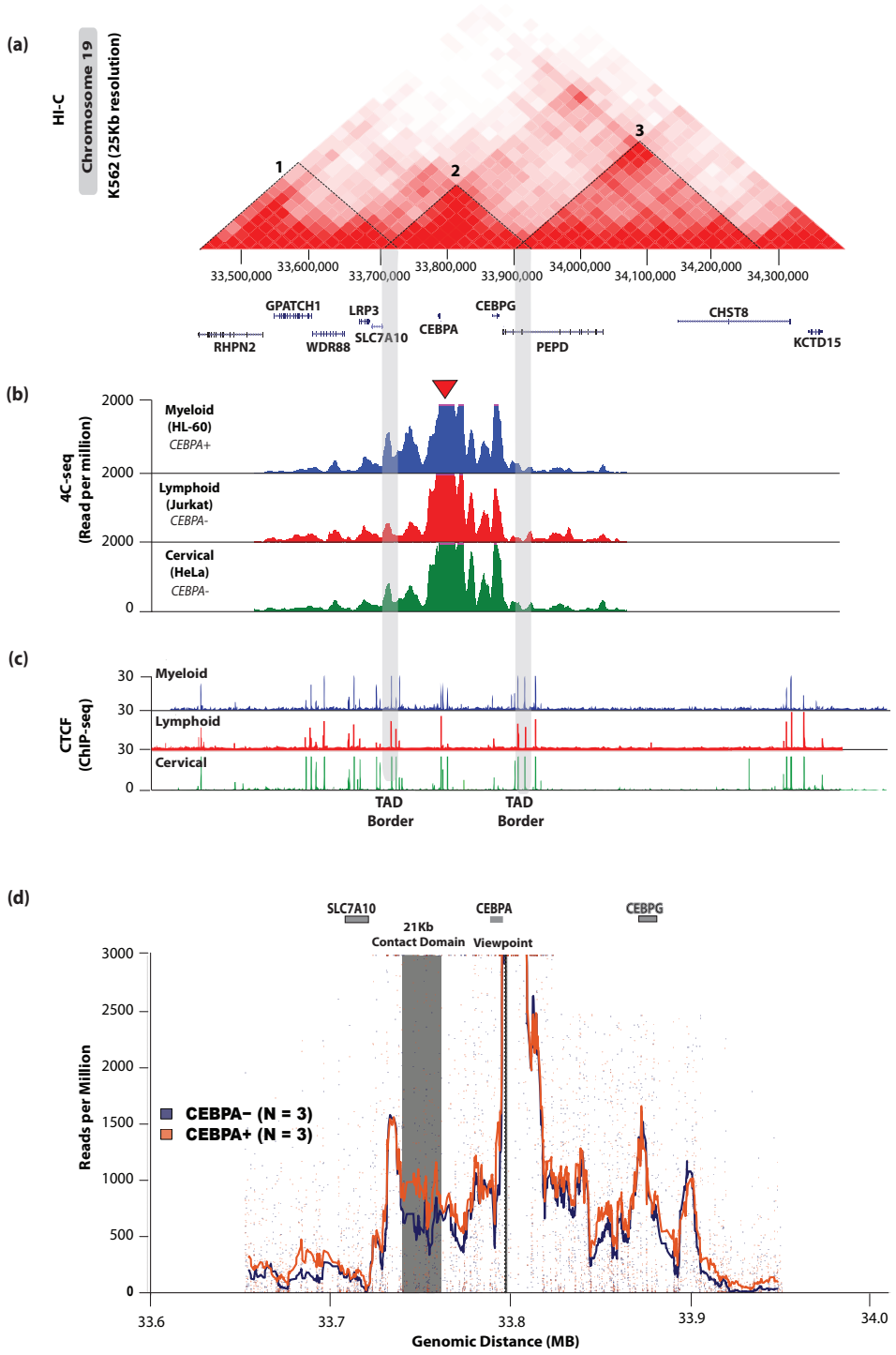

(a)
HI-C
Chromosome 19
K562 (25Kb resolution)
1
2
3
33,500,000
33,600,000
33,700,000
33,800,000
33,900,000
34,000,000
34,100,000
34,200,000
34,300,000
GPATCH1
LRP3
SLC7A10
CEBPA
CEBPG
CHST8
RHPN2
WDR88
PEPD
KCTD15
(b)
4C-seq
(Read per million)
Myeloid
(HL-60)
CEBPA+
Lymphoid
(Jurkat)
CEBPA-
Cervical
(HeLa)
CEBPA-
2000
2000
2000
0
(c)
CTCF
(ChIP-seq)
Myeloid
Lymphoid
Cervical
30
30
30
0
TAD
Border
TAD
Border
(d)
SLC7A10
CEBPA
CEBPG
21Kb
Contact Domain
Viewpoint
Reads per Million
3000
2500
2000
1500
1000
500
0
CEBPA− (N = 3)
CEBPA+ (N = 3)
33.6
33.7
33.8
33.9
34.0
Genomic Distance (MB)

Figure 1: The *CEBPA* promoter contacts multiple intra-TAD genomic sites: Stronger interaction with a 21kb genomic region in *CEBPA* expressing myeloid cells.(a) HI-C heatmap matrix (25kb resolution) in the K562 cell line on chromosome 19 reveals a 170kb *CEBPA* TAD (**2**), which is flanked by TADs **1** and **3**. The *CEBPA* TAD also contains *CEBPG* and part of *SLC7A10*. **(b)** Normalized 4C-seq profiles of myeloid *CEBPA*^{pos} HL-60 (Blue), lymphoid *CEBPA*^{neg} Jurkat (Red) and cervical *CEBPA*^{neg} HeLa (green) cell lines. The viewpoint (red triangle) located at the *CEBPA* promoter shows multiple interacting sites confined to the *CEBPA*-TAD (borders marked in grey). **(c)** CTCF ChIP-seq (ENCODE) in the myeloid HL-60, lymphoid Jurkat and cervical HeLa cell lines shows enrichment at the *CEBPA* TAD borders (grey) which overlap with the HI-C contact-matrix borders separating the *CEBPA*-containing TAD2 from TAD1 and TAD3. **(d)** Quantitative analysis of 4C-seq data to distinguish interacting regions occurring at higher contact frequencies in *CEBPA*^{pos} myeloid cells (orange; n=3) compared to *CEBPA*^{neg} cells (blue; n=3). The *CEBPA* viewpoint is marked with a dotted line. A specific region indicated in gray of around 21kb located 3' of *CEBPA* and with more than 250 reads per million shows a statistically significant higher contact frequency (FDR <0.05) in *CEBPA*^{pos} as compared to *CEBPA*^{neg} cell lines.

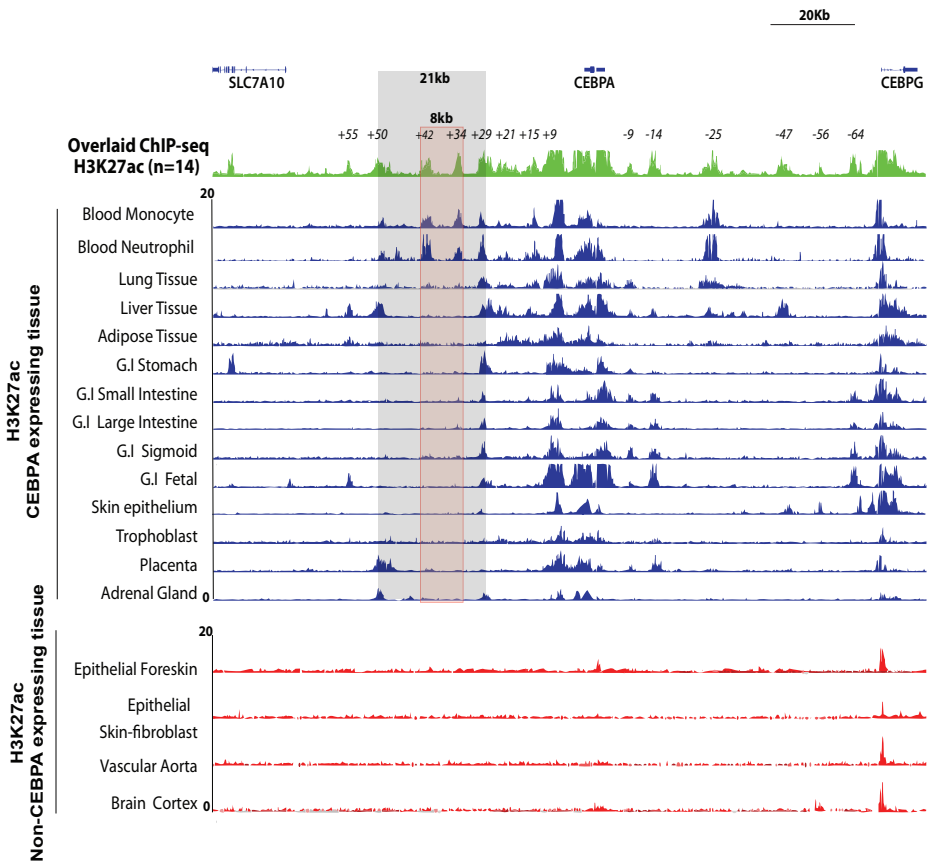


Figure 2: The *CEBPA* TAD exhibits a diverse combination of active enhancers in different *CEBPA*^{pos} tissues. ChIP-seq for H3K27ac conducted in terminally differentiated neutrophils and monocytes (in-house) was compared to publicly available ChIP-seq H3K27ac (www.roadmapepigenomics.org/). Superimposed H3K27ac (top; green) ChIP-seq profiles from 14 different *CEBPA*^{pos} tissue types shows 14 potential enhancers situated within the *CEBPA* TAD at 5' (-9, -14, -25, -47, -56, -64 kb) and at 3' (+9, +15, +21, +29, +34, +42, +50, +55 kb). Each individual *CEBPA*^{pos} tissue type (middle; blue) shows a different combinatorial set of active enhancers. *CEBPA*^{neg} tissue types (bottom; red) do not exhibit H3K27ac at the locus, except at *CEBPG*. An intergenic 8kb hotspot (red) located within the 21kb contact domain (grey), contains two potential enhancers (+34kb and +42kb) that are H3K27ac enriched in neutrophils and monocytes only.

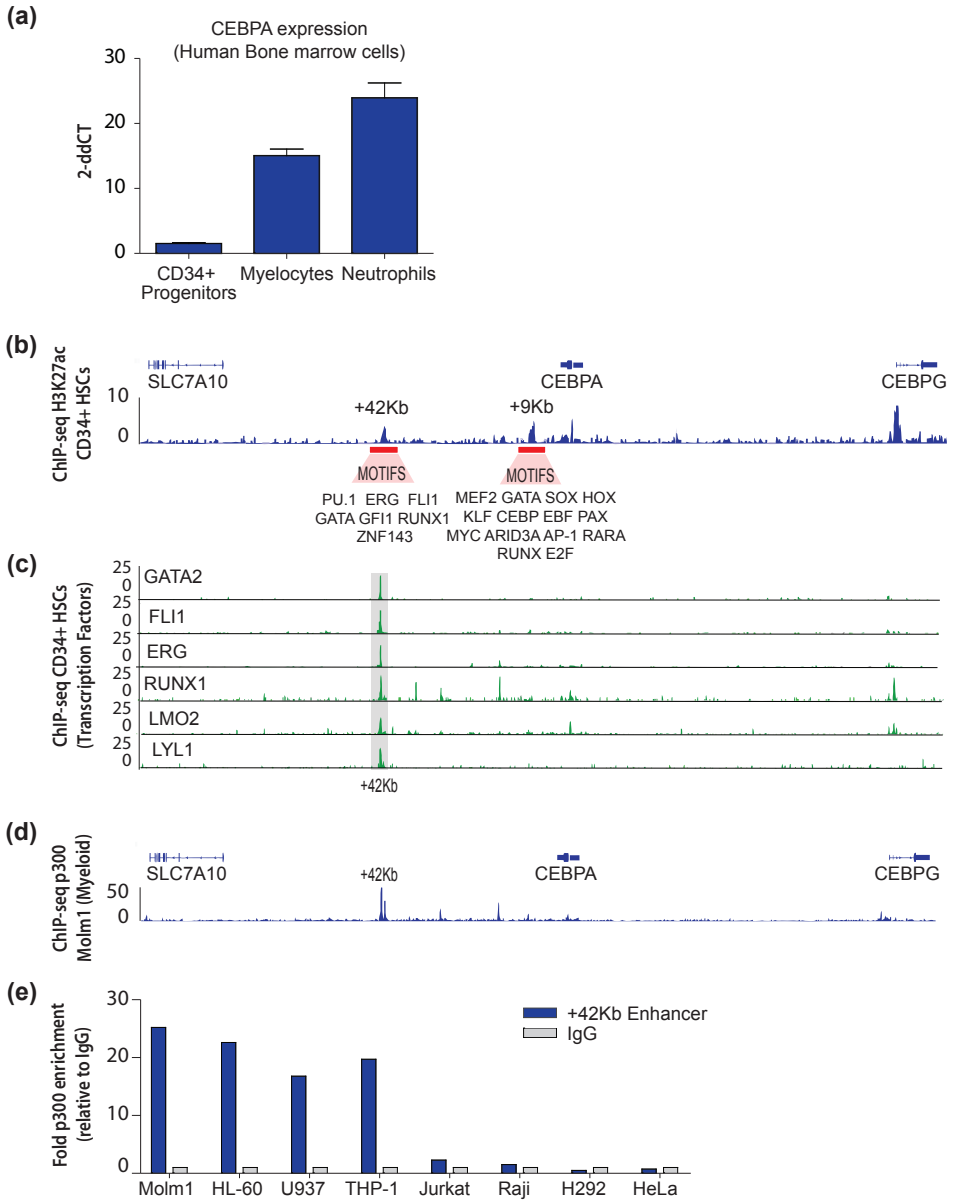


Figure 3: The +42kb region is specifically H3K27ac marked in CD34+ hematopoietic stem cells. (a) *CEBPA* mRNA expression determined by qPCR in FACS sorted populations of normal CD34+ bone marrow cells, metamyelocytes and neutrophils (n=3). (**b**) H3K27ac ChIP-seq in CD34+ cells, obtained from GCSF-mobilized peripheral blood cells, reveals enrichment at the +9kb and +42kb enhancers. Motifs that correspond to specific TF binding sites are depicted underneath each enhancer (For details, see Fig.S3a). (**c**) ChIP-seq for the indicated transcription factors carried out in CD34+ cells, shows specific binding at the +42kb enhancer. (**d**) ChIP-seq for p300 in MOLM-1 *CEBPA*^{pos} cell line MOLM-1, reveals the strongest interaction at +42kb. (**e)** ChIP-qPCR shows p300 enrichment within the +42kb region in the *CEBPA*–expressing cell lines MOLM-1, U937, HL-60, THP-1, but not in the *CEBPA*^{neg} cell lines Jurkat, Raji, H292 and HeLa. Enrichment was calculated as fold change relative to IgG control.

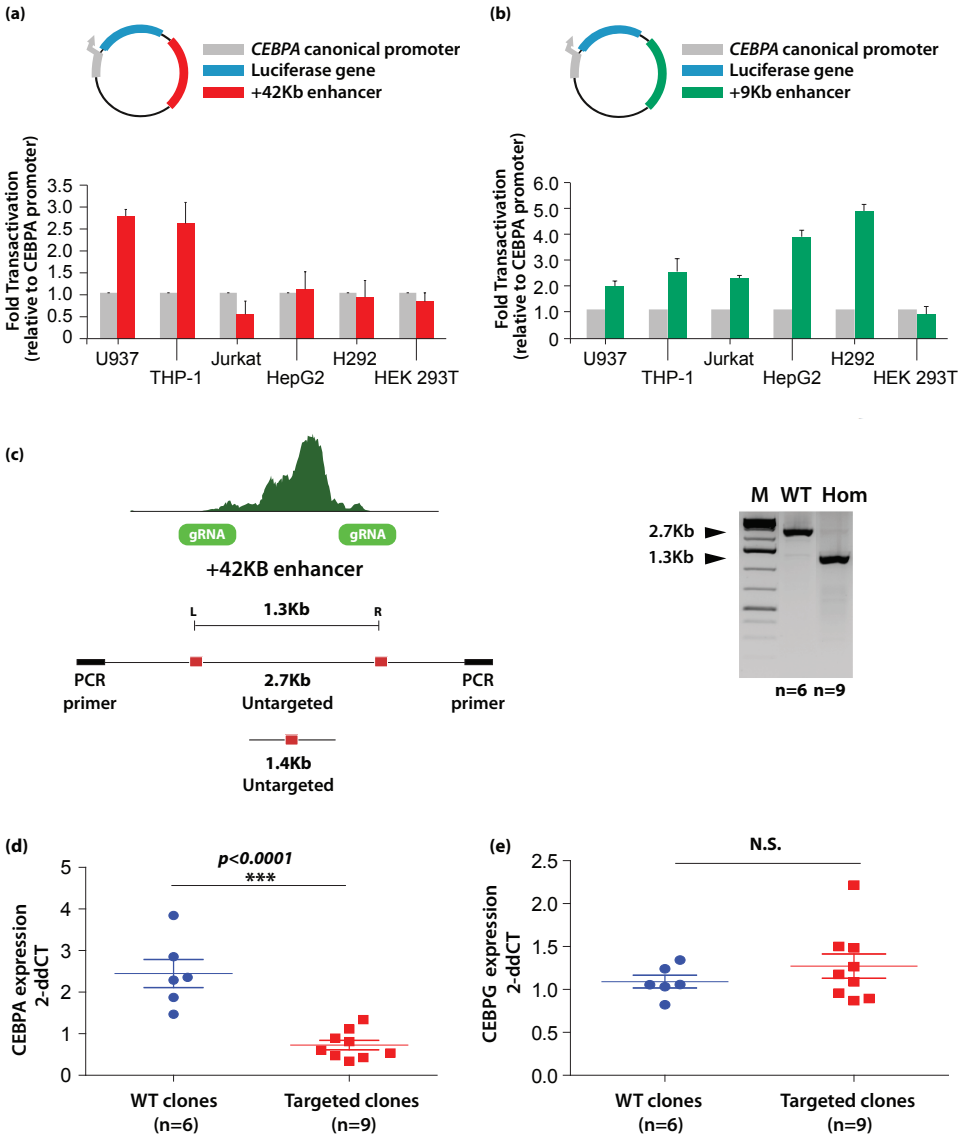


Figure 4: The +42kb enhancer is a myeloid-specific *CEBPA* transcriptional activator. (a,b) The +42kb and +9kb enhancer were cloned 3' of a luciferase reporter gene under the control of the full canonical *CEBPA* promoter. Results are presented as fold change of the +42kb enhancer in combination with the *CEBPA* promoter (blue=myeloid; red=lymphoid; green=*CEBPA*^{pos} non-hematopoietic; orange=*CEBPA*^{neg} non-hematopoietic cell lines) relative to *CEBPA* promoter alone (grey). **(c)** gRNA for the CRISPR/Cas9 system were designed to flank the p300 and TF binding sites within the +42Kb enhancer. Single cell clones were generated and genotyped using a PCR strategy. **(d and e)** Wild type clones (n=6) and homozygous clones (n=9) were selected and qPCR for *CEBPA* mRNA expression and for *CEBPG* was conducted. Statistical significance to compare mRNA expression levels between wild type and homozygous clones for both genes under investigation, was carried out using two-tailed T-test. *** = p<0.0001; N.S = not significant.

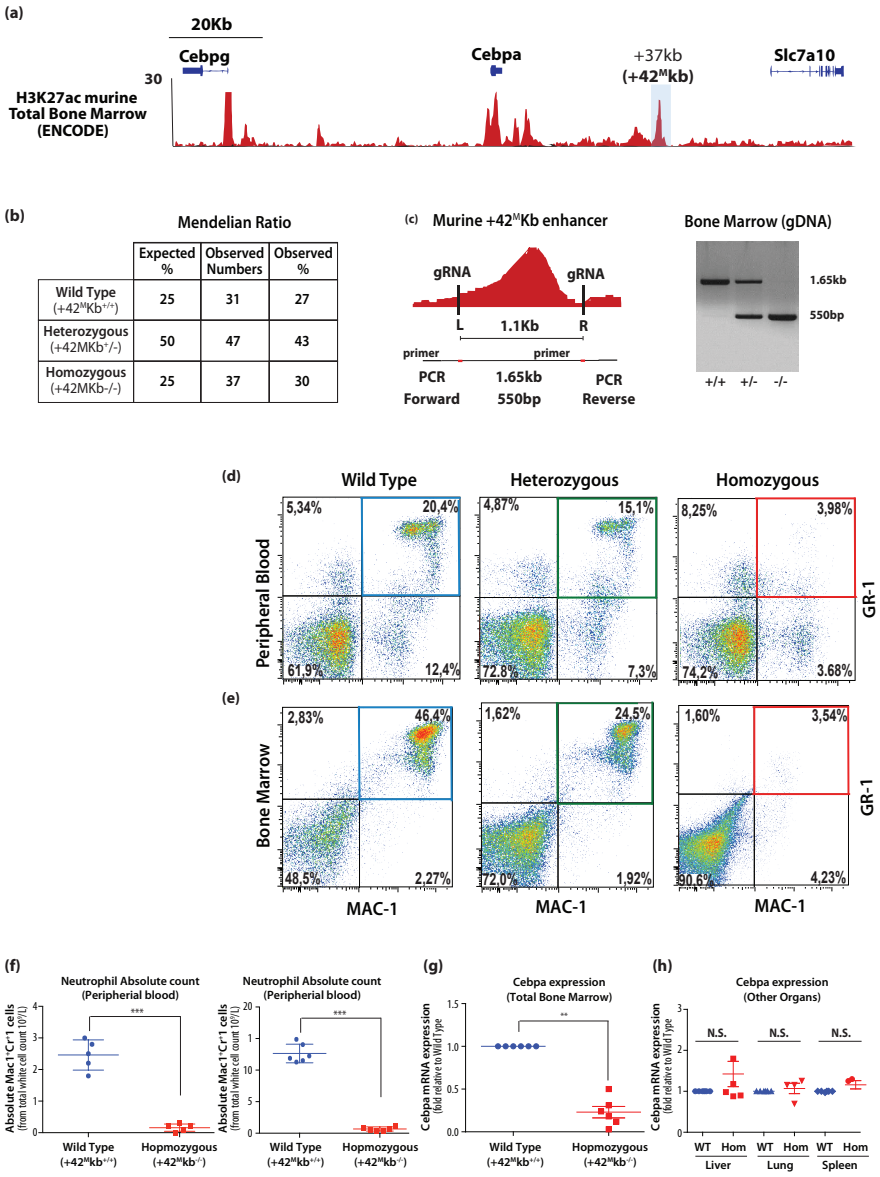


Figure 5: +37kb deleted mice (+42^{M}kb) show low *Cebpa* levels and develop neutropenia. (**a**) ChIP-seq H3K27ac in murine total bone marrow (ENCODE) shows multiple regions of open chromatin. A region located at +37kb in mice is highly homologous (Fig S5a) to the human +42kb enhancer and is H3K27ac marked. (**b**) Table showing Mendelian ratios. (**c**) PCR genotyping using primers flanking the gRNAs generate an amplicon of 1.65kb on the intact/wild type allele and an amplicon of 550bp on the deleted/rearranged allele. (**d and e**) Flow-cytometric analysis to distinguish neutrophils in peripheral blood or bone marrow of wild type (blue), heterozygous (green) and homozygous (red) mice using the myeloid differentiation markers Mac1 and GR1. (**f**) Neutrophil absolute counts in peripheral blood and bone marrows of wild type and homozygous mice (**g**) *Cebpa* mRNA expression from total bone marrow obtained from wild type (n=6) or homozygous +42^{M}kb knockout mice (n=6) is presented as fold change. (**g**) *Cebpa* mRNA expression from liver, lung and spleen does not show significant changes. *** = P<0.0001;** = P<0.001; N.S. = not significant.

The +42kb enhancer regulates *CEBPA* in myeloid cells

We next determined the activity of the +42kb enhancer using luciferase reporter assays. The +42kb enhancer was cloned into a luciferase reporter construct driven by the *CEBPA* promoter and its activity was investigated in different cell lines (Fig.4a). A 2.5-fold increase of luciferase activity was observed in the myeloid cell lines U937 and THP-1, compared to the luciferase-reporter driven by the *CEBPA* promoter only. The +42kb enhancer was not active in non-myeloid cell lines Jurkat, HepG2, H292 or HEK293T (Fig.4a). In contrast, the activity of the +9kb enhancer was more general across different cell types (Fig.4b). Using CRISPR/Cas9 genome editing technology, we next generated gRNAs flanking the TF and the p300 binding region of the +42kb enhancer (approximately 800bps) and co-electroporated them with Cas9 into the myeloid cell line THP-1. Targeted THP-1 cells generated heterozygous clones (Fig.4c), which were further tested for *CEBPA* expression by qPCR. Deletion of the +42kb enhancer resulted in 2-4 fold reduced *CEBPA* transcript levels as compared to wild type controls (Fig.4d). No changes in *CEBPG* mRNA expression levels were observed (Fig.4e). *SLC7A10,* located 5' of the 170kb *CEBPA* TAD is not expressed in THP-1 cells (data not shown). In contrast to the effects observed in THP-1 cells, deletion of the +42kb enhancer in the Hep3B hepatocyte cell line, revealed no changes in *CEBPA* or *CEBPG* expression compared to wild type clones (Fig.S4). These results suggest a tissue-specific role of the +42kb enhancer in the regulation of *CEBPA* levels in myeloid cells.

In vivo deletion of the murine* +42kb *homologous enhancer causes neutropenia

We hypothesized that deletion of the +42kb enhancer *in vivo* would cause neutropenia due to a selective decrease of *Cebpa* levels in myeloid progenitors, leaving other *Cebpa* expressing organs unaffected. A region located +37kb from the mouse *Cebpa* TSS shows approximately 90% homology with the human +42kb region and is H3K27ac enriched in mouse bone marrow (Fig.5a, S5a, ENCODE)[29]. Applying CRISPR/Cas9 nickase technology, we generated three +37kb (here designated +42^{M}kb) knock-out mouse lines (Fig.S5b). Genotyping, confirmed germ-line deletion of the enhancer in the three lines and revealed Mendelian distributions of wild type (+42^{M}kb$^{+/+}$), heterozygous- (+42^{M}kb$^{+/-}$) and homozygous-deleted (+42^{M}kb$^{-/-}$) mice (Fig.5b and c). In contrast to full *Cebpa* knockout mice[13], homozygous +42^{M}kb$^{-/-}$ mice were viable after birth and histopathological analysis of 4 to 5 week old mice did not reveal any major defects in lung, liver or spleen tissue (data not shown). Flow-cytometric analysis of blood and bone marrow showed a strong reduction of Mac1^{+}Gr1^{+} mature neutrophils in +42^{M}kb$^{-/-}$ mice compared to age-matched +42^{M}kb$^{+/-}$ and +42^{M}kb$^{+/+}$ control animals (Fig.5d, e and f). May-Grünwald staining of bone marrow cells showed the reduction of neutrophils in +42^{M}kb$^{-/-}$ mice compared to control mice (Fig.S5d). It is important to note that, in line with the fact that the neutrophil count was severely reduced, approximately 30% of the +42^{M}kb$^{-/-}$ mice died from bacterial infections 3 to 4 weeks after birth, as illustrated by the presence of bacteria in blood vessels of multiple tissues by histopathological analysis (Fig.S5f). Other

blood indices, including total white blood cell count and hemoglobin concentration revealed no differences between wild type and mutant mice (Fig.S5c and d). A slight increase in lymphocyte and monocyte counts was noticed, probably due to reduced space that is normally occupied by neutrophils in the bone marrow (data not shown). *Cebpa* transcript levels were 60-80% reduced in total bone marrow of +42^{M}kb$^{-/-}$ mice compared to wild type control mice (Fig.5f), but no changes in expression were observed in *Cebpg* expression levels, although *Cebpa* knockout shows increase in Cebpg levels[41]. No decrease of *Cebpa* transcript levels was observed in lung, liver and spleen of +42^{M}kb$^{-/-}$ mice (Fig.5g).

The +42^{M}kb enhancer controls *Cebpa* expression in GMPs and CMPs

We investigated whether the reduced neutrophil numbers are preceded by a decrease in *Cebpa* levels in bone marrow progenitor cells. Flow-cytometric analysis showed a significant reduction in absolute numbers of granulocytic/monocytic progenitors (GMP) and a significant increase in common myeloid progenitor (CMP) numbers in the bone marrow of +42^{M}kb$^{-/-}$ mice compared to +42^{M}kb$^{+/+}$ controls (Fig.6a and b). We performed RNA-seq on sorted progenitor fractions and observed that *Cebpa* levels were reduced more than 100-fold in CMPs and GMPs in +42^{M}kb$^{-/-}$ (n=3) compared to +42^{M}kb$^{+/+}$ (n=3) mice (Fig.6c). *Cebpa* levels were low to absent in MEP sorted populations from +42^{M}kb$^{+/+}$ and +42^{M}kb$^{-/-}$ mice (data not shown). CMPs and GMPs derived from +42^{M}kb$^{+/+}$ and +42^{M}kb$^{-/-}$ mice also showed major differences in expression levels of myeloid-associated genes (Fig.6c; Fig.S6a and b). One of these target genes is *Csf3r,* encoding the colony stimulating factor receptor Csf3r, required for GMP survival and neutrophilic differentiation. *Csf3r* transcript levels were decreased 20-fold in total marrow, as well as in CMP and GMP FACS sorted fractions from +42^{M}kb$^{-/-}$ mice (Fig.6d). Consequently, bone marrow cells from +42^{M}kb$^{-/-}$ mice failed to form colonies in response to GCSF (Fig.6e). These data show that the enhancer is required at early stages of myeloid development and acts as a main activator of the CSF3-driven myeloid differentiation program.

Loss of hematopoietic stem cells and expansion of multi-potent progenitors in +42^{M}kb$^{-/-}$ mice

We next investigated the effects of +42^{M}kb enhancer deletion on the distribution of HSCs and multipotent progenitor cells (MPPs) (Fig.7b). Absolute numbers of the lin^{neg}Sca-1^{pos}c-KIT^{pos} (LSK) cells were significantly higher in +42^{M}kb$^{-/-}$ mice than in control mice (Fig.7a and b). RNA-seq of the LSK fractions revealed that *Cebpa* levels were several folds lower in +42^{M}kb$^{-/-}$ (n=3) than in +42^{M}kb$^{+/+}$ LSK cells (n=3) (Fig.7c). Within the LSK population MPPs can be discriminated from short-term and long-term HSCs using SLAM (signaling lymphocyte activating molecules) code CD48/CD150 markers. The lin^{neg}CD48^{neg}CD150^{pos} LT-HSCs and lin^{neg}CD48^{neg}CD150^{neg} ST-HSCs were reduced by 10-20 fold in +42^{M}kb$^{-/-}$ mice (Fig.7a and b). These changes are in line with data from Porse and colleagues[43] using Mx-cre

Cebpa conditional knockout model, indicating that decreased CEBP/α levels disturbs the integrity of the HSC pool. Conversely, a significant increase of the lin^{neg}CD48^{pos}CD150^{neg} and lin^{neg}CD48^{pos}CD150^{pos} MPP population was observed in +42^{M}kb$^{-/-}$ mice. These data show that the enhancer activity at the HSPC stage is essential to maintain constant *Cebpa* levels in myeloid-primed progenitors during the course of myelopoiesis.

Sustained proliferation of +42^{M}kb$^{-/-}$ bone marrow progenitor cells

To investigate the effects of +42^{M}kb enhancer deletion on the proliferative behavior of bone marrow progenitors, colony cultures were carried out using a combination of IL-3, IL-6, SCF and GM-CSF. No differences in primary colony numbers were observed between +42^{M}kb$^{-/-}$, +42^{M}kb$^{+/-}$ or +42^{M}kb$^{+/+}$ mice. +42^{M}kb$^{-/-}$ bone marrow cells could be serially replated, whereas +42^{M}kb$^{+/-}$ and +42^{M}Kb$^{+/+}$ cells underwent exhaustion (Fig.7d). Flow-cytometric analysis revealed that the majority of the replated cells from +42^{M}kb$^{-/-}$ mice expressed lin^{neg}CD48^{pos}CD150^{neg} MPP and lin^{neg}CD16/32^{pos}CD34^{pos} GMP markers, with minimal neutrophilic differentiation (Fig.7d).

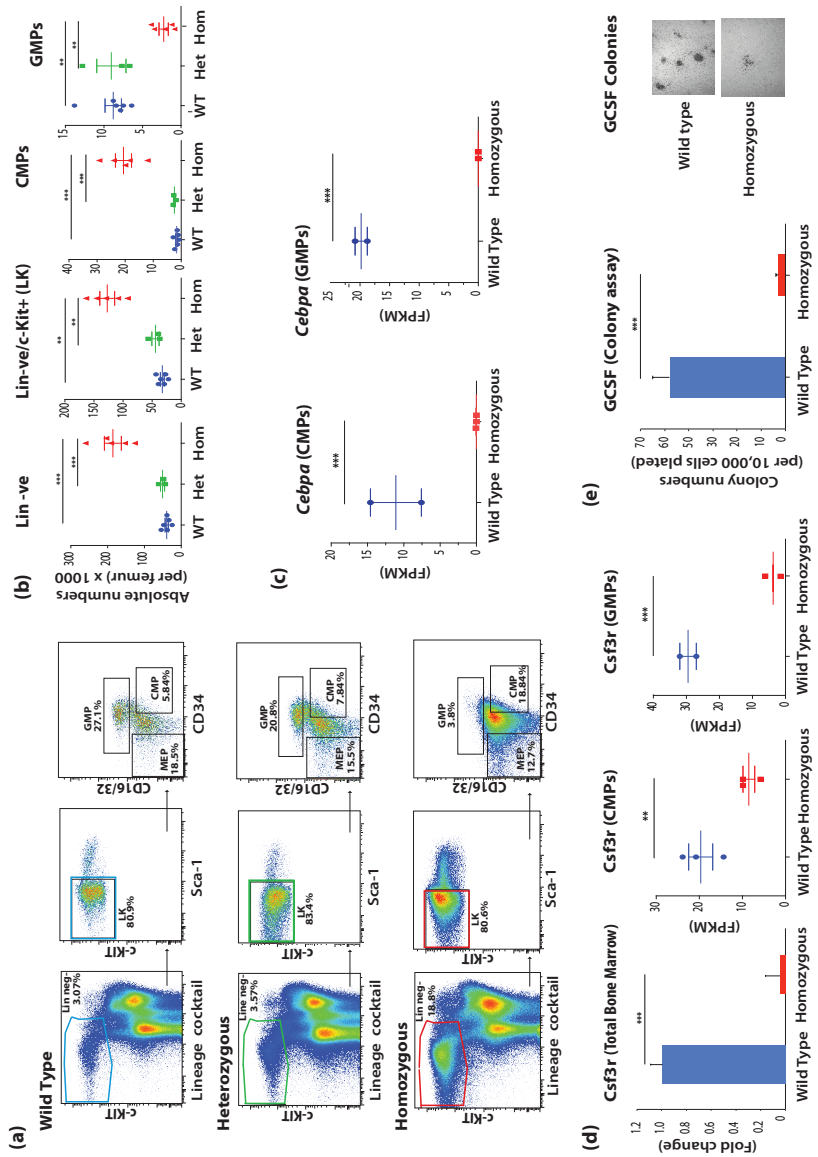


Figure 6: Reduction in GMPs, increase in CMPs and loss of GCSF response in +42^{M}kb enhancer deleted bone marrow. (**a**) Lineage negative cKIT^{pos}Sca-1^{neg} (LK) cells were derived from gated c-KIT^{pos} cells. The myeloid progenitor cell population including CMP, GMP and MEP was characterized using CD34 and CD16/32 markers gated from LK cells. (**b**) Absolute numbers for lineage negative cells, LK, CMP and GMP cell populations were calculated from bone marrow white cell count per femur. (**c**) *Cebpa* expression measured by RNA-seq expressed as FPKM values derived for wild type and homozygous mice in CMP (2wt vs. 3hom) and GMP (2wt vs. 2hom) sorted fractions. (**d**) Csf3r expression in total bone marrow by qPCRs, presented as fold change between wild type (n=3) and homozygous (n=3) mice. RNA-seq analysis of *Csf3r* in FACS-sorted CMP and GMP cell populations with values expressed as FPKM. (**e**) Numbers of CSF3 stimulated colonies per 10,000 cells plated obtained from wild type bone marrow or from +42^{M}kb homozygous deleted mice. Colony numbers represent the average of three independent experiments. Representative microphotographs of colonies show differences in sizes and numbers between wild type and homozygous mice.

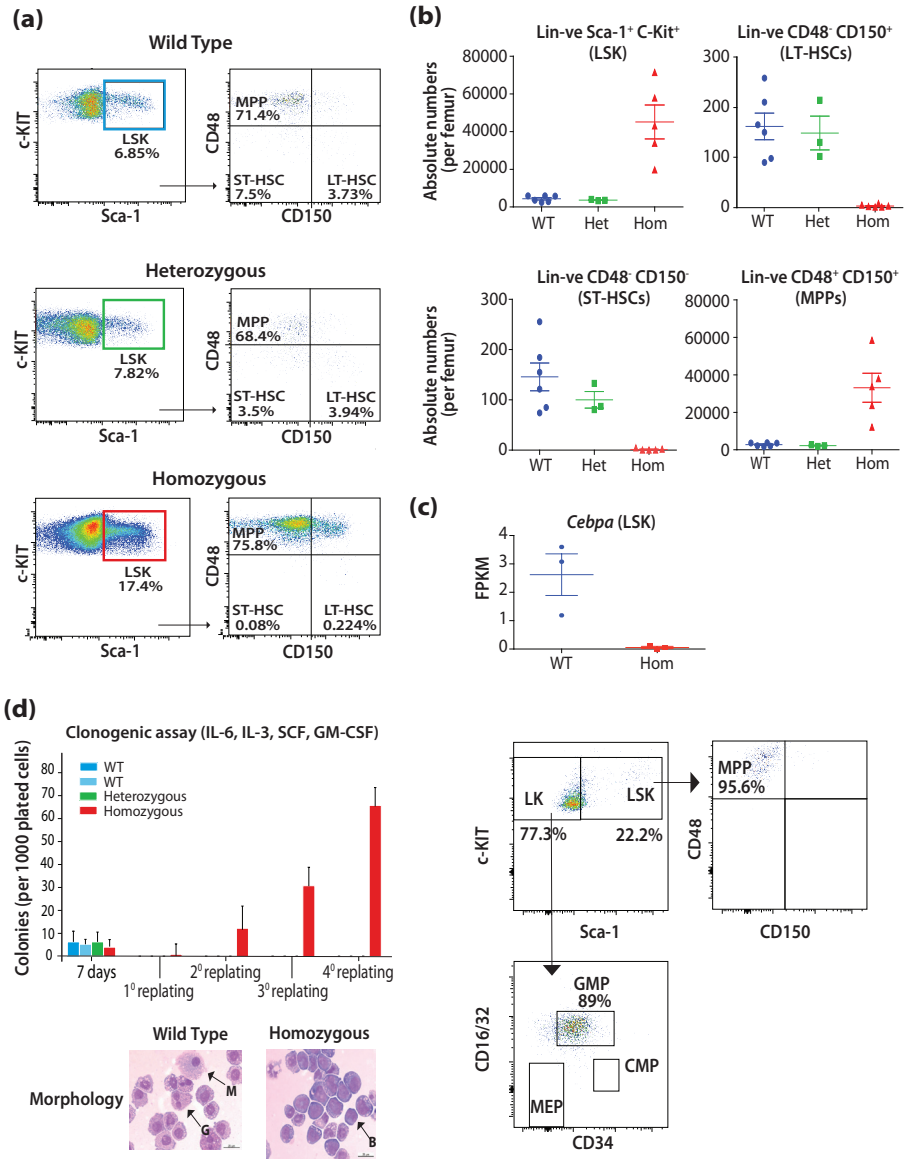


Figure 7: ***Loss of hematopoietic stem cells and expansion of multi-potent progenitors in +42^{M}kb$^{-/-}$ mice.*** (**a**) SLAM CD48+CD150+ markers were used to characterize cell distribution within the MPP, LT-HSC and ST-HSC cell populations gated from lin^{neg}Sca-1^{pos}c-kit^{pos} (LSK) cell populations. (**b**) Absolute cell numbers for LSK, MPP, LT-HSC and ST-HSC were calculated from bone marrow white cell count per femur. (**c**) *Cebpa* expression by RNA-seq (FPKM values of wild type vs. homozygous mice). (**d**) Total bone marrow cells from wild type, heterozygous and homozygous mice were cultured in semi-solid medium supplemented with IL-3, IL-6, SCF and GM-CSF. Colonies were counted and replated every seven days. FACS plots showing that majority of cells grown under these conditions are mainly LK/GMP cells, and to a less extent MPP/LSK cells. Morphological examination with May-Grünwald-Giemsa after 7 days, distinguishes normal granulocytic and macrophage differentiation in wild type cells as compared to homozygous cells that show blasts as the major cell population.

DISCUSSION

Applying functional genomics and genome editing to different models including human hematopoietic cell lines, human bone marrow progenitors and *in vivo* mouse models, we show that a single myeloid-specific enhancer is (1) autonomously responsible for *CEBPA* expression in myeloid primed HSCs, (2) initiates the myeloid gene expression program and (3) indispensable for neutrophil development.

Deletion of the +42kb enhancer in our model causes loss of *Cebpa* expression in the myeloid lineage and failure to induce complete neutrophilic differentiation. This suggests that the +42kb enhancer initiates the myeloid program by acting as a highly occupied target region for HSC-related TFs (Fig.3) and thereby activates *Cebpa*[34]. Upon myeloid commitment, the intergenic site between *CEBPA* and *SLC7A10* contains multiple enhancers which become active during myeloid differentiation (Fig.2, Fig.S5d), raising the question if these enhancers possess functional redundancy, thus, potentially acting as shadow enhancers[32,33]. We postulate that (1) the +42kb enhancer works autonomously at the HSPC stage to induce necessary *Cebpa* levels, (2) followed by activation of the myeloid transcription program inducing myeloid commitment, (3) upon myeloid commitment the other enhancers become active in order to serve as a transcription activation platform[35] and elevate *Cebpa* levels to a necessary level for terminal neutrophilic differentiation (Fig.S8). We predict that upon deletion of the +42kb enhancer the other enhancers within the locus will not become active (absence of H3K27ac) resulting in failure to prime myeloid differentiation.

The +42kb enhancer deletion causes a reduction in LT-HSCs and ST--HSCs, therefore recapitulating the phenotype of the Mx-Cre/*Cebpa* conditional knockout mice[42]. In this study, as well as in other studies[43,44,] it has been demonstrated that *Cebpa* levels are critical to maintain HSC numbers and survival under a quiescent state. Given that only a small population of HSCs expresses *Cebpa*[45], it remains unclear what causes the severe loss of HSCs upon deletion of either Cebpa or the +42^{M}kb enhancer. The LSK fraction (including the MPPs, LT-HSCs and ST-HSCs) shows significant reduction of *Cebpa* levels in the +42^{M}kb-/- mice, suggesting a critical role for the enhancer in *Cebpa* regulation in LSKs. However, given that the MPP fraction (CD48+CD150-) constitutes the majority of the LSK population, *Cebpa* downregulation in the LSK fraction (Fig.7b) mainly reflects *Cebpa* level changes in the MPP population. Given that C/EBPa negatively regulates cell cycle genes to keep a constant balance of proliferation and differentiation[46], it is possible that the block in differentiation leads to a constant demand for myeloid progenitor production, causing HSC exhaustion. In our model, the expansion of the progenitor population argues in favor of an increased progenitor state as a negative feedback mechanism to compensate for the differentiation block.

CEBPA is located in an enhancer-rich TAD and its promoter contacts eight intergenic sites. One question to be addressed is which potential architectural proteins or protein complexes that mediate the +42kb enhancer (or any of the other interacting enhancers) to *CEBPA* promoter interaction. The TAD is confined to a genomic region of 230kb with borders demarcated by CTCF (Fig.1C), an architectural protein involved in looping interactions within and across TADs[23,37,38]. CTCF also binds to the promoter of *CEBPA*, possibly, by forming multiple extrusions of the 5' and 3' interacting intergenic sites of *CEBPA*[39]. The intergenic sites contacting *CEBPA* are highly enriched for H3K27ac thereby marking potential enhancers, but they lack CTCF or cohesion binding. From our motif analysis data, (Fig.S3) we found that the +42kb enhancer harbors a ZNF143 DNA binding motif (CAGCCTTCATGCATTG). ZNF143 is a zinc finger TF that associates with CTCF to allocate enhancers close to promoters and facilitate transcription regulation[40]. It is therefore possible that ZNF143 has implications in initiating this interaction by binding the +42kb enhancer to associate with CTCF on the *CEBPA* promoter, thus causing the +42kb enhancer- *CEBPA* promoter interaction (Fig.S7). To test this hypothesis, functional experiments including genome editing of the ZNF143 binding site followed by 4C-seq are required to reveal the association between ZNF143 binding and *CEBPA* regulation in terms of a ZNF143-dependent *CEBPA* promoter to enhancer interaction.

Diverse oncogenic mechanisms that affect C/EBPa expression or function were reported in various subsets of AML[7,8,9,10,11,12]. It is possible that mutations in the +42kb enhancer could relate to transforming events. The expansion of the MPP population in the enhancer deleted mice suggests a pre-leukemic potential, which can only be confirmed by conducting serial transplantation experiments. The sustained replating of the +42^{M}kb MPPs (Fig.7d) are in concordance with a pre-leukemic state of the cells. It is also possible that the enhancer is involved in epigenetic deregulation of the *CEBPA* gene in certain AMLs. Patients with a chromosomal translocation t(8;21) present with low *CEBPA* expression levels. The t(8;21) generates the AML1-ETO (i.e., the *RUNX1-RUNX1T1* fusion transcript) fusion protein that binds to sites that are usually bound by RUNX1. The +42kb enhancer carries multiple RUNX1 binding sites and ChIP-seq experiments in CD34+ cells revealed that RUNX1 binds to the enhancer (Fig.3). Knock-down of *RUNX1-RUNX1T1* in Kasumi-1 cell line demonstrated a significant up-regulation of *CEBPA* mRNA and protein levels[48] but the mechanism by which this happens has not been resolved. Our data suggest that the +42kb enhancer is major interaction site for AML1-ETO, which may deregulate *CEBPA* expression. Interestingly, the transforming EVI1 protein (unpublished observation) also binds the +42kb

Acknowledgements

The authors wish to thank Erik Kreiter for critically reading the manuscript. This work was supported by the Dutch Cancer Foundation KWF (Grant # EMCR 2013-5829), Worldwide Cancer Research (Grant # 12-1309) and the Tata Memorial Trust Foundation. H.J.G. v.d. W. was supported by a Zenith grant (93511036) from the Netherlands Genomics Initiative (NGI).

Author contribution

R.A, E.B., and R.D. conceived of and designed the study. M.R., C.G, J.P., D.B., and R.D. provided study materials or patient samples. R.A., M.H., C.E., M.S., R.H., H.vd.W., E.R., K.v.L., C.G., M.R., J.P., D.B., I.T., S.G., T.K, E.B., and R.D. provided data analysis and interpretation, R.A., M.H., C.E., P.S., C.G., D.B., T.K., H.d.L. and E.B. carried out experiments, R.A., E.B, I.T. and R.D. wrote the manuscript.

The authors have no conflict of interest.

REFERENCES

1. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol,* 2015; **3,**144-54.
2. Miyamoto T, Iwasaki H, Reizis B, Ye M, Graf T, Weissman IL, Akashi K. Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Dev Cell*. 2002; **1,**137-47.
3. Pina C, Fugazza C, Tipping AJ, Brown J, Soneji S, Teles J, Peterson C, Enver T. Inferring rules of lineage commitment in haematopoiesis. *Nat Cell Biol*. 2012; **3,**287-94.
4. Sive JI, Göttgens B. Transcriptional network control of normal and leukaemic haematopoiesis. *Exp Cell Res*. 2014; **2,**255-64.
5. Iwasaki H, Akashi K. Myeloid lineage commitment from the hematopoietic stem cell. *Immunity*. 2007; **6,**726-40.
6. Iwasaki H, Mizuno S, Arinobu Y, Ozawa H, Mori, Y, Shigematsu H, Takatsu K, Tenen DG, Akashi K. The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. *Genes Dev*. 2006; **21,**3010-21.
7. Figueroa ME, Wouters BJ, Skrabanek L, Glass J, Li Y, Erpelinck-Verschueren CA, Langerak AW, Löwenberg B, Fazzari M, Greally JM,Valk PJ, Melnick A, Delwel R. Genome-wide epigenetic analysis delineates a biologically distinct immature acute leukemia with myeloid/T-lymphoid features. *Blood*. 2009; **12,**2795-804.
8. Pabst T, Mueller BU, Harakawa N, Schoch C, Haferlach T, Behre G, Hiddemann W, Zhang DE, Tenen DG. AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia. *Nat Med*. 2001; **4,**444-51.
9. Perrotti D, Cesi V, Trotta R, Guerzoni C, Santilli G, Campbell K, Iervolino A, Condorelli F, Gambacorti-Passerini C, Caligiuri MA, Calabretta B. BCR-ABL suppresses C/EBPalpha expression through inhibitory action of hnRNP E2. *Nat Genet.* 2002; **1,**48-58.
10. Radomska HS, Bassères DS, Zheng R, Zhang P, Dayaram T, Yamamoto Y, Sternberg DW, Lokker N, Giese NA, Bohlander SK, Schnittger S, Delmotte MH, Davis RJ, Small D, Hiddemann W, Gilliland DG, Tenen D.G. Block of C/EBP alpha function by phosphorylation in acute myeloid leukemia with FLT3 activating mutations. *J Exp Med*. 2006; **2,** 371-81.
11. Pabst T, Mueller BU, Zhang P, Radomska HS, Narravula S, Schnittger S, Behre G, Hiddemann W, Tenen DG. Dominant-negative mutations of *CEBPA*, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia. *Nat Genet*. 2001; **3,** 263-70.
12. Wouters BJ, Löwenberg B, Erpelinck-Verschueren CA, van Putten WL, Valk PJ, Delwel R. 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*. 2009; **13,** 3088-91.
13. Zhang DE, Zhang P, Wang ND, Hetherington CJ, Darlington GJ, Tenen DG. Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. *Proc Natl Acad Sci U S A*. 1997; **2,** 569-74.
14. Akai Y, Oitate T, Koike T, Shiojiri N. Impaired hepatocyte maturation, abnormal expression of biliary transcription factors and liver fibrosis in C/EBPα(*Cebpa*)-knockout mice. *Histol Histopathol*. 2014; **1,** 107-25.
15. Bassères DS, Levantini E, Ji H, Monti S, Elf S, Dayaram T, Fenyus M, Kocher O, Golub T, Wong KK, Halmos B, Tenen D.G. Respiratory failure due to differentiation arrest and expansion of alveolar cells following lung-specific loss of the transcription factor C/EBPalpha in mice. *Mol Cell Biol*. 2006; **3,** 1109-23.
16. Chen W, Zhu G, Hao L, Wu M, Ci H, Li YP. C/EBPα regulates osteoclast lineage commitment. *Proc Natl Acad Sci U S A*. 2013; **18,** 7294-9.

17. Chandrasekaran C, Gordon J. Cell lineage-specific and differentiation-dependent patterns of CCAAT/enhancer binding protein alpha expression in the gut epithelium of normal and transgenic mice. *Proc Natl Acad Sci U S A*. 1993; **19,** 8871-5.

18. Risca VI, Greenleaf WJ. Unraveling the 3D genome: genomics tools for multiscale exploration. *Trends Genet*. 2015; **7,** 357-72.

19. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010; **50,** 21931-6.

20. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011; **7333,** 279-83.

21. Spitz F, Furlong EE. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet*. 2012; **9**, 613-26.

22. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009; **5950,** 289-93.

23. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012; **7398,** 376-80

24. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014; **7,** 1665-80.

25. de Laat W, Duboule D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature*. 2013; **7472,** 499-506.

26. Cooper S, Guo H, Friedman AD. The +37 kb CEBPA-enhancer Is Critical for Cebpa Myeloid Gene Expression and Contains Functional Sites that Bind SCL, GATA2, C/EBPα, PU.1, and Additional Ets Factors. *PLoS One*. 2015; **5,** e0126385.

27. Hadrys T, Punnamoottil B, Pieper M, Kikuta H, Pezeron G, Becker TS, Prince V, Baker R, Rinkwitz S. Conserved co-regulation and promoter sharing of hoxb3a and hoxb4a in zebrafish. *Dev Biol*. 2016; **1,** 26-43.

28. Michalak P. Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. *Genomics*. 2008; **3,** 243-8.

29. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; **7414,** 57-74.

30. Beck D, Thoms JA, Perera D, Schütte J, Unnikrishnan A, Knezevic K, Kinston SJ, Wilson NK, O'Brien TA, Göttgens B, Wong JW, Pimanda JE. Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes. *Blood*. 2013; **14**, 12-22.

31. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature*. 2009; **7231,** 854-8.

32. Frankel N, Davis GK, Vargas D, Wang S, Payre F, Stern DL. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature*. 2010; **7305,** 490-3 .

33. Barolo S. Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy. *Bioessays*. **2,** 2012**;** 135-41.

34. Kvon EZ, Stampfel G, Yáñez-Cuna JO, Dickson BJ, Stark A. HOT regions function as patterned developmental enhancers and have a distinct cis-regulatory signature. *Genes Dev*. 2012; **9,** 908-13.

35. Koch F, Fenouil R, Gut, M, Cauchy P, Albert TK, Zacarias-Cabeza J, Spicuglia S, de la Chapelle AL, Heidemann M, Hintermair C, Eick D, Gut I, Ferrier P, Andrau JC. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat Struct Mol Biol*. 2011; **8,** 956-63.

36. Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, Furlong EE. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature*. 2014; **7512,** 96-100.

37. Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, Ong CT, Hookway TA, Guo C, Sun Y, Bland MJ, Wagstaff W, Dalton S, McDevitt TC, Sen R, Dekker J, Taylor J, Corces VG. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell*. 2013; **6,** 1281-95.

38. Van Bortle K, Nichols MH, Li L, Ong CT, Takenaka N, Qin ZS, Corces VG. Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol.* 2014; **6,** R82.

39. Sanborn AL, Rao SS, Huang SC, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J, Geeting KP, Gnirke A, Melnikov A, McKenna D, Stamenova EK, Lander ES, Aiden EL. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A*. 2015; **47,** E6456-65.

40. Bailey SD, Zhang X, Desai K, Aid M, Corradin O, Cowper-Sal Lari R, Akhtar-Zaidi B, Scacheri PC, Haibe-Kains B, Lupien M. ZNF143 provides sequence specificity to secure chromatin interactions at gene promoters. *Nat Communications*. 2015; **2,** 6186.

41. Alberich-Jordà. M, Wouters B, Balastik M, Shapiro-Koss C, Zhang H, Di Ruscio A, Radomska HS et al. C/EBPγ deregulation results in differentiation arrest in acute myeloid leukemia. *J Clin Invest*. 2012; **12,** 4490-504.

42. Zhang P, Iwasaki-Arai J, Iwasaki H, Fenyus ML, Dayaram T, Owens BM, Shigematsu H, Levantini E, Huettner CS, Lekstrom-Himes JA, Akashi K, Tenen D.G. Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity*. 2004 **6,** 853-63.

43. Hasemann MS, Lauridsen FK, Waage J, Jakobsen JS, Frank AK, Schuster MB, Rapin N, Bagger FO, Hoppe PS, Schroeder T, Porse BT. C/EBPα is required for long-term self-renewal and lineage priming of hematopoietic stem cells and for the maintenance of epigenetic configurations in multipotent progenitors. *PLoS Genet*. 2014; **1,** e1004079.

44. Ye M, Zhang H, Amabile G, Yang H, Staber PB, Zhang P, Levantini E, Alberich-Jordà, M, Zhang J, Kawasaki A, Tenen D.G. C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence. *Nat Cell Biol*. 2013; **4,** 385-94.

45. Wölfler A, Danen-van Oorschot AA, Haanstra JR, Valkhof M, Bodner C, Vroegindeweij E, van Strien P, Novak A, Cupedo T, Touw IP. Lineage-instructive function of C/EBPα in multipotent hematopoietic cells and early thymic progenitors. *Blood*. 2010; **20,** 4116-25.

46. Porse BT, Bryder D, Theilgaard-Mönch K, Hasemann MS, Anderson K, Damgaard I, Jacobsen SE, Nerlov C. Loss of C/EBP alpha cell cycle control increases myeloid progenitor proliferation and transforms the neutrophil granulocyte lineage. *J Exp Med*. 2005; **1,** 85-96.

47. Ma O, Hong S, Guo H, Ghiaur G, Friedman AD. Granulopoiesis requires increased C/EBPα compared to monopoiesis, correlated with elevated Cebpa in immature G-CSF receptor versus M-CSF receptor expressing cells. *PLoS One*. 2014; **4,** e95784.

48. Ptasinska A, Assi SA, Martinez-Soria N, Imperato MR, Piper J, Cauchy P, Pickin A, James SR, Hoogenkamp M, Williamson D, Wu M, Tenen DG, Ott S, Westhead DR, Cockerill PN, Heidenreich O, Bonifer C. Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal. *Cell Rep*. **6,** 2014**;** 1974-88 (2014).

49. Chen L, Kostadima M, Martens JH, Canu G, Garcia SP, et al. Transcriptional diversity during lineage commitment of human blood progenitors. *Science*. 2014; **6204,** 1251033.

50. van Galen P, Kreso A, Wienholds E, Laurenti E, Eppert K, Lechman ER, Mbong N, Hermans K, Dobson S, April C, Fan JB, Dick JE. Reduced lymphoid lineage priming promotes human hematopoietic stem cell expansion. *Cell Stem Cell*. 2014; **2,** 94-106.
51. Hübner M.R and Spector DL. Chromatin Dynamics. *Annual Review of Biophysics*. 2010; **39**, 471-489 (2010).
52. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, Yen CA, Schmitt AD, Espinoza CA, Ren B. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*. 2013; **14**, 290-294.
53. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010; **38**, 576–589.
54. R Core Team. R: A Language and Environment for Statistical Computing. (2015)
55. van de Werken HJG, de Vree PJP, Splinter E, Holwerda SJB, Klous P, de Wit E, and de Laat W. 4C Technology: Protocols and Data Analysis. *Methods Enzymol*. **513**, 2012(a) 89–112.
56. van de Werken HJG, Landan G, Holwerda SJB, Hoichman M, Klous P, Chachik R, Splinter E, Valdes-Quezada C, Öz Y, Bouwman, B.A.M, et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* 2012(b); **9**, 969–972.
57. Y.Benjamini, Y.H. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing When researchers tend to select pursuing multiple the (statistically) and support of conclusions. An unguarded use in a greatly results of single-inference inc. *J. R. Stat. Soc. Ser. B* 2015; **57**, 289–300.
58. Bindels EM, Havermans M, Lugthart S, Erpelinck C, Wocjtowicz E, Krivtsov AV, Rombouts E, Armstrong SA, Taskesen E, Haanstra JR, Beverloo HB, Döhner H, Hudson WA, Kersey JH, Delwel, R, Kumar AR. EVI1 is critical for the pathogenesis of a subset of MLL-AF9-rearranged AMLs. *Blood*. 2012; **24**, 5838-49.
59. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, Kim JS. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*. 2014; 1, 132-41.
60. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology*. 2013;**14**, R36.
61. Trapnell C, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*. 2012;**7**, 562-578.
62. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;**31**, 166-169 (2015).
63. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*. 2014; **15**, 550.
64. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, Zaretsky I, Jaitin DA, David E, Keren-Shaul H, Mildner A, Winter D, Jung S, Friedman N, Amit I. Immunogenetics. Chromatin state dynamics during blood formation. *Science*. 2014; **22**, 943-9.

SUPPLEMENTARY MATERIALS AND METHODS

4C-seq data analysis

4C-seq data analysis was performed as previously described[28]. In brief, the 4C-seq read primer sequences with their barcodes were used to demultiplex the reads and to trim the reads from the 5'-end to the first restriction enzyme recognition site. The sequences were mapped, while ignoring quality scores of the read bases and not allowing for a mismatch, to a database of digested genome fragment-ends using the mouse reference genome build mm9. All 4C-seq samples passed the quality control threshold values[28].

We normalized the data taking library size and 4C-seq fragment-end types into account[28,58]. The 4C-seq contact frequency profiles were generated with the median value of the CEPBA+ and CEPBA- 4C-seq data. The 4C-seq data was smoothed by applying a running trimmed (10%) mean approach using 21 fragment-ends in a single window.

We further determined regions of differentially contact frequencies. 4C-seq data are biased for each fragment-end differently and therefore do not follow a specific distribution. Hence, we used a non-parametric approach to test for statistical significance between two phenotypes. First, we ranked all normalized data for each fragment-end independently, and set ties to the minimal value. Subsequently, we binned, along the locus, the ranks of 21 fragment-ends and calculated the m x n rank frequency matrix (m = 2; i.e. the number of different phenotypes; n = number of samples). We merged the columns of the frequency matrix based on the order of each phenotype and the number of samples, e.g. if the first ranked phenotype, based on the enumeration of the rank frequencies multiplied by the rank number, has been sampled three times, the first three columns of the m x n matrix are merged by adding up the row values, and the other columns are merged similarly resulting in a 2 x 2 matrix. Moreover, a X^2-test was applied on this 2 x 2 contact frequency matrix. The p-values were corrected for multiple hypothesis testing using the Benjamini-Hochberg method[59] and areas with significant contact differences are indicated in gray in Figures 1d, S1e. The R statistical package version 3.2.2 was used for the statistical calculations and for generating the 4C-seq contact plots[60].

ChIP-seq alignment and peak calling

The immunoprecipitated DNA was processed according to the Illumina TruSeq ChIP Sample Preparation Protocol (Illumina) and single-end sequenced (1 × 50 bp) on the HiSeq 2500 platform. Reads were aligned to human genome build 19 (hg19). PCR bias amplicons were removed by SAMtools (http://samtools.sourceforge.net/). Subsequently, a peak-calling algorithm was applied by comparing target enrichment to its related input control using MACS 1.4.2, according to standard procedures

CRISPR of the +42kb enhancer in human cell lines

Two guide RNAs (gRNAs) flanking the +42Kb enhancer were designed using the **crispr.mit.edu** *tool. Guide RNAs were purchased as gBlocks (IDT) and cloned into a pCR2.1 TOPO blunt backbone using EcoRV blunt restriction sites (Invitrogen).*

Sorting of bulk GFP positive THP-1 cells was carried 48 hours post nucleofection (Amaxa, Lonza, using Kit-V) of 1x10^6 cells with Cas9-GFP (Addgene: 44719) together with gRNAs (1ug per construct). Bulk GFP positive THP-1 cells were grown as single cell derived colonies in methocult semi-solid medium. Hep3B cells (2x10^6) were transfected using Fugene 6 (Promega) with Cas9-GFP together with gRNAs (1ug per construct). Single cells were sorted in 96 well plates. Per cell line, 100 clones were expanded in 48 and 24 well plates. DNA was harvested for genotyping using primers flanking the gRNAs. Homozygous and wild type cell clones were harvested for total RNA isolation by Trizol (Invitrogen) extraction followed by DNase digestion. The SuperScript II Reverse Transcriptase kit (Invitrogen) was used for cDNA-synthesis. Quantitative RT-PCR was carried out on the 7500 Fast Real-time PCR System (Applied Biosystems). Relative levels of gene expression were calculated using the 2-ΔΔ**Ct** method. Statistical significance was calculated by a T-test (two-tailed) using GraphPad Prism software.

CRISPR gRNA sequence:

5': CCGGGAGCAGTCAGGATATC
3': AGGATCCACATAGACCCGAT
CRISPR PCR genotyping :
Forward : CAGGGTATTCCCTGTGGGGAAGCTTGGAGC
Reverse : GAAGGATCTCACAGGCTCCTGGGCTCAGGC
QPCR primers :
CEBPA
Forward : GGATAACCTTGTGCCTTG
Reverse : CTCCCCTCCTTCTCTCAT
CEBPG
Forward : GGCTAGAGGAGCAGGTACAT
Reverse : GCCTGGGTATGGATAACACTA

Generation of CRISPR mouse models using CRISPR/Cas9 nickase

The nickase system was used to reduce potential off-target effects by CRISPR/Cas9[61]. Two pairs of gRNAs were designed on either side of the mouse +42^{M}kb enhancer. Guide RNAs were purchased as gBlocks (IDT) and cloned into a pCR2.1 TOPO blunt backbone using EcoRV blunt restriction sites (Invitrogen). *In vitro* RNA transcription (ABI: MEGAscript T7 Transcription kit) from each of the four gRNAs was generated using T7 dependent RNA production from gBlocks (IDT) whereas Cas9n mRNA was purchased commercially (Sigma-Aldrich).

In brief, microinjections of 15ng/ul of gRNA and 30ng/ul of Cas9n mRNA were carried out in one cell stage zygotes derived from C57/BL6, which were then transferred into foster mice. Three mouse lines were generated by CRISPR/Cas9 via random on-target site recombination, producing three different deleted sites occurring within the enhancer (Fig.S5b). Using a PCR genotyping approach, each line produced different amplicon sizes. The biggest cut was of 1.1kb; a middle cut of approximately 1.05kb; smallest cut of 0.65kb. Heterozygous littermates were crossbred to obtain F1 generation mice and beyond for analysis in order to avoid mosaic deletion. All mouse lines were sacrificed at two time points: 4-5 weeks and 9-10 weeks after birth. RNAs from total bone marrow of wild type and homozygous mice were harvested by Trizol extraction, followed by DNase digestion. The SuperScript II Reverse Transcriptase kit (Invitrogen) was used for cDNA-synthesis. Quantitative real-time RT-PCR for ***Cebpa***, ***Cebpg*** and ***Csf3r*** was carried out on the 7500 Fast Real-time PCR System (Applied Biosystems). Relative levels of gene expression were calculated using the 2-ΔΔ**Ct** method. Statistical significance was calculated by a T-test (two-tailed) using GraphPad Prism software.

CRISPR gRNA sequence (nickase system):

5' : TGAAGCCTACACTACTTTGT and AGAGGTAGGAACTCCATTCC

3' : AGAGCCTCGCTCAAGCCCAT and TTGAGACATCTGGTAACCTT

CRISPR PCR genotyping :

Forward : GAGGTGACAGTCTGTGCAGCTGGGACACAAC

Reverse : GGATACTGATGGCTGATCCTCCCATTCCTC

Primers for QPCRs :

Csf3r

Forward : CCTGGATGATAGAACCTAACGGG

Reverse : CTCTCCAGCGAAGGTGTAGACA

Cebpa

Forward : GCAAAGCCAAGAAGTCGGTGGA

Reverse : CCTTCTGTTGCGTCTCCACGTT

Cebpg

Forward : GCTTACAGCAGGTTCCTCAGCT

Reverse : GGCGGTATTCGTCACTATTCCG

RNA-seq analysis

In brief, all reads were aligned against genes annotated in the RefSeq Transcriptome database and remaining non-aligned reads were aligned against the full genome. Gene expression levels were quantified by the fragments per kilobase of exon per million fragments mapped (FPKM) statistic as calculated by Cufflinks[59]. Hierarchical clustering analysis was performed on the FPKM values using complete linkage as clustering methodology and Euclidean distance

as distance measure using the Gplots package in the R environment. Read counts were determined with HTSeq-count[62] and subsequently used for differential expression analysis in DESeq2[63], with default parameters, in the R environment. Multiple testing correction was achieved by performing the Benjamini-Hochberg procedure[59] on the calculated p-values to control the False Discovery Rate (FDR).

Histopathological examination

Tissues (bone marrow, brain, heart, intestine, kidney, liver, lung, pancreas, spleen, thymus) were fixed in 10% neutral-buffered formalin and embedded in paraffin. Tissue sections of 3 µm thick were stained routinely with hematoxylin and eosin and examined by light microscopy. Bacterial infection was detected In all tissues except brain and intestine, based on the presence of bacteria in the lumina of blood vessels. In many places, these bacteria formed long chains. Postmortem invasion of tissues by bacteria was ruled out by the lack of autolysis of the tissues, together with the specific localization of bacteria in blood vessel lumina.

Luciferase reporter constructs: Primer sequences.

Full *CEBPA* canonical promoter primer sequences:
Forward: ATCACTGATATCGCCGACTCCATGGGGGAGTTAGAGTTCT;
Reverse: ATCACTAAGCTTGCCAGGCCTAAGGCCACTGTCGGTGAAG.

+42Kb enhancer primer sequences :
Forward : ATTATGTCGACGGATCCAGGCCCTATCCCAGGGTATT
Reverse: ATTATGTCGACGAGGCTGAGGAGCAAATCAC

+9Kb enhancer primer sequences :
Forward : attaatGTCGACTCTAGAAGCACGTGGGAATCATTAGC
Reverse: attaatGTCGACCAAGCCCTCTTGGATCTGAA

Accession numbers

ChIP-Seq data derived from cell lines are available in the ArrayExpress Database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-2225.
Blood-cell ChIP-seq data have been deposited in the NCBI GEO database (accession codes GSE40668).
ChIP-Seq of transcription factors on CD34+ HSPC fractions are available at http://149.171.101.136/python/BloodChIP/
The 4C-Seq results and RNA-Seq data of sorted mouse progenitors are deposited at the European Nucleotide Archive (ENA, http://www.ebi.ac.uk/ena/), which is hosted by the EBI.

SUPPLEMENTARY FIGURE LEGENDS

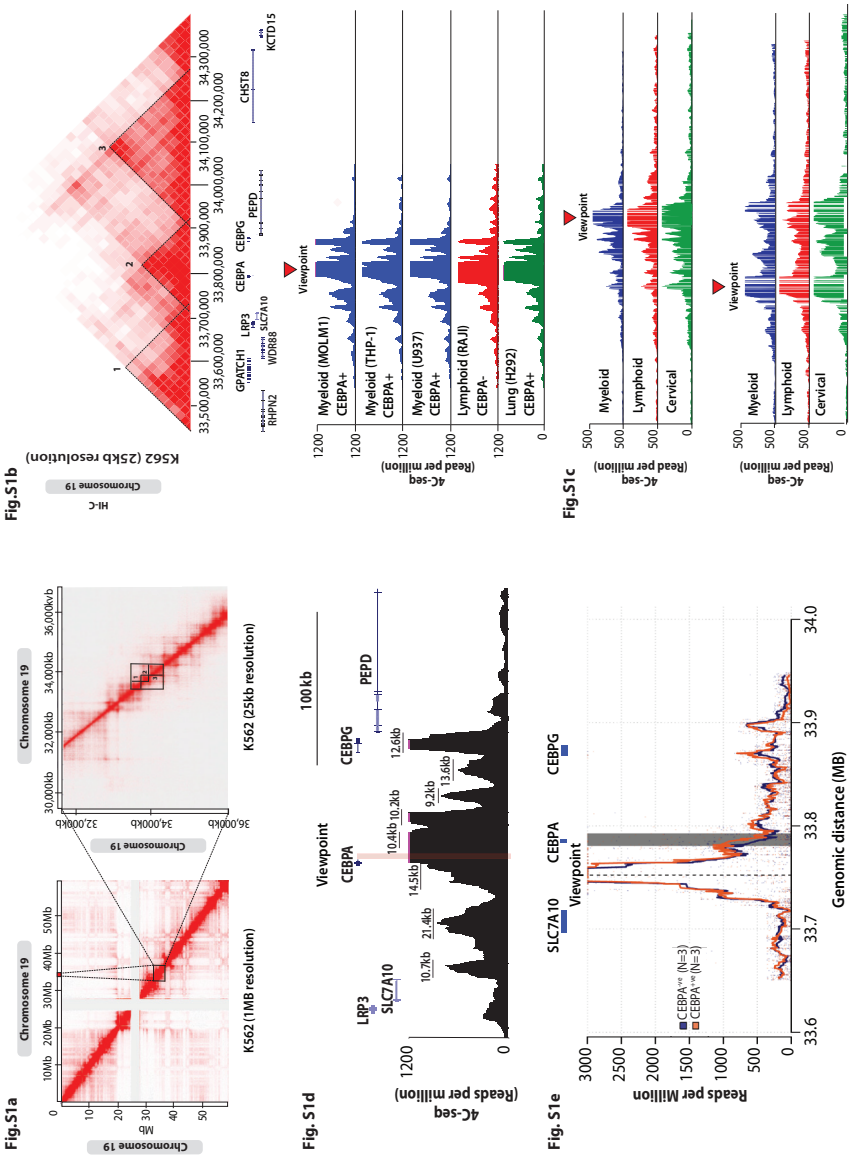


Fig.S1 : ***CEBPA* is confined within a 170kb TAD**. **(a)** HI-C heatmap matrix of chromosome 19 in K562 at 1MB resolution (top) and at 25Kb resolution (bottom), showing the *CEBPA*-TAD (**2**) and its adjacent TADs (**1** and **3**). (**b**) Normalized 4C-seq in *CEBPA*^{pos} myeloid cell lines (MOLM-11, THP-1, U937), *CEBPA*^{neg} lymphoid Raji cell line and *CEBPA*^{pos} lung H292 cell line show interactions held in TAD2. **(c)** Normalized 4C-seq in myeloid HL-60 (Blue), lymphoid Jurkat (red) and HeLa (green) using *CEBPG* (middle) and *SLC7A10* (bottom) as a viewpoint. **(d)** Normalized 4C-seq profile which represents the 8 common contact domains (median = 11.65Kb) observed in all cell lines investigated. **(e)** Quantitative analysis of 4C-seq data to distinguish interacting regions occurring at higher contact frequencies in *CEBPA*$^{pos+}$ cells (orange, n=3) compared to *CEBPA*$^{neg-}$ cells (blue, n=3). The viewpoint is marked as a dotted line which contacts *CEBPA* (gray bar) with a significant higher contact frequency of more than 250 reads per million (FDR <0.05) in *CEBPA*$^{pos+}$ vs. *CEBPA*$^{neg-}$ cell lines.

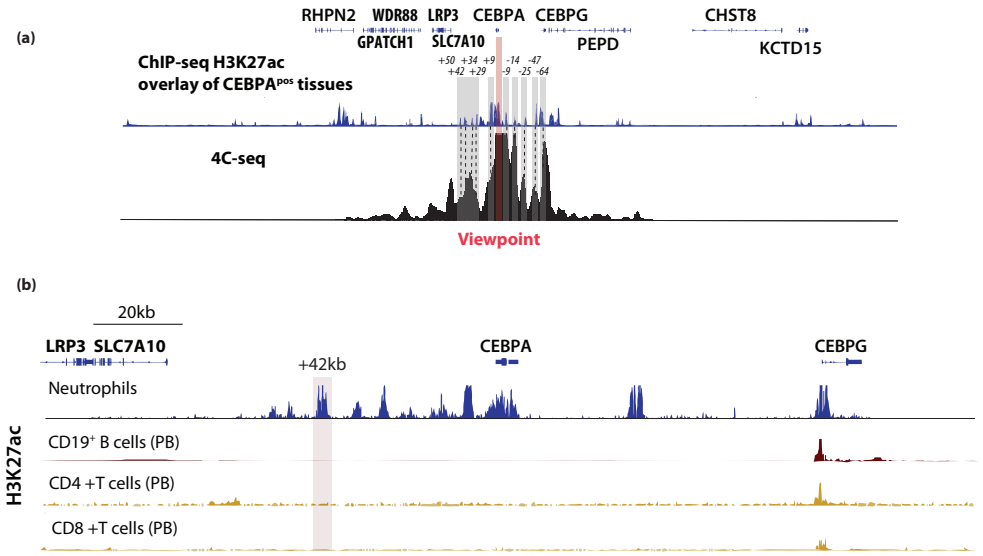


Fig.S2 : Overlap between potential enhancers and intra-TAD contact domains that interact with *CEBPA*. (a) Superimposed H3K27ac ChIP-seq from 14 *CEBPA* expressing tissues with 4C-seq profile showing potential enhancers allocated within contact domains. (b) ChIP-seq H3K27ac of human neutrophils, CD4+ T-cells, CD8+ T-cells, CD19+ B cells.

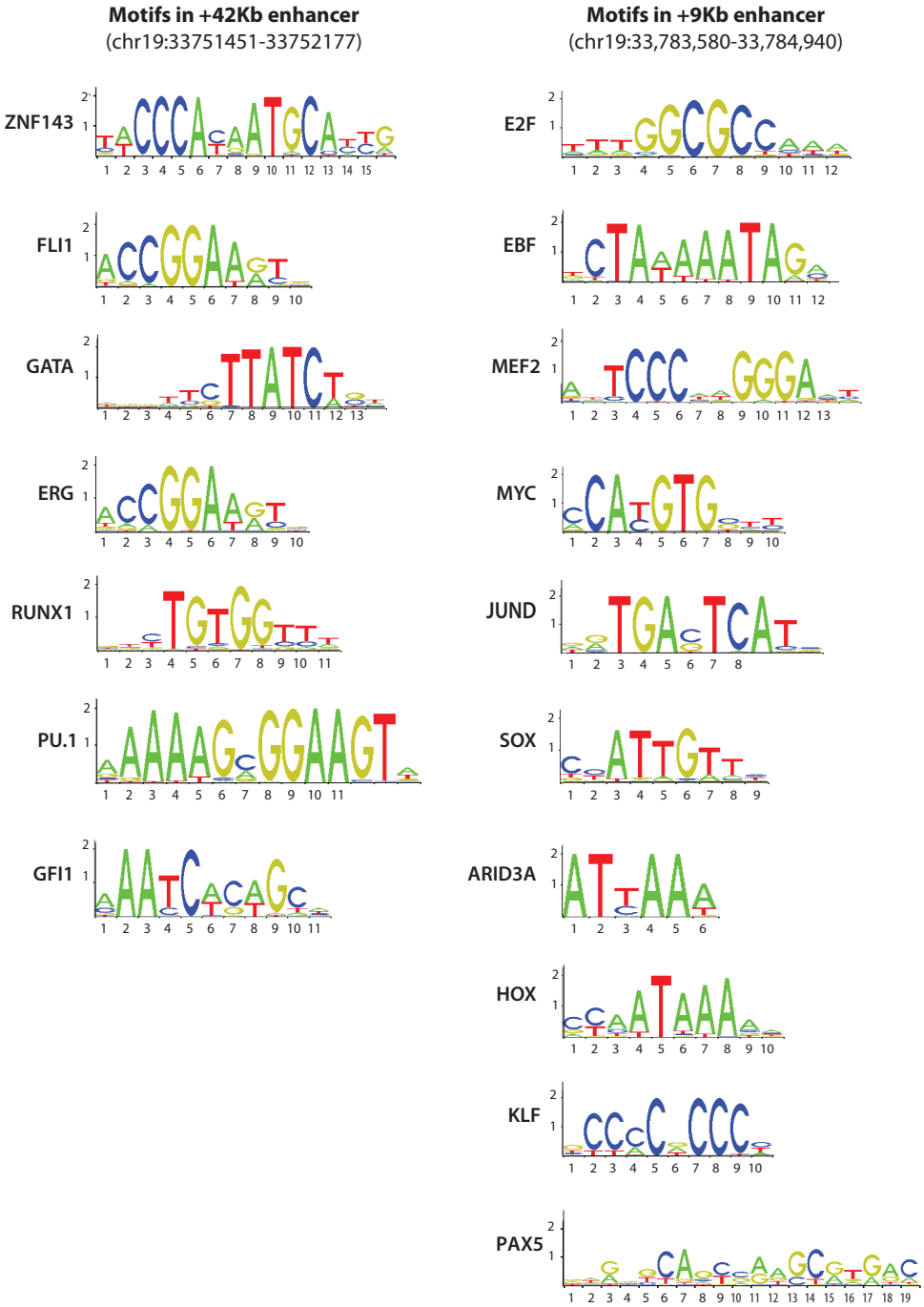


Fig.S3 : Selective haematopoietic specific transcription factor motifs in +9kb and +42kb enhancers using Jaspar motif analysis (Relative profile score threshold : 0.9).

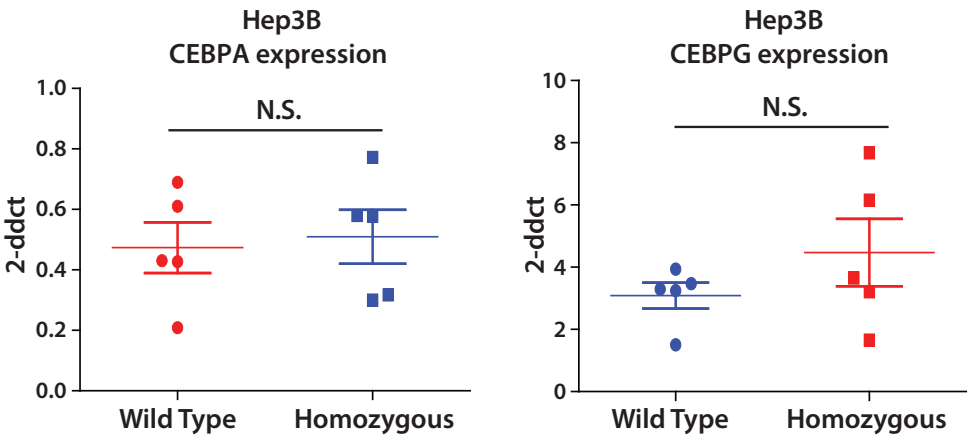
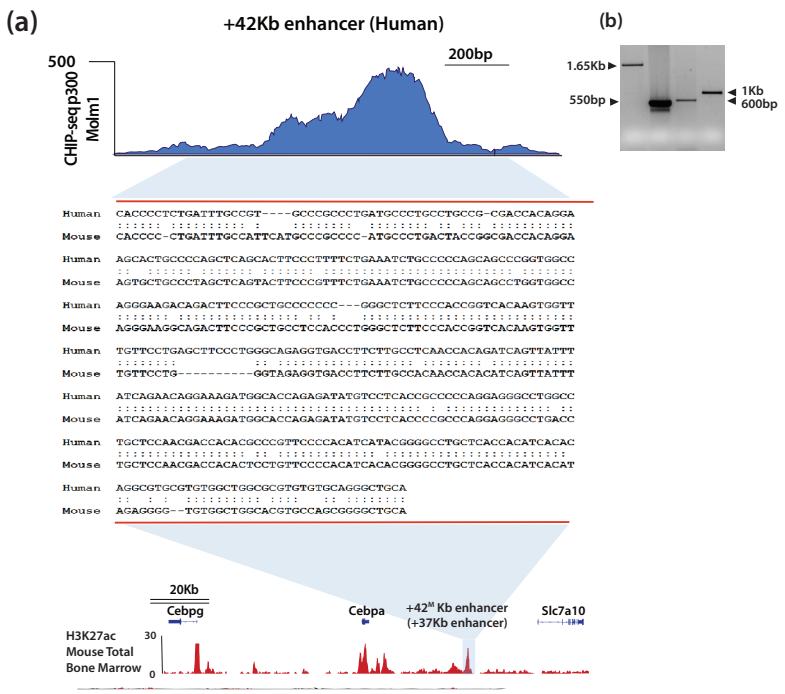


Fig.S4. CRISPR/Cas9 deletion of the +42Kb enhancer in Hep3B. Wild type (n=5) and Homozygous (n=5) clones were selected from single cell-based colonies formed in semi-solid medium. QPCRs for *CEBPA* and *CEBPG* showed no significant (N.S) difference (two-tailed T-test) between wild type and homozygous clones for both genes.



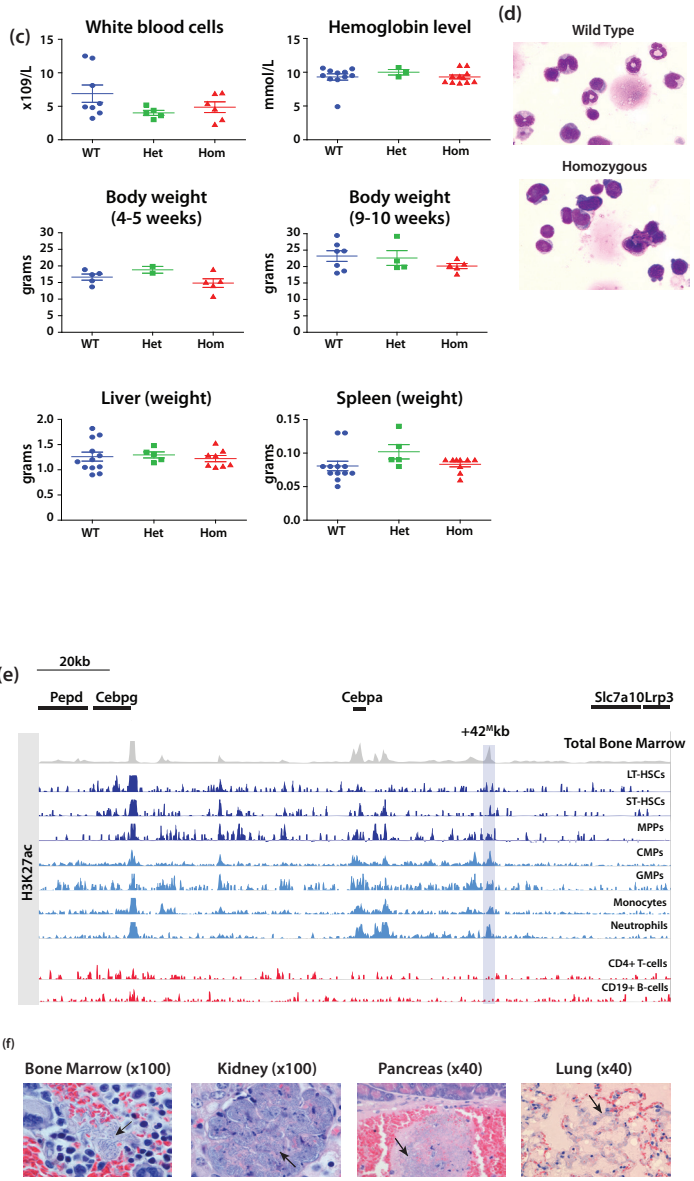


Fig.S5. Generation of the +42Kb enhancer deletion mouse model by CRISPR/Cas9. (**a**) Human +42Kb enhancer DNA sequence is 90% conserved with mouse +37Kb H3K27ac enhancer DNA sequence. (**b**) Genotyping by PCR showing three different lines with three different random NHEJ cuts generated by CRISPR/Cas9. Each cut represents a different amplicon of either 550bps, 600bps or 1Kb. (**c**) White blood cell count (x10^9/L) and Hb levels (grams) for wild type, heterozygous and homozygous mice using ABC blood counter. Total body, liver and spleen weight in grams. (**d**) May-Grünwald-Giemsa staining of bone marrow cytospins of wild type and homozygous mice. **(e)** H3K27ac ChIP-seq in LT-HSCs, ST-HSCs, MPPs, CMPs, GMPs, Neutrophils, Monocytes, CD4+ T cells, CD19 B cells[64]. **(f)** Presence of bacteria (arrows) in blood vessels of femoral bone marrow, glomeruli of the kidney, pancreas, and lung of a +42^{M}kb$^{-/-}$ mouse that died at 3 to 4 weeks after birth. In the lung, the presence of bacteria in capillaries of the pulmonary alveoli is associated with abundant proteinaceous fluid (oedema fluid) and fibrin strands in the alveolar lumina, suggesting damage to the alveolar wall and leakage of fluid from blood vessels into the air spaces. Hematoxylin and eosin stain. Objective magnification: 100X (bone marrow, kidney), 40X (pancreas, lung).

THESIS_Roberto_Avellino.indd 91 07-05-18 10:17

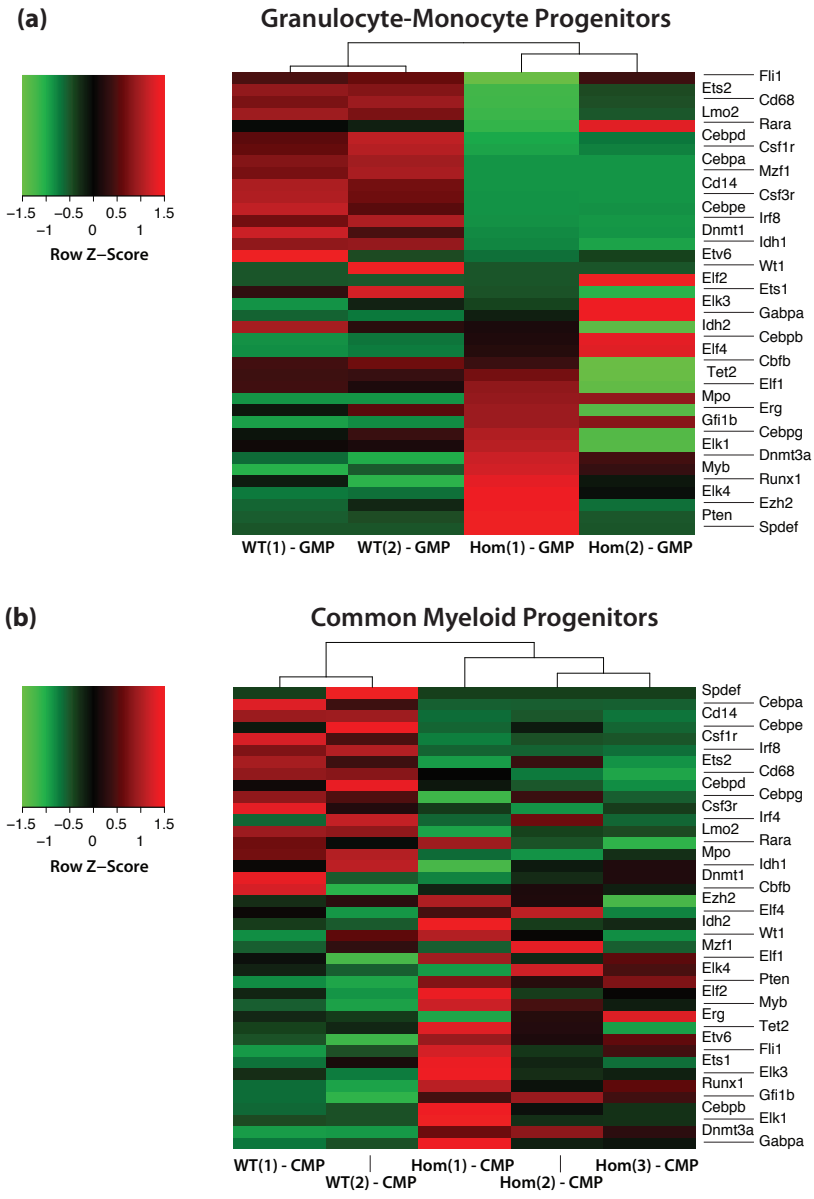


Fig.S6. Deregulation of myeloid associated genes in +42kb$^{-/-}$ GMPs and CMPs sorted fractions. Heatmap generated from RNA-seq representing selected myeloid associated genes (**a**) conducted on wild type (n=2) and homozygous (n=2) GMP sorted fractions, (**b**) on wild type (n=2) and homozygous (n=3) CMP sorted fractions. For both heatmaps a Z-score with a -/+ 1.5 standard deviation was used that represents upregulated (red) and downregulated (green) genes between wild type vs homozygous sorted fractions.

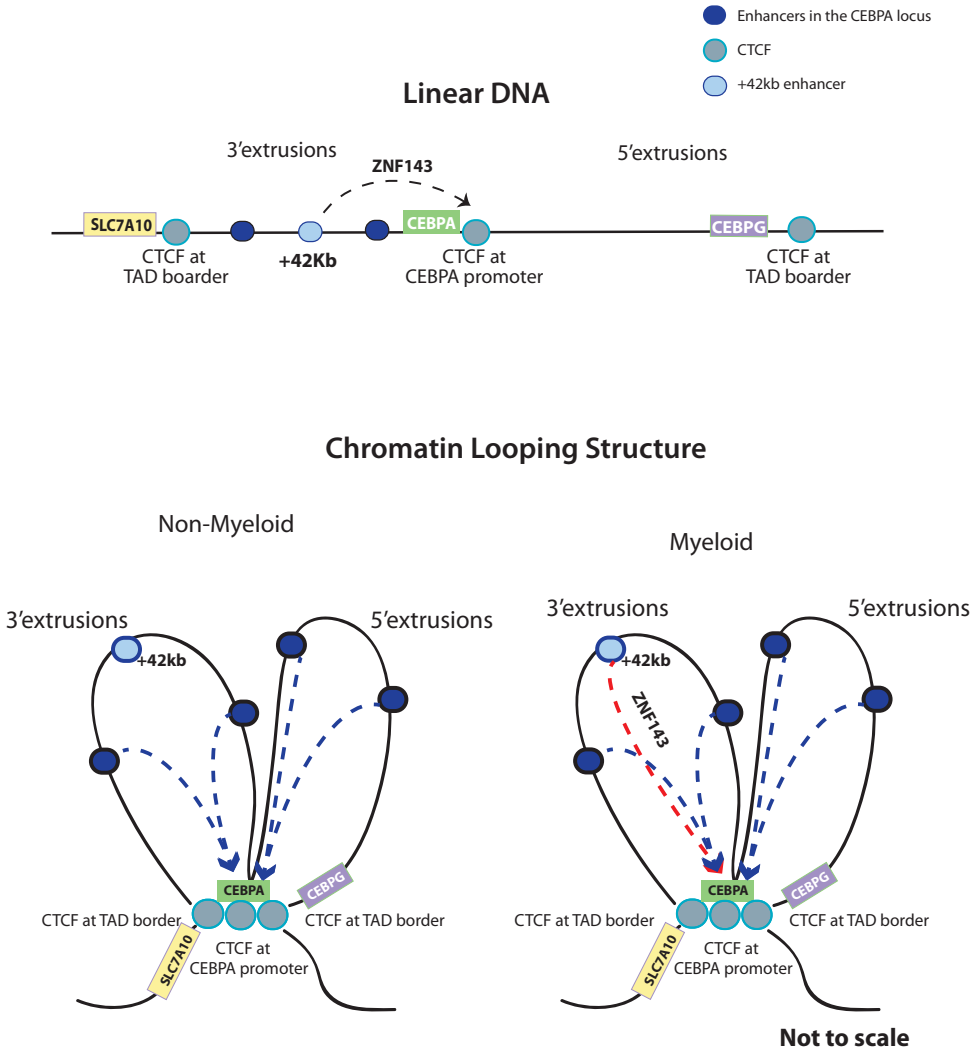


Fig.S7. Schematic figure showing the involvement of ZNF143 transcription factor in looping between the +42kb enhancer and the *CEBPA* promoter on (upper) linear DNA and in (lower) chromatin looping. protein levels and terminal neutrophilic differentiation.

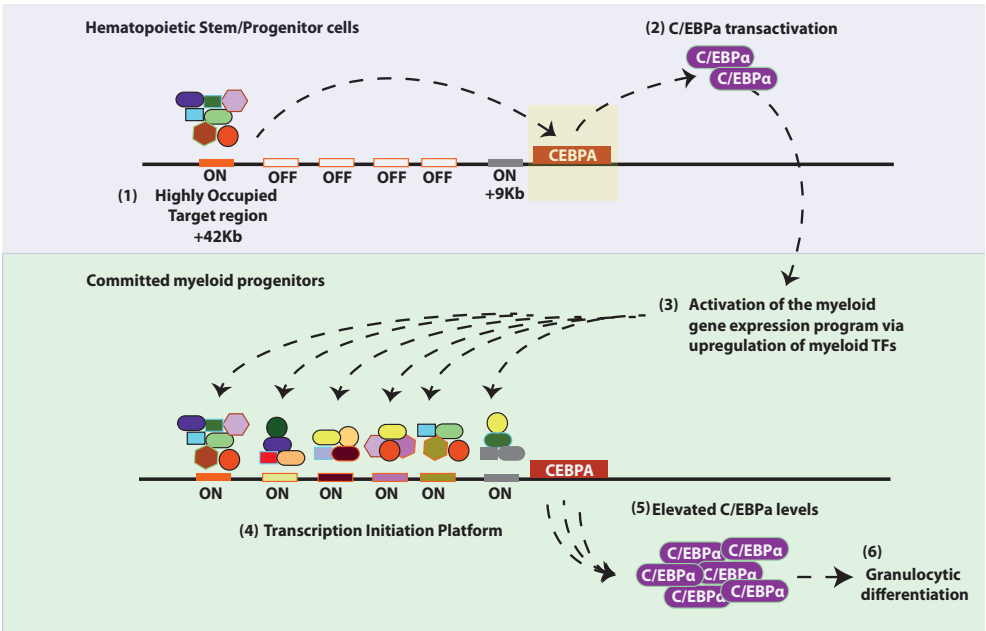


Fig.S8: Myeloid lineage-priming by the +42kb enhancer in HSCs initiate *CEBPA* transcription followed by the activation of the myeloid gene expression program. This induces myeloid commitment and activation of the *CEBPA* locus, leading to increased C/EBPα protein levels and terminal neutrophilic differentiation.

REFERENCES FOR SUPPLEMENTARY MATERIAL

58. van de Werken HJG, Landan G, Holwerda SJB, Hoichman M, Klous P, Chachik R, Splinter E, Valdes-Quezada C, Öz Y, Bouwman, B.A.M, et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* 2012(b); 9, 969–972.
59. Y. Benjamini, Y.H. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing When researchers tend to select pursuing multiple the (statistically) and support of conclusions. An unguarded use in a greatly results of single-inference inc. *J. R. Stat. Soc. Ser. B* 2015; 57, 289–300.
60. R Core Team. R: A Language and Environment for Statistical Computing. (2015)
61. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, Kim JS. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*. 2014; 1, 132-41.
62. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31, 166-169 (2015).
63. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*. 2014; 15, 550.
64. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, Zaretsky I, Jaitin DA, David E, Keren-Shaul H, Mildner A, Winter D, Jung S, Friedman N, Amit I. Immunogenetics. Chromatin state dynamics during blood formation. *Science*. 2014; 22, 943-9.

4

THESIS_Roberto_Avellino.indd 96 07-05-18 10:17

CEBPA-ENHANCER TARGETING MUTATIONS ARE INFREQUENT IN ACUTE MYELOID LEUKEMIA

Roberto Avellino[1], Claudia Erpelinck[1], Elodie C.G. Stoetman[1], Ruud Delwel[1]

[1]Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands

Work in progress

THESIS_Roberto_Avellino.indd 97 07-05-18 10:17

ABSTRACT

The non-coding genome hosts the major part of regulatory enhancer elements required for transcriptional control. Chromosomal abnormalities and point mutations in the non-coding genome are becoming more evident as disease-causing mutations in cancer, including acute myeloid leukemia (AML). Many mutations in AML generate oncogenic mechanisms and block neutrophilic differentiation by interfering with the neutrophil differentiation transcription factor C/EBPα. Low *CEBPA* levels are commonly observed in acute myeloid leukemia (AML), thus we hypothesized that point mutations or cryptic chromosomal abnormalities at previously identified *CEBPA*-enhancers decrease are responsible for low transcriptional output. A specific AML subgroup with relative low *CEBPA* expression levels, *CEBPA*^{LOW}(n=84), was selected from a cohort of 640 AML patients based on ranking of *CEBPA* levels from microarray data. Point mutations were not found in the whole locus in all patients investigated, whereas gross deletions occurring across multiple chromosomes were observed in one patient. We observe that deletion breakpoints do not happen randomly but occur at CTCF sites; located either at borders or inside topological associated domains (TADs). Deletions involved loss of genes important for differentiation, DNA replication, recombination and genes encoding architectural proteins involved in genome maintenance. Genes outside the deletion breakpoints were also deregulated due to disturbed insulated neighbor-hoods between TADs. Our results indicate that the *CEBPA*-enhancers are not targeted by DNA mutations except in a patient harboring a high rate of genomic instability accompanied by large deletions with breakpoints occurring at CTCF binding site located at topological associated domains, which are transcriptional domains with high susceptibility for DNA breakage.

INTRODUCTION

The folding and packaging of chromatin within the nuclear space is fundamental for gene regulation[1]. At the highest organizational level of chromatin, chromosomes are segregated into discreet chromosome territories (CT) in interphase nuclei[2]. Chromosomes are composed of condensed chromatin sub-divided into megabase-scale domains known as topological associated domains (TADs)[3, 4]. These TADs contain genes that engage with corresponding cis-regulatory modules (or enhancers) at relatively close proximity, forming looping structures to facilitate gene regulation [5]. Architectural proteins, such as CTCF and cohesin, form complexes with DNA and they form borders at TADs to secure and restrain their regulatory content from any interactions with genes or enhancers outside the neighborhood [6-9]. Such organization allows for a few genes to be co-regulated within the same insulated neighborhood, thus providing an efficient system to transcribe genes simultaneously and decreasing the risk of global gene deregulation by a one-hit mutation [9, 10]. Point mutations or structural variants in non-coding regions of the genome interfere with the organization of the nucleus and disturbs gene regulation[11]. Structural variants, such as chromosomal inversions, duplications and translocations, cause TAD reconfiguration that disturbs insulated neighborhoods and relocate ectopic enhancers close to or away from genes leading to aberrant gene expression programs [12, 13]. Such mechanisms have been revealed in different types of congenital disorders as well as in wide variety of tumors [12, 14-16]. Point mutations in transcription factor DNA binding motifs can either lead to low affinity transcription factor (TF) binding or they can generate a binding consensus site for a TF[13, 17-19].

The neutrophilic transcription factor C/EBPα is a common target for deregulation in AML. Many oncogenic mechanisms modulate *CEBPA* mRNA levels or compromise its protein function[20]. *CEBPA* is located in a 240kb TAD together with *CEBPG* [21]. *Cebpa* expression levels in the bone marrow are mainly upregulated by the autonomous +37kb enhancer (+42kb human homolog enhancer) in early myeloid-biased progenitors. Upon myeloid commitment and differentiation, additional enhancers become active in the *CEBPA* locus resulting in a further increase of *CEBPA* expression[21]. We aimed to investigate the *CEBPA* locus for possible DNA mutations that disrupt enhancer function in AML patients. No mutations were found in the *CEBPA* locus except for one patient that exhibits bi-allelic deletion of *CEBPA*. This patient also exhibits several TAD deletions and duplications. In conclusion, this is the first patient to be reported that exhibit *CEBPA* null mutations in AML and our findings indicate that TAD borders are fragile sites that increase susceptibility for chromosomal abnormalities, possibly by involving a common factor which might be responsible for TAD maintenance during DNA replication and transcription.

MATERIALS AND METHODS

Patient samples and Sanger sequencing

AML patient samples were collected from three AML cohorts (n=641) that were part of the HOVON study (Netherlands, Austria and Germany) and were processed as previously described. Genomic DNA was harvested from all patients in this cohort. Two sets of PCR primers covering the +42kb enhancer were designed for sequencing (F1: CATCAGATCTGGAGGACGCC/
R1: GGTATTCCCTGTGGGGAAGC; F2:CTCTGAGTGCAGAGAGGAGCC/
R2: CATCAGATCTGGAGGACGCC). Amplicons from each patient were sequenced on ABI 3500 sanger sequencing machine and sequences were analysed manually for mutations.

Custom Capture DNA-seq.

A custom probe library that covers 2.3 mega-basepairs on chromosome 19 covering the *CEBPA* locus was designed from Nimblegen. In brief, 100ngs of genomic DNA was subjected to enzymatic shearing followed by an end-repair and A-tailing reaction. Illumina adaptors were ligated and a double size selection for the optimal size of DNA to be sequenced was carried out. The right DNA size was confirmed using a bioanalyzer (Agilent) and DNA was hybridized for 16-20 hours onto probe captures. The captured DNA was cleaned and amplified again before run on Illumina 2500 sequencing platform. See Supplementary methods for analysis.

RESULTS

CEBPA expression-based ranking identifies an AML subgroup without or low *CEBPA* level

We hypothesized that low *CEBPA* expression levels in AML are driven by enhancer deregulation. To study our hypothesis, we studied *CEBPA* mRNA expression using a microarray dataset available for a cohort of 641 AML patients. Based on different logarithmic cutoff values, we stratified the patients based on *CEBPA* expression levels as follows: *CEBPA*^{LOW} group (cut-off: -1 to -4); *CEBPA*$^{LOW\text{-}MED}$ (cutoff: 0 to -1); *CEBPA*$^{MED\text{-}HIGH}$ (cutoff: 0 to 1) and *CEBPA*^{HIGH} (cutoff: 1 to 3) (Fig.1a). The *CEBPA*^{LOW} subgroup comprises 84 patients, of which 47%, had a known recurrent AML-associated abnormality including translocation t(8,21) (n=20), *CEBPA*^{methylated} (n= 11)[22], AMLs with chromosome 3q26 abnormalities with EVI1 overexpression (n=4) and a subset of AML patients with FLT3-ITD (n= 2) (Fig.1b). The remaining *CEBPA*^{LOW} cases (53%) did not belong to any AML subtype of known recurrent abnormalities. We investigated by Sanger nucleotide sequencing for potential mutations in the +42kb enhancer, which was previously characterized as an indispensable *CEBPA*-enhancer for neutrophilic differentiation. Within the conserved +42kb *CEBPA*-enhancer (~500bps) no mutations were found in any of the 642 AML cases studied,which is in line with a previous analysis that was conducted on a small number of AML patients (n=110)[23-25]. These findings led us to investigate for mutations occuring in other potential regulatory elements within the *CEBPA* locus[21].

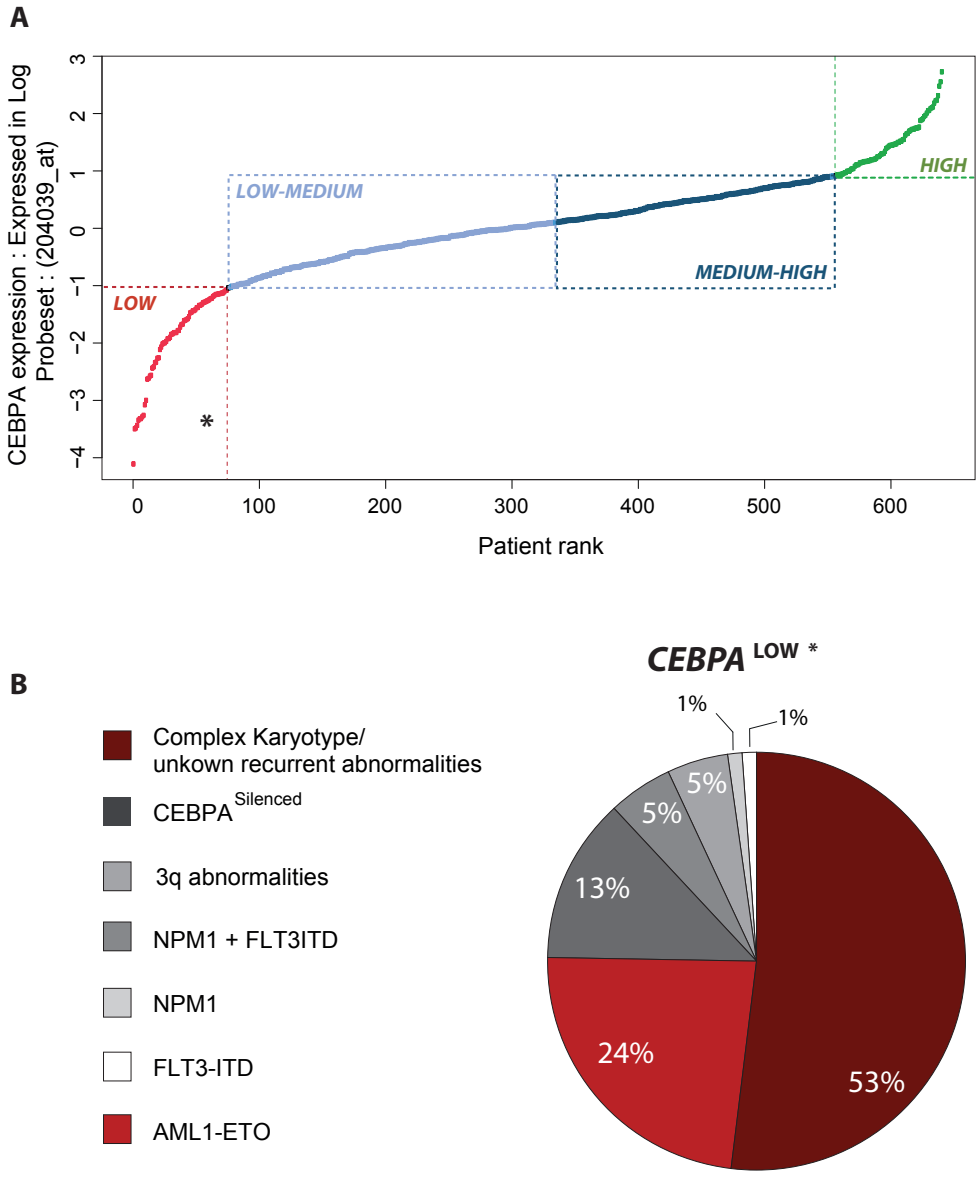


Figure 1. *CEBPA expression-based ranking analysis in AML recognizes 4 AML subgroups*. (a) Ranking of *CEBPA* expression based on mRNA microarray analysis classifies a cohort of 642 AML patients into four groups: low group (cut-off: -1 to -4); low-to-medium (cutoff: 0 to -1); medium-to-high (cutoff: 0 to 1) and high (cutoff: 1 to 3). The x-axis represents the patient rank and the y-axis represents the *CEBPA* levels expressed as log values. (b) Pie-chart showing the different genetically defined AML subgroups found in the *CEBPA* low AML subgroup.

DNA-sequencing and SNP array CGH reveal a bi-allelic deletion of the *CEBPA* locus in one AML patient

In addition to the +42kb enhancer, the *CEBPA* locus harbors at least 14 potential enhancer regions, which are all contained in a 240kb topological associated domain (TAD)[21]. The *CEBPA* TAD also includes the *CEBPG* gene, the *SLC7A10* promoter and partially *PEPD*. By applying custom capture DNA-sequencing (DNA-seq), we extended our strategy to screen 2.3 mega base-pairs on 19q11, including the *CEBPA* TAD (Fig.2a). Surprisingly, after filtering single nucleotide polymorphisms (SNPs), no mutations were detected except in one patient (#7120) that harbored no DNA-seq reads at the *CEBPA* TAD, suggesting a *CEBPA*-TAD null mutation (Fig.2b). This finding is in line with the *CEBPA* expression-based ranking results showing that this patient ranks the lowest from all the patients in the cohort (Supp. Fig.1a). Moreover, this patient sample was previously identified as a *CEBPA* silenced leukemia, without evidence of hyper-methylation of the *CEBPA* gene[22]. Clinical details of this patient claim a secondary AML following colorectal cancer, which potentially is a therapy-induced AML. It is clear that, when compared to other DNA-sequenced patients, the sequencing-read depth of the whole captured 2.3mb region is low at both centromeric and telomeric regions of the *CEBPA* locus (Fig.1b and Supp. Fig.1b). This suggests that one allele exhibits a larger deletion than the 2.3mb captured by DNA-sequencing. To map for precise deletion breakpoints on chromosome 19, we applied SNP array comparative genomic hybridization on genomic DNA of patient 7120 and validated the presence of a 242kb deletion at one allele, but also found a gross deletion of 28 megabases on the other chromosome 19q (Supp. Fig.2). Together these data suggest that mutations in the non-coding genome of the *CEBPA* locus are extremely rare, with one patient exhibiting a homozygous deletion of the *CEBPA* locus i.e. one allele exhibiting loss of heterozygosity of chromosome 19q and a 242kb deletion on the other allele that targets the *CEBPA* TAD.

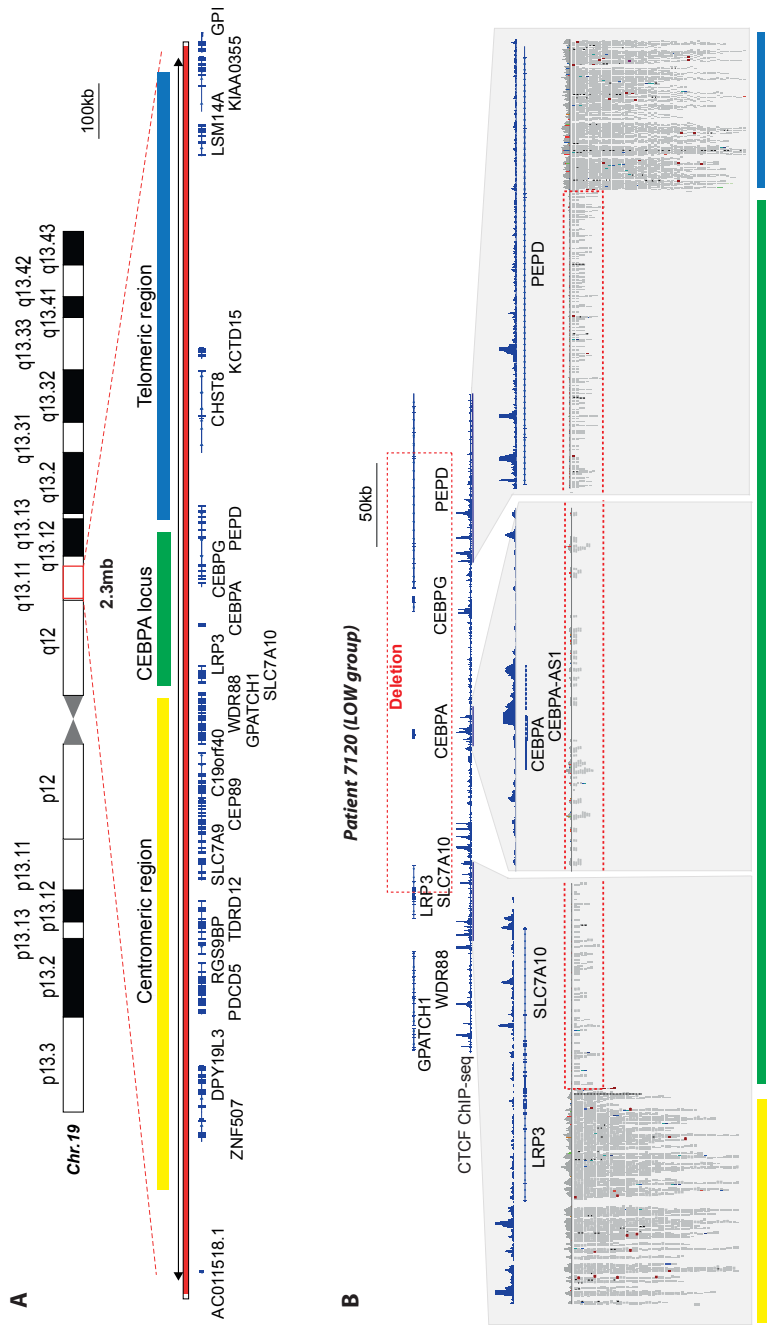


Figure 2. *DNA-seq reveals bi-allelic deletion of CEBPA in patient 7120*. (a) Capture beads for DNA-sequencing cover a region of 2.3mega basepairs on chromosome 19q13.12, including the *CEBPA* locus (green bar), telomeric region (blue bar) and centromeric region (yellow bar). (b). DNA-seq reads visualized on the integrated genome browser showing deletion of the entire *CEBPA* locus (green bar) with deletion breakpoints occurring close to CTCF regions from CTCF ChIP-seq tracks (ENCODE) in K562 myeloid cell line.

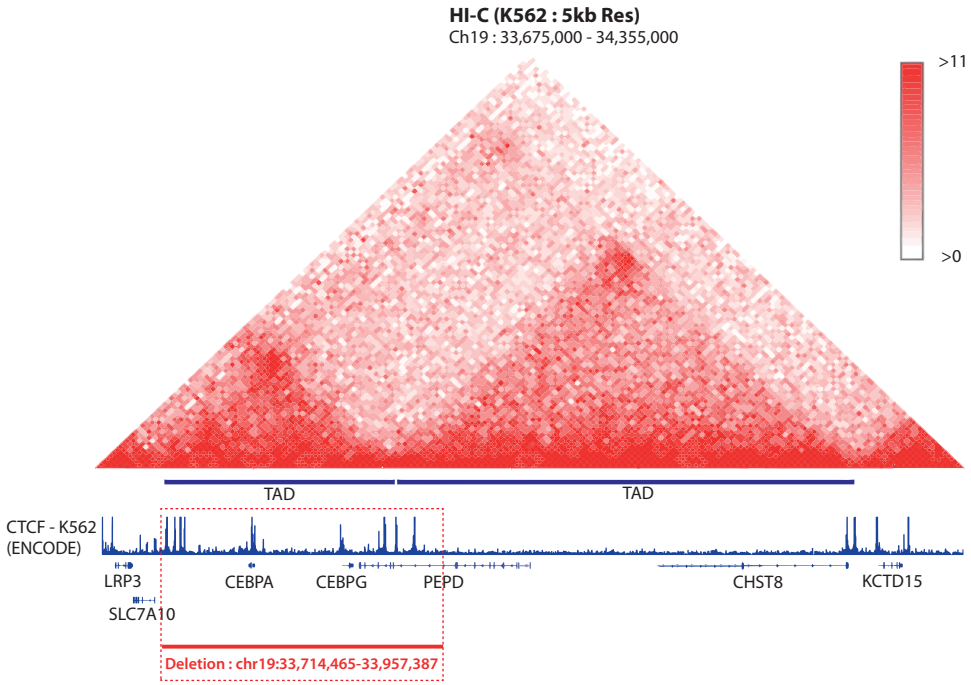


Figure 3. *Homozygous deletion breakpoints at the CEBPA locus occur at the TAD borders*. HI-C contact frequency from K562 cell line (at 5kb resolution) representing a genomic region on the long arm of chromosome 19 (Ch19 : 33,675,000 - 34,355,000) overlaid with CTCF Chip-seq tracks from K562 cell lines. Deleted region (chr19:33,714,465-33,957,387) is shown in red dotted box.

Deletion breakpoints occur close to or at CTCF-bound regions at borders or inside topological associated domains

To investigate whether breakpoints are associated with the chromosome architecture[26], we used a combinatorial approach by mapping the deletion breakpoints at the *CEBPA* locus from DNA-seq and SNP array, and overlaid them with contact frequencies derived from HI-C experiments conducted in K562 cells. In addition, CTCF ChIP-seq was used to define borders of TADs or sub-TADs. Interestingly, the deletion breakpoints occurred close to or at CTCF sites at the borders of the *CEBPA* TAD (Fig.3). In addition to the chromosome 19q deletions, more deletions were revealed on other chromosomes in the patient (7120) applying the SNP-array. We used the same combinatorial approach to investigate breakpoints on other chromosomes and found that they also occur close to or at CTCF bound TAD borders or sub-TADs (Fig.4). Deletion sizes varied from one TAD to several adjacent ones (See Supp. Table 4). In total, 5322 genes were found located inside deleted TADs. The majority of these deleted genes were heterozygous i.e. loss of heterozygosity (LOH). Since the number of deleted genes is large, it was difficult to assign the deleted genes to biological pathways.

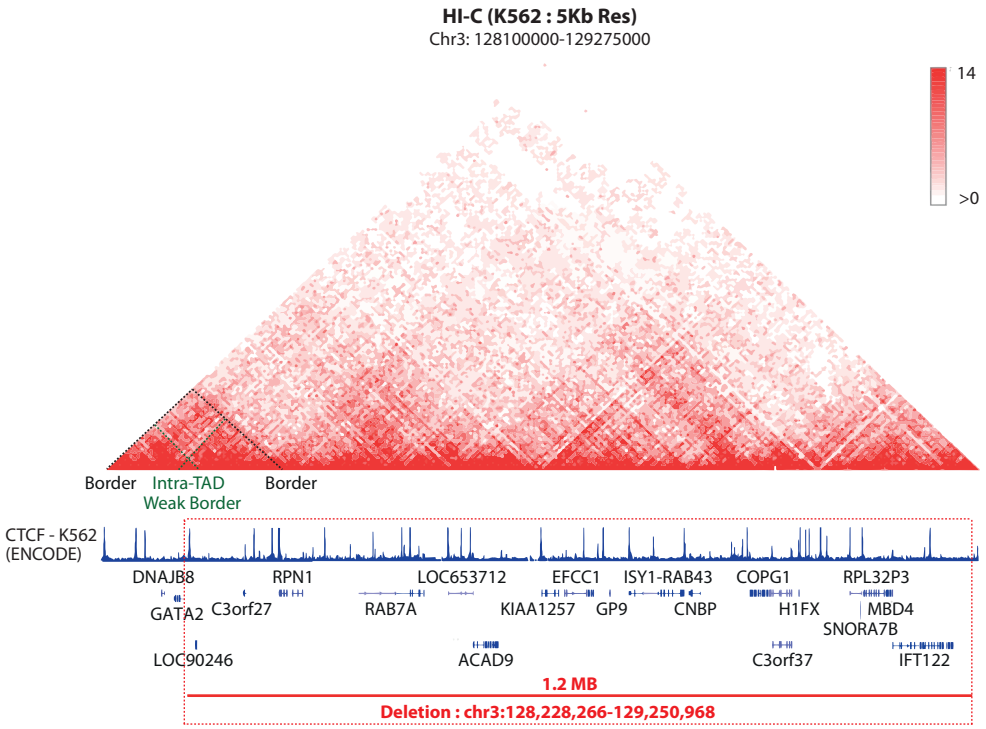


Figure 4. *Heterozygous deletion on chromosome 3q showing a gross deletion of 1.2 megabases involving several TADs and sub-TADs*. HI-C contact frequency from K562 cell line (at 5kb resolution) representing a genomic region on the long arm of chromosome 3 (chr3: 128100000-129275000) overlaid with CTCF Chip-seq tracks from K562 cell lines. Deleted region (chr3:128,228,266-129,250,968) is shown in red dotted box.

Many deleted genes encode proteins that belong to specific functions such as genome architecture (*CTCF, SMC1*), chromatin modifiers (*Ep300*, *EZH2*, *HDACs*, *MED21*, *KMT2* and 5, *KDM3*,7,8), DNA damage response (*APOBEC3*, *ERCC1* and 2), myeloid membrane markers (*CSF1R*, *CD14*, *CD33*), 100 long non-coding RNAs including HOTAIR and HOTTIP and transcription factors (*CEBPA*, *CEBPE*, *FOSB*, *FOX*, *HOX*, *IRF8*, *CBFB*) . Genome duplications were also observed, although at a minimal frequency when compared to deletions (Supp. Table 4 and Supp. Fig.4). In conclusion, the CTCF binding regions are potential fragile sites for chromosomal breakage leading to a high genomic instability profile and loss of heterozygosity to thousands of genes.

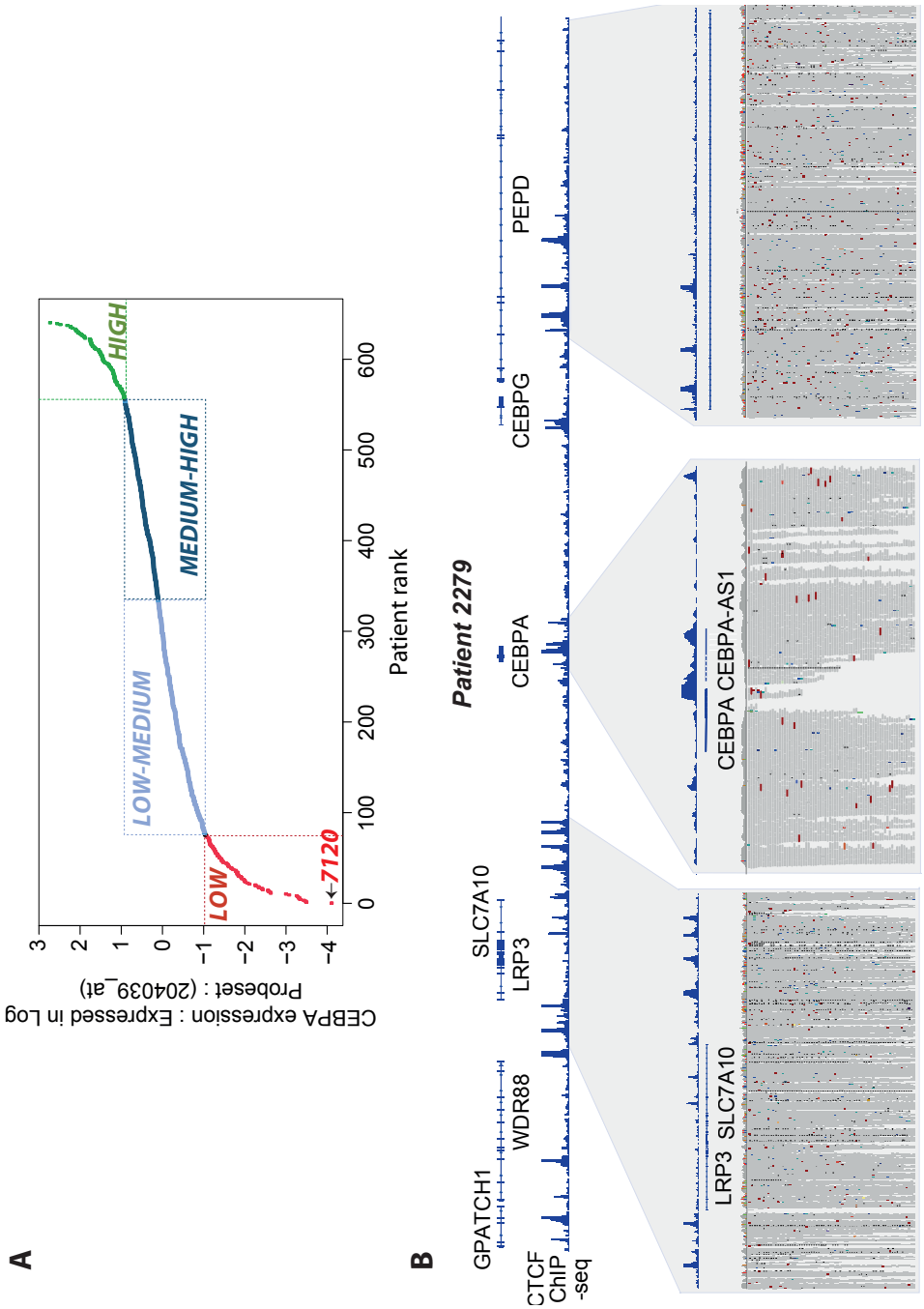
DISCUSSION

We here report that mutations in regulatory elements near *CEBPA* are highly infrequent. A subgroup of AML with low *CEBPA* expression levels was selected based on the hypothesis that low *CEBPA* expression levels in AML are caused by DNA mutations, such as insertions and deletions of enhancer DNA sequences. We predicted that mutations in enhancers might disturb specific TF binding motifs and hence abrogate or reduce the affinity of transcription factors and their co-activators to bind the enhancer and regulate *CEBPA* transcription. Examples of mechanisms for this sort of transcriptional deregulation have been reported previously. For example, mutations in an intergenic region on chromosome 9p13 attenuate the expression of the tumor suppressor gene *PAX5* in patients with pre-B-cell precursor acute lymphoblastic leukemia[27]. On the other hand, 7bp insertions of DNA sequences generate a MYB TF binding site in enhancers of proto-oncogenes such as LMO2 and TAL1, and elevate their expression levels in T-ALL [18]. Similar mechanisms were also reported in other non-hematopoietic tumors, indicating a more common mechanism used in different tumor types [28]. Although these reports show gene deregulation by enhancer-targeted mutations, the *CEBPA*-enhancers are most probably targeted by different oncogenic mechanisms in leukemia.

Breakpoints generated by chromosomal abnormalities such as translocations, deletions or inversions are commonly found in the coding part of genes in AML [29-31]. However, breakpoints occur less frequent in the non-coding part of the genome [12]. By combining two technologies, DNA-seq and SNP-array, we confirmed that the deletion at the *CEBPA* locus involves both alleles; one allele affected as part of a large deletion on the long arm of chromosome 19 (MB) and the other allele exhibited a smaller deletion occurring close to boundaries of the *CEBPA* TAD. In fact, we found many more deletions, most of which occupied whole TADs or sub-TADs. The breakpoints were located close to or at a CTCF binding site. There is supporting evidence that CTCF binding occurs next to genomic fragile sites, which are highly susceptible for chromosomal breakage and instability[26]. The genomic instability observed in this patient is possibly caused as a secondary event post-therapy that was administered for the primary solid tumor. Increase in therapy-related DNA damage might have mutated genes that safeguard the genome such as TP53, as well as genes involved cell cycle checkpoints, DNA replication, recombination and repair. However, this can only be confirmed by applying whole genome exome-sequencing and search for potential mutated candidates responsible for genomic instability [32-34].

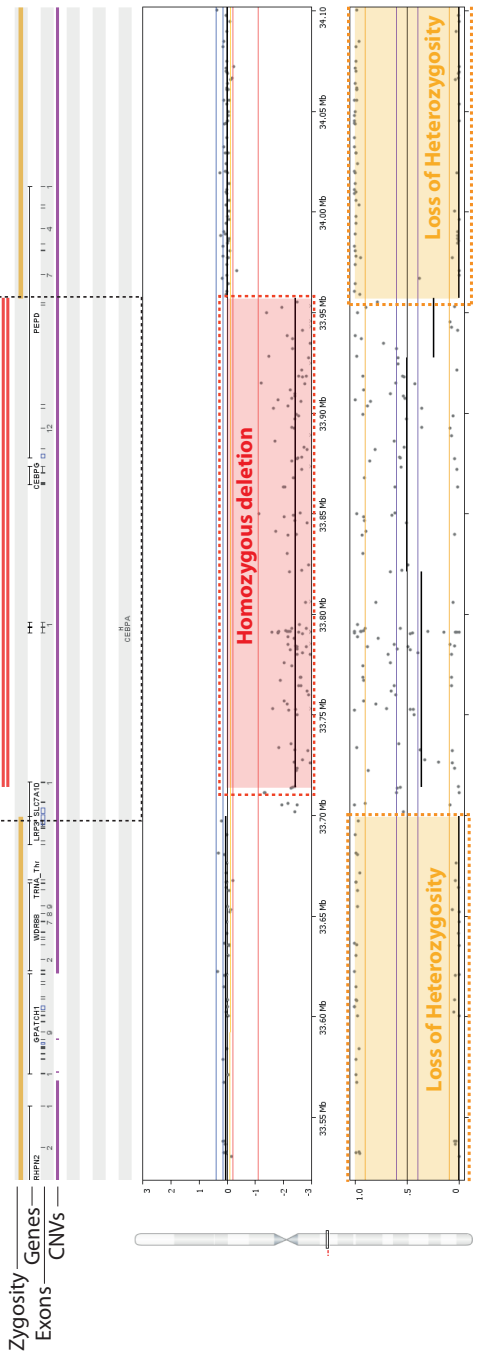
We conclude that mutations in the non-coding genome of the *CEBPA* locus are infrequent in AML. This suggests that other unknown oncogenic mechanisms in the *CEBPA*^{LOW} subgroup interfere with transcriptional control of the gene. Oncoproteins such as AML1-ETO and EVI1 have been reported to recruit protein complexes that reverse active chromatin and repress genes [35]. Thus, it is highly possible that low *CEBPA* expression levels in 47% of

the *CEBPA*^{LOW} subgroup are directly linked to the oncoprotein generated by the underlying abnormality including AML1-ETO in t(8;21), EVI1 in 3q abnormalities, mutations in coding sequence of DNMT3, and FLT3-ITD (internal tandem duplication) (Fig.1).The mechanism by which these oncoproteins deregulate *CEBPA* expression is yet to be determined. The other 53% cases with no associated abnormality and with no detected mutations in the *CEBPA*-TAD, could be potentially targeted by other causative mechanisms that require further investigations. It is of major importance to understand how oncoproteins such as AML1-ETO and EVI1 hijack differentiation-related enhancers and deregulate genes involved in myeloid differentiation, in order to provide more insights on epigenetic mechanisms and improve therapeutic strategies.In the next chapter of this thesis, we investigated thoroughly the potential mechanism of enhancer deregulation in AML by studying the epigenetic influence of AML1-ETO on *CEBPA* deregulation.

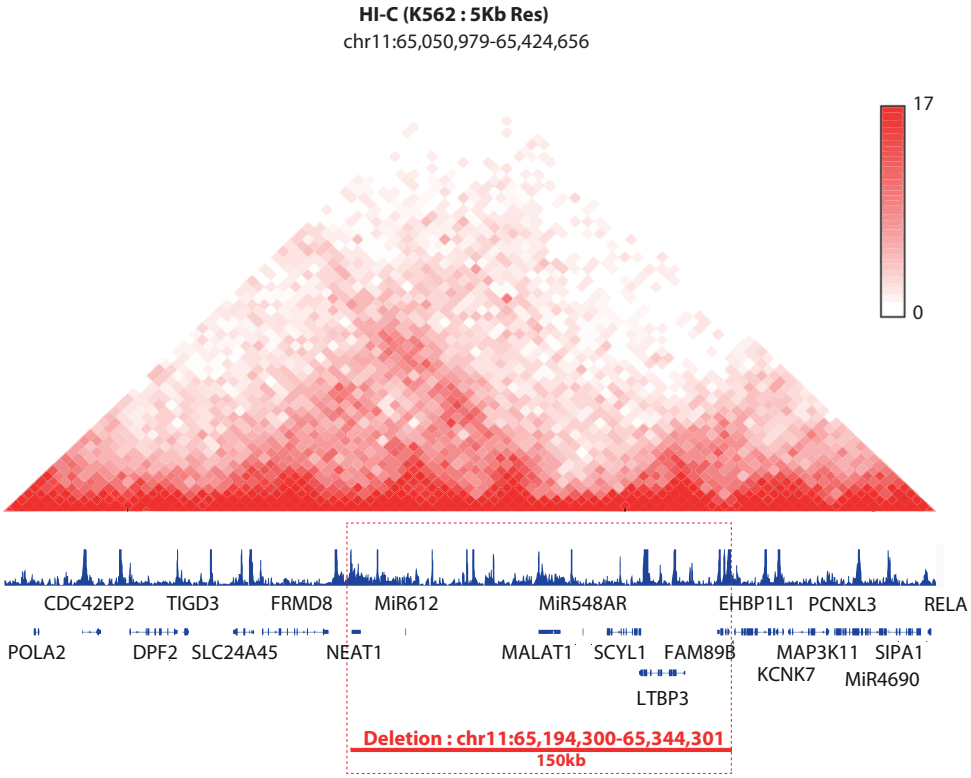


Supplementary Figure 1. (a) Ranking of *CEBPA* expression showing placement of patient 7120 in the *CEBPA*^{LOW} subgroup. (b) DNA-seq reads from patient 2279 show no deletions.

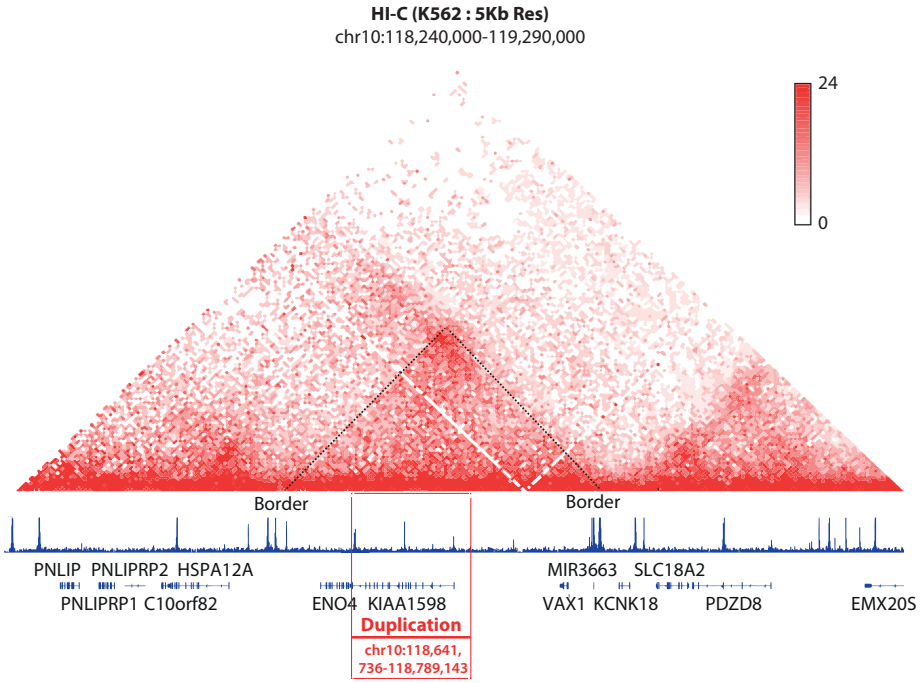
Supplementary Figure 2



Supplementary Figure 2. Snapshot of SNP array showing the *CEBPA* TAD on chromosome 19. The presented region showed a biallelic chromosomal loss at the *CEBPA* TAD shown in red, flanked by monoallelic chromosomal loss, telomeric and and centromeric to the *CEBPA* TAD shown in light orange.



Supplementary Figure 3. *Heteroygous deletion on chromosome 11q showing a deletion of 150kb*. HI-C contact frequency from K562 cell line (at 5kb resolution) representing a genomic region on the long arm of chromosome 11 (65,050,979-65,424,656) overlaid with CTCF Chip-seq tracks from K562 cell lines. The deleted region (chr11:65,194,300-65,344,301) is shown in red dotted box that covers a region of 150kb encompassing two sub-TADs.



Supplementary Figure 4. HI-C contact frequency from K562 cell line (at 5kb resolution) representing a genomic region on the long arm of chromosome 10 (118,240,000-119,290,000) overlaid with CTCF Chip-seq tracks from K562 cell lines. The amplified/duplicated region (118,641,736 - 118,789,143) is shown in red dotted box that covers a region of 148kb occurring within one sub-TAD.

SUPPLEMENTARY METHODS FOR CUSTOM CAPTURE DNA-SEQ

Alignment

Quality control of the sequencing run was assessed by using ShortRead. The alignment to the reference genome was accomplished using the Burrows-Wheeler Alignment Tool (BWA-MEM) (1). The Human Genome version 19 (hg19) (Santa Cruz (UCSC), CA) was used as the reference genome. SAMtools was used to sort and index the compressed binary format (BAM) files outputted by BWA.

Variant Calling

Variant calling was performed on the outputted BAM files using the algorithms: GATK MuTect(2), GATK MuTect2(2), GATK UnifiedGenotyper (3), SAMtools/BCFtools (4), Varscan2 (5), GATK IndelGenotyperV2 (6) and Pindel (7). All tools from GATK were run with the BadCigar filter. Each variant file was subsequently annotated by Annovar and merged to create one variant file per sample.

Variant Analysis

Variants were filtered according to the regions used in the capture. All variants found to be non-synonymous or flagged as known germline variants by either the snp137 or snp138 database (Santa Cruz (UCSC), CA) except if also flagged as somatic by the COSMIC70 database, were removed. Any variants with lower than 4 supporting high quality reads, an allele frequency lower than 15% high quality reads or found within more than 50% of samples were removed. The remaining variants were separated into the two groups based upon low or high *CEBPA* expression. Any variant found to be significantly more present in either group (p-value < 0.05) were manually analyzed using the integrative genome viewer (IGV).

ROH Analysis

Runs of homozygosity (ROH) were detected using the hidden Markov model from the H3M2 tool for whole exome with default parameters (8).

1. Li H. and Durbin R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics, 25:1754-60. [PMID: 19451168]
2. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G: Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013, 31: 213-219. 10.1038/nbt.2514.
3. DePristo M. A., Banks E., Poplin R., et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics. 2011;43(5):491–501. doi: 10.1038/ng.806.
4. Li et al. (2009) Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–2079. doi: 10.1093/bioinformatics/btp352.
5. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Research 22:568-576
6. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA, 2010 GENOME RESEARCH 20:1297-303
7. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics. 2009 Nov 1;25(21):2865-71. Epub 2009 Jun 26.
8. Alberto Magi, Lorenzo Tattini, Flavia Palombo, Matteo Benelli, Alessandro Gialluisi, Betti Giusti, Rosanna Abbate, Marco Seri, Gian Franco Gensini, Giovanni Romeo, Tommaso Pippucci; H3M2 : detection of runs of homozygosity from whole-exome sequencing data , Bioinformatics, Volume 30, Issue 20, 15 October 2014, Pages 2852–2859, https://doi.org/10.1093/bioinformatics/btu401

REFERENCES

1. Felsenfeld, G., et al., *Chromatin structure and gene expression.* Proc Natl Acad Sci U S A, 1996. **93**(18): p. 9384-8.
2. Cremer, T. and M. Cremer, *Chromosome territories.* Cold Spring Harb Perspect Biol, 2010. **2**(3): p. a003889.
3. Dixon, J.R., et al., *Topological domains in mammalian genomes identified by analysis of chromatin interactions.* Nature, 2012. **485**(7398): p. 376-80.
4. Zuin, J., et al., *Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells.* Proc Natl Acad Sci U S A, 2014. **111**(3): p. 996-1001.
5. Lieberman-Aiden, E., et al., *Comprehensive mapping of long-range interactions reveals folding principles of the human genome.* Science, 2009. **326**(5950): p. 289-93.
6. Carretero, M., S. Remeseiro, and A. Losada, *Cohesin ties up the genome.* Curr Opin Cell Biol, 2010. **22**(6): p. 781-7.
7. Gause, M., C.A. Schaaf, and D. Dorsett, *Cohesin and CTCF: cooperating to control chromosome conformation?* Bioessays, 2008. **30**(8): p. 715-8.
8. Parelho, V., et al., *Cohesins functionally associate with CTCF on mammalian chromosome arms.* Cell, 2008. **132**(3): p. 422-33.
9. Rao, S.S.P., et al., *Cohesin Loss Eliminates All Loop Domains.* Cell, 2017. **171**(2): p. 305-320 e24.
10. Wendt, K.S. and J.M. Peters, *How cohesin and CTCF cooperate in regulating gene expression.* Chromosome Res, 2009. **17**(2): p. 201-14.
11. Krijger, P.H. and W. de Laat, *Regulation of disease-associated gene expression in the 3D genome.* Nat Rev Mol Cell Biol, 2016. **17**(12): p. 771-782.
12. Groschel, S., et al., *A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia.* Cell, 2014. **157**(2): p. 369-381.
13. Dreijerink, K.M.A., et al., *Enhancer-Mediated Oncogenic Function of the Menin Tumor Suppressor in Breast Cancer.* Cell Rep, 2017. **18**(10): p. 2359-2372.
14. Franke, M., et al., *Formation of new chromatin domains determines pathogenicity of genomic duplications.* Nature, 2016. **538**(7624): p. 265-269.
15. Kraft, K., et al., *Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice.* Cell Rep, 2015.
16. Lupianez, D.G., et al., *Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions.* Cell, 2015. **161**(5): p. 1012-1025.
17. Li, Z., et al., *APOBEC signature mutation generates an oncogenic enhancer that drives LMO1 expression in T-ALL.* Leukemia, 2017. **31**(10): p. 2057-2064.
18. Mansour, M.R., et al., *Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element.* Science, 2014. **346**(6215): p. 1373-7.
19. Rahman, S., et al., *Activation of the LMO2 oncogene through a somatically acquired neomorphic promoter in T-cell acute lymphoblastic leukemia.* Blood, 2017. **129**(24): p. 3221-3226.
20. Avellino, R. and R. Delwel, *Expression and regulation of C/EBPalpha in normal myelopoiesis and in malignant transformation.* Blood, 2017. **129**(15): p. 2083-2091.
21. Avellino, R., et al., *An autonomous CEBPA-enhancer specific for myeloid-lineage priming and neutrophilic differentiation.* Blood, 2016. **127**(24): p. 2991-3003.
22. Wouters, B.J., et al., *Distinct gene expression profiles of acute myeloid/T-lymphoid leukemia with silenced CEBPA and mutations in NOTCH1.* Blood, 2007. **110**(10): p. 3706-14.

23. Cooper, S., H. Guo, and A.D. Friedman, *The +37 kb CEBPA-enhancer Is Critical for Cebpa Myeloid Gene Expression and Contains Functional Sites that Bind SCL, GATA2, C/EBPalpha, PU.1, and Additional Ets Factors.* PLoS One, 2015. **10**(5): p. e0126385.
24. Cancer Genome Atlas Research, N., et al., *The Cancer Genome Atlas Pan-Cancer analysis project.* Nat Genet, 2013. **45**(10): p. 1113-20.
25. Cancer Genome Atlas Research, N., et al., *Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia.* N Engl J Med, 2013. **368**(22): p. 2059-74.
26. Canela, A., et al., *Genome Organization Drives Chromosome Fragility.* Cell, 2017. **170**(3): p. 507-521 e18.
27. Puente, X.S., et al., *Non-coding recurrent mutations in chronic lymphocytic leukaemia.* Nature, 2015. **526**(7574): p. 519-24.
28. Abraham, B.J., et al., *Small genomic insertions form enhancers that misregulate oncogenes.* Nat Commun, 2017. **8**: p. 14385.
29. Nimer, S.D. and M.A. Moore, *Effects of the leukemia-associated AML1-ETO protein on hematopoietic stem and progenitor cells.* Oncogene, 2004. **23**(24): p. 4249-54.
30. Chen, Z. and S.J. Chen, *RARA and PML genes in acute promyelocytic leukemia.* Leuk Lymphoma, 1992. **8**(4-5): p. 253-60.
31. Liu, P.P., et al., *Identification of the chimeric protein product of the CBFB-MYH11 fusion gene in inv(16) leukemia cells.* Genes Chromosomes Cancer, 1996. **16**(2): p. 77-87.
32. Godley, L.A. and R.A. Larson, *Therapy-related myeloid leukemia.* Semin Oncol, 2008. **35**(4): p. 418-29.
33. Ganser, A. and M. Heuser, *Therapy-related myeloid neoplasms.* Curr Opin Hematol, 2017. **24**(2): p. 152-158.
34. Fabiani, E., et al., *Clonal evolution in therapy-related neoplasms.* Oncotarget, 2017. **8**(7): p. 12031-12040.
35. Loke, J., et al., *RUNX1-ETO and RUNX1-EVI1 Differentially Reprogram the Chromatin Landscape in t(8;21) and t(3;21) AML.* Cell Rep, 2017. **19**(8): p. 1654-1668.

THESIS_Roberto_Avellino.indd 117 07-05-18 10:17

5

THESIS_Roberto_Avellino.indd 118 07-05-18 10:17

AML1-ETO REVERSES THE ACTIVE CHROMATIN STATE AND CONFORMATION OF A MYELOID SPECIFIC ENHANCER TO INHIBIT *CEBPA* EXPRESSION

Roberto Avellino[1], Claudia Erpelinck[1], Roger Mulet Lazaro[1], Leonie Smeenk[1], Michael Vermeulen[1], Anna Pickin[2], Michael Rehli[3,4], Claudia Gebhard[3,4], Eric Bindels[1], Connie Bonnifer[2], Ruud Delwel[1]

[1]Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands
[2]Institute for Cancer and Genomic Sciences, College of Medicine and Dentistry, University of Birmingham, B15 2TT Birmingham, UK.
[3]Department of Internal Medicine III, University Hospital Regensburg, Regensburg, Germany.
[4]Regensburg Centre for Interventional Immunology (RCI), Regensburg, Germany

Work in progress

THESIS_Roberto_Avellino.indd 119 07-05-18 10:17

ABSTRACT

Oncogene activation by recurrent genetic mutations rewires the chromatin states of genes responsible for bone marrow cell differentiation and proliferation. A diverse range of oncogenes target *CEBPA* transcriptional levels or its function in acute myeloid leukemia (AML). The mechanisms by which such oncogenes halt interfere with *CEBPA* are largely unknown. *CEBPA* expression levels are low in AML patients with the AML1-ETO (AE) fusion oncoprotein. We here show that AE directly influences *CEBPA* expression by binding to its canonical +42kb enhancer, reverses its active chromatin state at the enhancer and disrupts the three-dimensional topology in human AML patients and in the AE cell line model, Kasumi-1. Gene deregulation by AE involves the recruitment of histone de-acetyl transferases (HDACs), which might potentially explain the selective loss of H3K9ac and H3K27ac as well as the loss of *CEBPA* gene to enhancer interaction. Revealing further how AE interferes with histone acetyl-transferase enzymes responsible for H3K9ac and H3k27ac, may help to find a therapeutic platform to reactivate *CEBPA* expression in myeloid malignancies. Our study show how a fusion oncoprotein influences the dynamics of the myeloid differentiation program by hijacking a specific enhancer, reverses its active state, interferes with its chromatin conformation and reduces gene expression levels of the myeloid transcription factor, *CEBPA*.

INTRODUCTION

Acute myeloid leukemia (AML) is a heterogeneous malignancy with an aberrant differentiation program of myeloid cells in the bone marrow. It is becoming more evident that bone marrow progenitors require myeloid priming and commitment to develop the right molecular environment to initiate AML. Several studies have recently shown that substantial levels of *CEBPA*, the major driver of neutrophilic differentiation, are required to generate myeloid progenitors and initiate AML[1, 2]. In fact, up to our knowledge, there are no studies that report any *CEBPA* null mutations in AML patients[3-5] (except the patient we report in chapter 4 of this thesis), which supports the idea that an adequate *CEBPA* dosage is required for some degree of myeloid differentiation to occur and hence trigger the onset of AML.

Several mutations that occur in AML are linked to *CEBPA* deregulation; either by interfering with its expression levels and/or with its function. Such mutations frequently occur in transcription factors and other nuclear proteins that act upstream of *CEBPA*. One of the major upstream transcription factors of *CEBPA* in normal hematopoiesis is RUNX1[6, 7]. The normal function of RUNX1 is critical in the generation of hematopoietic stem cells (HSCs) from the hemogenic endothelium during development [8, 9]. RUNX1 has two important domains: the transcriptional activating domain at the C-terminal and the DNA binding domain at the N-terminal site of the protein. Investigations in conditional *Runx1* knockout mouse models show that that RUNX1 protein in the bone marrow acts as an upstream transcriptional co-activator of *Cebpa* and, upon deletion of *Runx1, Cebpa* levels are diminished[6]. Moreover, we showed that in human CD34+ hematopoietic stem and progenitor cells (HSPCs) RUNX1 physically binds to its consensus motifs located at the +42kb enhancer and at the gene promoter of *CEBPA[7]*, highly suggesting an important role of RUNX1-*CEBPA* regulatory axis in myelopoiesis.

In AML, RUNX1 is a common target for genomic mutations[10, 11]; either deregulated by recurrent chromosomal translocations generating partner fusion genes such as AML1-ETO (AE) also known as RUNX1-RUNX1T1, also known as) and AML1-MDS1-EVI1 (AME) or RUNX1-MDS1-EVI1 or by point mutations occurring in the RUNX1 coding sequence of both the DNA binding and the trans-activating domain[12-14]. Fusion genes involving RUNX1 exhibit breakpoints within the gene that result in the loss of its transcriptional activating domain but retain the DNA binding domain. The RUNX1 fusion partner genes are known to bind DNA via the RUNX1 DNA-binding domain, whereas their fusion partners ETO or EVI1 act as strong transcriptional repressors[15]. These fusion oncoproteins bind to RUNX1 recognition motifs embedded within DNA sequences of gene promoters and regulatory elements and exert their function by deregulating target genes to perturb the myeloid differentiation program. In both AE and AME driven AMLs, the *CEBPA* expression levels are relatively lower when compared to other AMLs[15], suggesting that *CEBPA* is one of the main differentiation factors to be deregulated in these two subtypes of AML.

Throughout the last decades, it has been studied thoroughly how mutations activate oncoproteins in tumorigenesis. Many of these mutations have been modelled in cell lines and *in vivo* models to understand and recapitulate the phenotype observed in cancer patient samples. However, how these oncoproteins modulate the chromatin landscape to interfere with the expression of their genes is still poorly understood. In this study, we sought to understand the mechanism of how AE negatively[16] influence *CEBPA* expression in AML, by studying the chromatin state and topology of the *CEBPA* locus using cell line models and AE patient samples.

MATERIALS AND METHODS

Patient Samples and gene expression profiling

AML patient samples were collected from three AML cohorts (n=528) that were part of the HOVON study (Netherlands, Austria and Germany) and were processed as previously described. Gene expression analysis for these patients using Affymetrix Human Genome 133 Plus2.0 GeneChips (Affymetrix, Santa Clara, CA, USA) has been published previously. Labelling, hybridization, scanning and data normalization were performed as previously described[17, 18].

Cell lines, culturing and transductions

The cell lines used in this study were grown in the following conditions: Tet-off inducible U937 cell lines (Tet off inducible empty vector and AML1-ETO) in RPMI 1640 supplemented with TET free serum and 1ug/mL of tetracycline (Clontech); Kasumi-1 (parental cell line) in RPMI1640 supplemented with 20% fecal calf serum; Tet on inducible shRNA system Kasumi-1 cell lines (scrambled shRNA and AE shRNA) in RPMI 1640 supplemented with Tet free serum (Clontech), induced with 1ug/mL of tetracycline. Tet-off Inducible cell lines were washed three times with PBS prior to starting with inducible experiments. U937 Tet-off cell lines and Kasumi-1 Tet-on cell lines were a kind gift from D.Zhang and C.Bonifer, respectively. The HL-60 cell line was grown in RPMI 1640 supplemented with 10% FCS. Viral particles were produced by chemical transfection (Fugene 6, Promega) in human embryonic kidney 293T cells with equal concentrations of gag-pol-env construct and different isoforms of AML1-ETO (AE long form, AE short form 9a and AE DNA binding mutant R135G) cloned in pMSCV-IRES-GFP construct. Supernatants were harvested after 48hours and were immediately used for transduction experiments. Retro-viral transduction of these constructs in HL-60 was carried out using Retronectin (Clontech). GFP positive cells were sorted on FACS ARIA 72 hours after transduction and cells were harvested for further investigations. The AML1-ETO (AE long form, AE short form 9a and AE DNA binding mutant R135G) constructs were a kind gift from D.Zhang.

mRNA and protein expression

RNA was harvested from cell lines ($2x10^6$) in Trizol at different time points as previously described and converted to cDNA using reverse transcriptase kit (Invitrogen). QPCRs for CEBPA and AML1-ETO gene expression were quantified using SYBR green FAST and were further run on ABI7500. Nuclear lysates were harvested from the Tet-off inducible U937 lines at different time points; 0 hours, 24hours, 48 hours, 72 hours. Nuclear lysates were prepared using the following buffers: Protein concentration was measured by the pierce BCA Protein assay kit (Thermofisher) and then run on an illuminometer ( ). For every sample, 40ug of protein lysates were first boiled for denaturation at 95^0C then loaded on a 4-10% gel. Gels were blotted on () for an hour, and the blots were further blocked for one hour with milk. Primary antibody for CEBPA, AML1-ETO and GAPDH were incubated at 4^0C overnight, then washed the next morning with PBS/Tween for three times. Secondary antibody (donkey or goat anti-human) was then incubated for 1 hour at room temperature.

Chromatin immunoprecipitation (ChIP)

ChIP experiments were performed as previously described. Cells were crosslinked at room temperature for 10 minutes with 1% formaldehyde, sonicated for chromatin shearing and nuclear lysates were prepared. Immunoprecipitation of crosslinked chromatin was performed overnight at 4^0C with antibodies directed against the histone mark H3K27ac, H3K9ac, AML1-ETO. An input sample for normalization was included for all cell lines and patient sample investigated.

ChIP-seq analysis

The immunoprecipitated DNA was processed according to the Illumina TruSeq ChIP Sample Preparation Protocol (Illumina) and single-end sequenced (1 × 50 bp) on the HiSeq 2500 platform. The quality of the sequence reads was first assessed with *FASTQC* [1] and summarized with *MultiqQC* [2]; all quality checks were considered acceptable. The reads were then uniquely aligned to the human genome build 19 (hg19) using bowtie v1.1.1 with the following settings: *--mm --tryhard -m 1 --best --chunkmbs 1024 --strata*. Peak calling was performed using MACS2 v.2.1.1 [4] with default parameters, taking the immunoprecipitated (IP) sample as the foreground data and its matched genomic DNA (input) as the control background. Further quality checks were conducted with *PicardTools* and a custom-made script that implements metrics developed by the ENCODE Consortium.

Next, the peaks of individual samples were combined into a consensus set of non-overlapping peaks, merging those that overlap by at least one base. Peaks that could not be identified in a minimum of two primary AML samples or overlapped with ENCODE blacklist regions [7] were excluded. Subsequently, the number of reads contained in the peaks of the consensus set was counted individually for every sample and corrected for background signal by subtracting the reads in the matching input sample. The background-subtracted counts were then normalized by trimmed mean of M values (TMM) [8]. The merging of

peaks, counting and normalization were conducted using the *DiffBind* package [9] in R. DiffBind was also employed to conduct unsupervised clustering on the TMM-normalized counts and supervised comparisons between the AE and the AML^{CEBPA+} groups. Namely, fold-change and FDR values were obtained for such comparisons by feeding raw background-subtracted counts into *DESeq;* peaks with a FDR < 0.05 were considered to be significant.

Finally, bigwig tracks were generated for visualization with *deepTools* by counting the number of background-subtracted reads in bins of 20 bp with a 200-bp smoothing window, and normalized by reads per kilobase per million (RPKM).

High resolution circularized chromatin conformation capture sequencing (4C-seq)

High resolution 4C-seq was conducted as previously described. In brief, 10x10^6 cells were crosslinked with 2% formaldehyde for 10 minutes at room temperature. Glycine (0.125M) was added to quench the crosslinking reaction and cells were centrifuged and suspended in lysis buffer to disrupt membranes and isolate chromatin. A primary four-base cutter, either DPNII, was used for digestion, followed by diluted ligations. After precipitation, chromatin was further subjected to a second round of digestions with a different 4 base cutter NLAIII and ligated to small-circularized plasmids. Primers for *CEBPA* viewpoint (Forward: ACTGCTTCTTTACTGCGATC; Reverse: AAATCAAAAAGCACCAAGAG) and for the +42Kkb contact domain viewpoint (Forward: GCCCAGGAGCCTGTGAGATC; Reverse: ACTCTGAGTGCAGAGAGGAG) were designed as previously reported(7). Inverse PCR was carried out to amplify sample libraries that were pooled and spiked with 40% PhiX viral genome sequencing library to increase sample diversity. Multiplexed sequencing was performed on the HiSeq2500 platform. 4C-seq data analysis is explained in the Supplementary Methods.

4C-seq analysis

4C-seq data analysis was performed as previously described. In brief, the 4C-seq read primer sequences with their barcodes were used to demultiplex the reads and to trim the reads from the 5'-end to the first restriction enzyme recognition site. The sequences were mapped, while ignoring quality scores of the read bases and not allowing for a mismatch, to a database of digested genome fragment-ends using the human reference genome build hg19. All 4C-seq samples passed the quality control threshold values.

We normalized the data taking library size and 4C-seq fragment-end types into account. The 4C-seq contact frequency profiles were generated with the median value of the 4C-seq data. The 4C-seq data was smoothed by applying a running trimmed (10%) mean approach using 21 fragment-ends in a single window.

We further determined regions of differentially contact frequencies. 4C-seq data are biased for each fragment-end differently and therefore do not follow a specific distribution. Hence, we used a non-parametric approach to test for statistical significance between two phenotypes. First, we ranked all normalized data for each fragment-end independently, and set ties to the minimal value. Subsequently, we binned, along the locus, the ranks of 21

fragment-ends and calculated the m x n rank frequency matrix (m = 2; i.e. the number of different phenotypes; n = number of samples). We merged the columns of the frequency matrix based on the order of each phenotype and the number of samples, e.g. if the first ranked phenotype, based on the enumeration of the rank frequencies multiplied by the rank number, has been sampled three times, the first three columns of the m x n matrix are merged by adding up the row values, and the other columns are merged similarly resulting in a 2 x 2 matrix. Moreover, a X^2-test was applied on this 2 x 2 contact frequency matrix. The p-values were corrected for multiple hypothesis testing using the Benjamini-Hochberg method. The R statistical package version 3.2.2 was used for the statistical calculations and for generating the 4C-seq contact plots.

RESULTS

AML1-ETO attenuates *CEBPA* expression in myeloid cells

From unsupervised clustering of gene expression profiling data conducted in 528 AML patients, AML1-ETO patients form an independent gene expression cluster from the rest of the AML patients (Fig.1a). All AE patients harbor relatively low levels of *CEBPA* transcripts (Fig.1a and 1b). To investigate the causal relationship between AE fusion oncoprotein and low *CEBPA* expression levels, we retrovirally introduced an IRES-GFP construct containing either the full-length form of AE^{long}, the DNA binding mutant AE^{DBmut} form of AML1-ETO[19] or an empty vector (EV) control into HL-60 cells. Sorted GFP+ HL60 cells transduced with AE^{long}, but not cells with AE^{DBmut} or empty control vector, showed significantly lower *CEBPA* mRNA levels (Figure 1c). Given that extended time points of AE expression exhibit increase in cell death in cell cultures[20], it limited the number of cells to study the dynamics of *CEBPA* deregulation caused by AE^{long}. Using a TET-off inducible AE model in a U937 cell line [20], AE expression was majorly increased after 96 hours (Fig.S1), which was accompanied by a significant decrease in *CEBPA* expression at the mRNA (Fig.1d) and protein level (Fig.S1). To exclude that these AE-driven dynamic effects on *CEBPA* expression are accompanied by an additional genetic drift found in the cell line models used so far, we studied the effects of AE in purified human CD34+ umbilical cord blood (UCB) cells. AE-expressing CD34+ UCB cells showed similar low expression levels as the AE^{long} positive cell lines Kasumi-1 and SKNO-1 when compared with CD34+ MLL-AF9 UCB cells (Fig.1e). These data are in line with a previous study by Link et al, demonstrating loss of *CEBPA* in AE^{long} CD34+ cells, whereas mutant AE constructs showed no effects [21]**.** Inducible shRNA knockdown of AE in Kasumi-1 cell line showed an increase in *CEBPA* expression over a timeframe of 72 hours (Figure 1fi and ii), similar to what was shown previously using a siRNA system targeting the AE fusion gene in Kasumi-1 cell line [22-24]. Together these data demonstrate an inverse causal relation between the presence of the fusion AE oncoprotein and expression of *CEBPA* in patients.

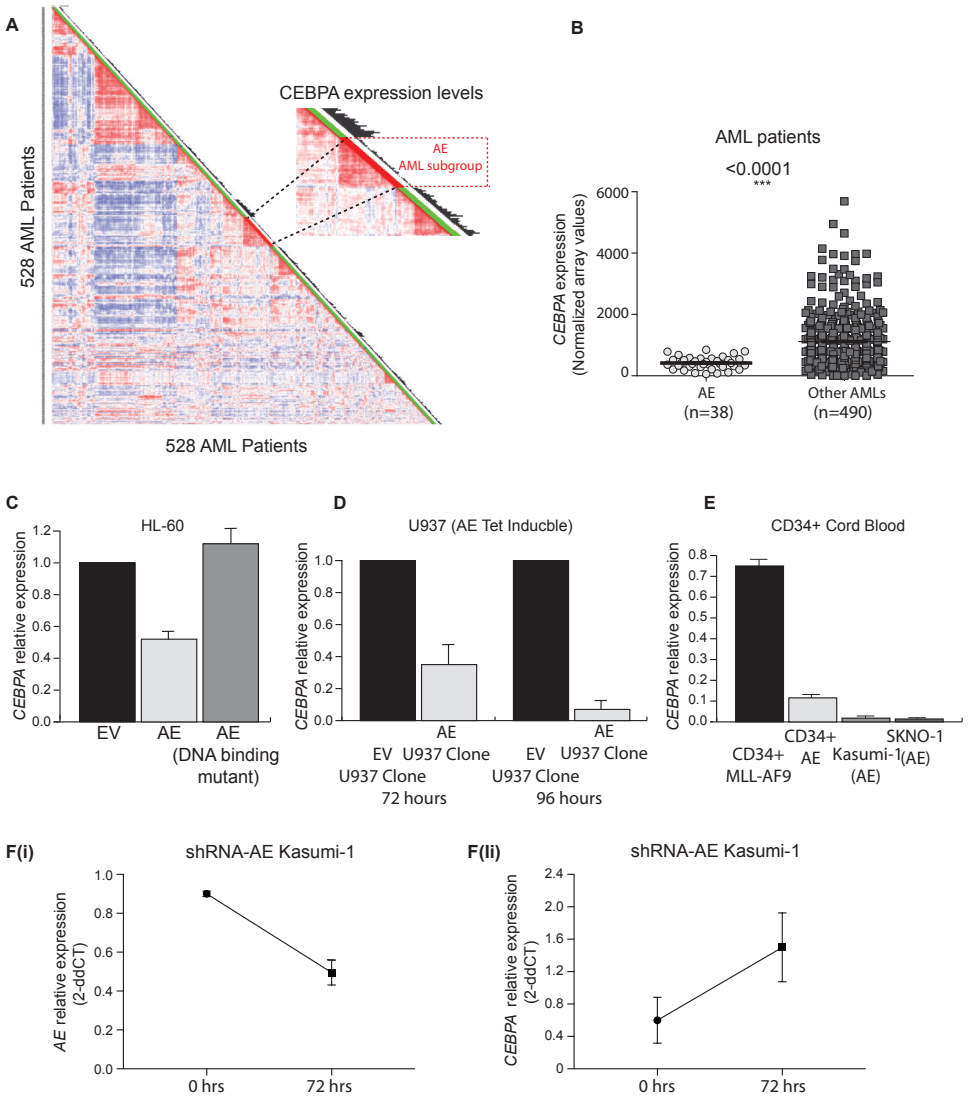


Figure 1. Low *CEBPA* expression levels in the presence of AML1-ETO. (1a) Heat map showing clustering of a 528 AML patient cohort based on gene expression profiling. AML1-ETO patients cluster together in a distinct group from the other AMLs. *CEBPA* expression levels are shown in black across the whole cohort. (1b) Average difference of *CEBPA* expression levels based on normalized array values, comparing AE AML patients with the other AML samples of the cohort. (1c) Relative *CEBPA* mRNA expression levels (2-ddCt) in HL-60 cell line transduced with an empty vector, AE and DNA binding mutated AE constructs. (1d) Relative *CEBPA* mRNA expression levels (2-ddCt) in Tet-off AE inducible U937 cell line after 72 and 96 hours without tetracycline. (1e) QPCR measuring *AE* and *CEBPA* expression on RNA harvested from CD34+ cells transduced with AE or MLL-AF9 [21] and grown in xenograft models. Expression was calculated using the 2-ddct method. (1fi and 1fii) Relative mRNA expression levels (2-ddCt) of *AE* and *CEBPA* in doxoxcyline induced AE shRNA Kasumi-1 relative to a shRNA scrambled control Kasumi-1.

AML1-ETO highjacks the +42kb enhancer and reverses the active chromatin state of the *CEBPA* locus

We hypothesized that AE binding to the *CEBPA* locus downregulates its expression. The genome-wide AE binding profile was determined by ChIP-Seq in the t(8;21) cell line model Kasumi-1 and in three AML patient bone marrow samples carrying the AE fusion gene as the result of a t(8;21) translocation (Fig.2a). AE binds predominantly to the +42kb *CEBPA*-enhancer [25], whereas enrichment was also observed at the +54kb potential enhancer and at the *CEBPG* promoter. The +42kb enhancer contains four conserved RUNX1 binding motifs within its conserved DNA sequence (Fig.S2a), which might all be occupied by AE, explaining the strong binding to this enhancer.

The chromatin state of the *CEBPA* locus in myeloid cells is highly enriched with widespread acetylation of H3K27 and H3K9 across the whole *CEBPA* TAD [25]. We next sought to interrogate the *CEBPA* locus for active chromatin state changes that correlate with low *CEBPA* expression levels and the presence of AE. Genome-wide chromatin profiling in AML patients dissected AE patients (n=5) from other AML patients (n=5) in two distinctive groups, based on H3K9ac and H3K27ac profiling (Fig.S3 and S4). H3K9ac profiling showed significant enrichment-changes, predominantly occurring at +53kb, +42kb, +9kb, +3kb, -1.8kb and -29kb genomic regions (Fig.2b-2h). In addition, profiling for H3K27ac shows that most of the enhancers located downstream (+54kb, +42kb, +35kb, +29kb, +9kb, +3kb, *CEBPA* gene) and upstream (-29kb) of *CEBPA* have significant low H3K27ac enriched sites in AE patients expressed as trimmed mean of M-values (TMM), when compared to the other AMLs (Fig.3a-3j). The most significant enrichment-change is observed at the +9kb enhancer (Fig.2b, 2e, 3 and 3h). This enhancer is active in all *CEBPA* expressing tissues [25], suggesting that all the other 3' enhancers in the locus collaborate with *CEBPA* through interactions with the +9kb enhancer, independent of the tissue or organ.

Based on all the findings obtained in AML patients, we hypothesize that AE binding to the *CEBPA* locus reverses the active chromatin, hence the de-regulation of *CEBPA* mRNA expression. Using the CD34+ cord blood cell system, the *CEBPA* locus was devoid of any of the H3K9ac or H3K27ac in the presence of AE (Figure 1e). In contrast, MLL-AF9 CD34+ cells the *CEBPA* locus was highly H3K9 and H3K27 acetylated (Fig.S5a and S5b), in line with *CEBPA* expression in these cell models (Fig1e). We then used an inducible AE shRNA knockdown system in Kasumi1 model and showed that upon shRNA activation, H3K27ac chromatin was widely enriched throughout the *CEBPA* locus (Fig.S6), recapitulating previous reported findings H3K9ac results derived from AE knockdown using a siRNA system [22-24]. We conclude that AE binds the +42kb enhancer to reverse the active chromatin state of the whole locus, which explains the increased *CEBPA* mRNA levels upon knockdown of AE in Kasumi1 (Fig.1e).

+42kb enhancer looping with *CEBPA* through the +9kb enhancer is lost in AE AML blasts.
CEBPA engages into a chromatin loop with its 3' enhancers in myeloid cells. We hypothesized that enhancer de-acetylation in AE cells coincides disturb chromatin looping. To determine whether the enhancer loses engagement with *CEBPA* in the presence of AE, we applied high resolution 4C-seq in a cohort of AML samples with a translocation t(8;21) and low *CEBPA* expression (n=3), compared with AMLs without AE fusion gene and with mid-to high *CEBPA* expression levels (n=4). Patient characteristics are shown in Table 1. We used the +42kb enhancer as the anchor point since it is the main region at which AE binds in the *CEBPA* locus. A semi-quantitative analysis of 4C-seq revealed a strong interaction of the +42Kb region with the +9kb enhancer, and to a lesser extent with the *CEBPA* gene and promoter. The +9kb enhancer forms a chromatin complex with +42kb enhancer possibly to regulate *CEBPA* gene expression, which is in line with the strongest H3K27ac or H3K9ac enrichment observed at the +9kb enhancer in the absence of AE (Figure 2b and 3a). In fact, AE AML patients show a significant decrease in the interaction between the +42kb enhancer and the +9kb enhancer, which complements with the loss of active chromatin state and *CEBPA* expression in these patients. The chromatin interactions with the *CEBPG* locus (Fig.4a and b) appeared stable and were not different between the AE AML patients and control AMLs. (Fig.2b and Fig.3a) Altogether, these findings compliment the line of thought that AE disturbs *CEBPA* expression interfering with looping interactions with its enhancers, possibly via a de-acetylation mechanism.

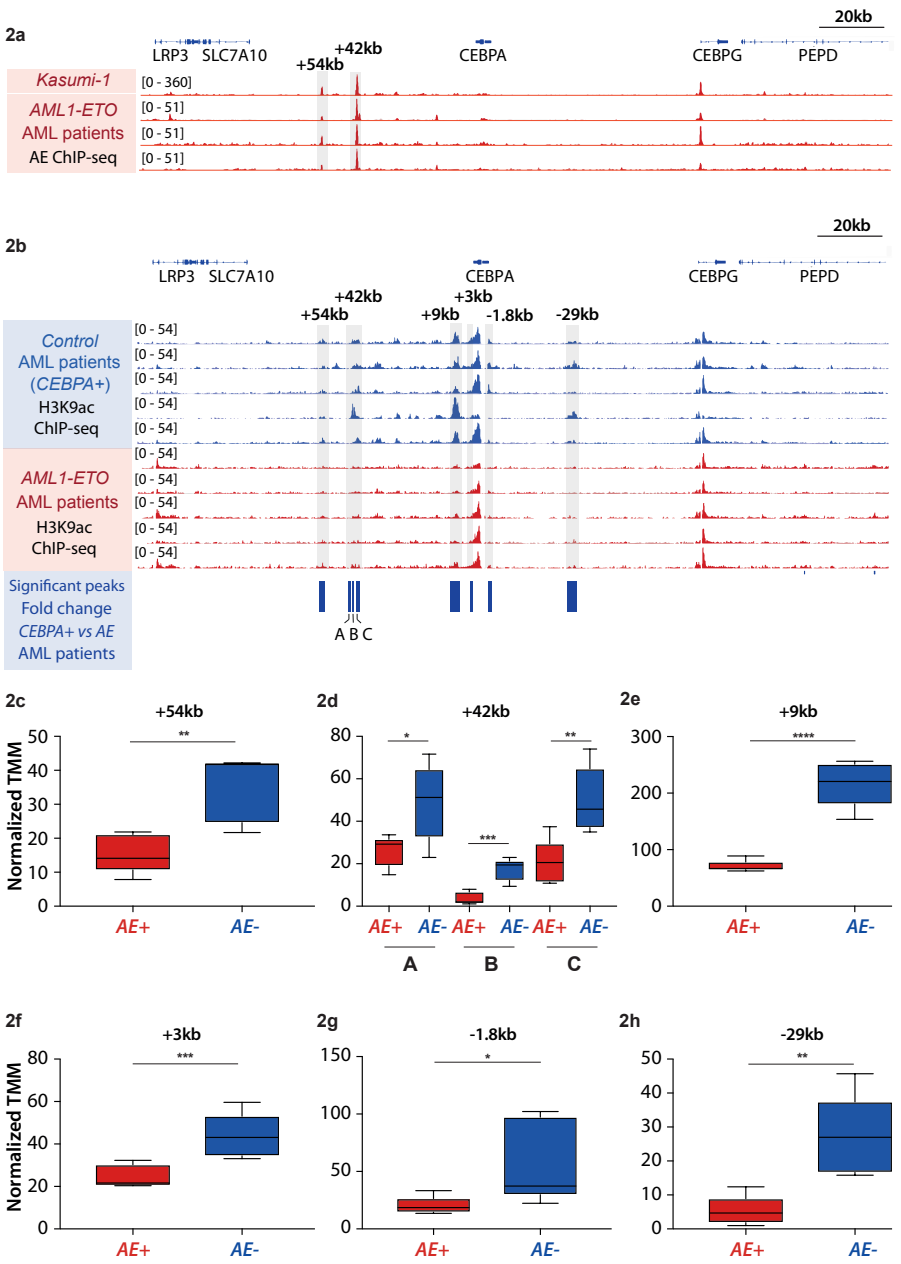


Figure 2. *AML1-ETO binds the +42kb enhancer and reverses the active chromatin state*. (a) ChIP-seq of AE binding at the *CEBPA* locus, in Kasumi-1 cell line and AE patients (n=3). (b) H3K9ac Chip-seq in AML patients (n=5) without AE fusion (light blue) and AE patients (n=5) (light red). Significant peaks (blue) calculated via a fold enrichment ratio with an FDR threshold of 0.05, comparing AMLs without AE and with AE. The significant peaks occur at +54kb, +42kb (A,B,C), +9kb, +3kb, -1.8kb, and -29kb. (c-h) Differential enrichment in H3K9ac at the distinct defined putative enhancers, expressed as trimmed mean of M-values (TMM values) between AE (red) and the control AMLs using student t-Test ($p<0.05$).

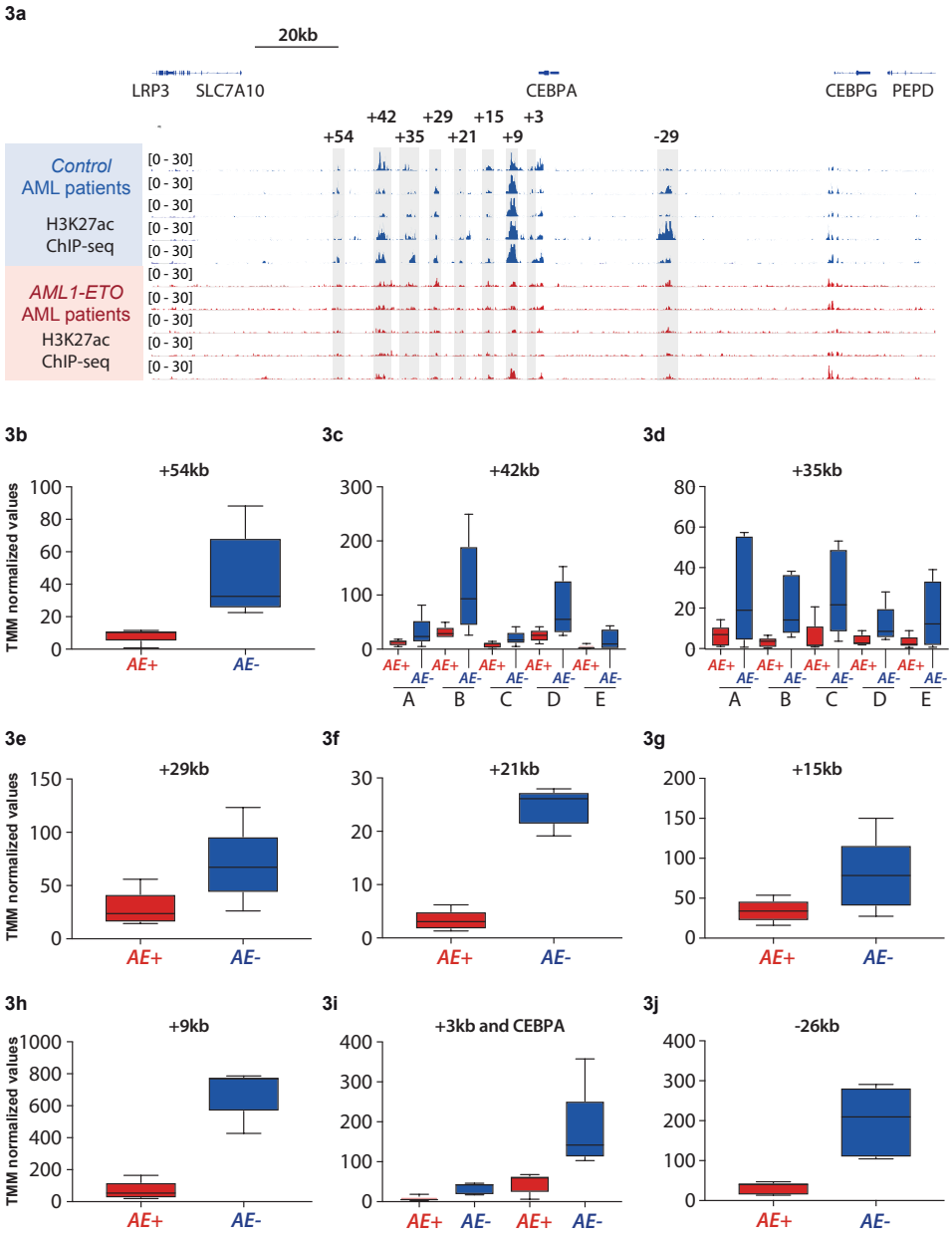
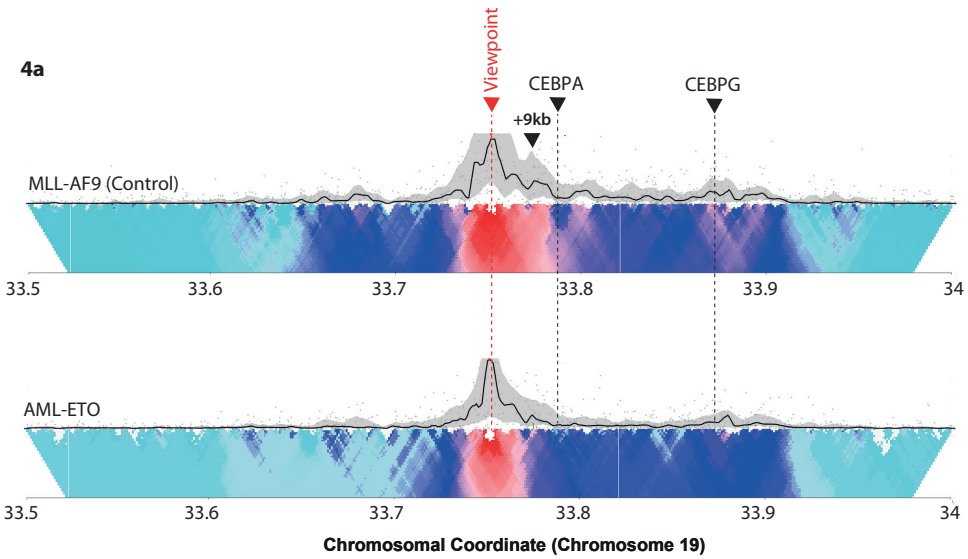


Figure 3. *AE deactivates H3K27ac at active enhancers within the CEBPA locus*. (a) ChIP-seq of H3K27ac in control AML patients (light blue) (n=5) compared with AE AML patients (light red) (n=5); (b-j) Differential H3K27ac enrichment at distinct enhancers (see Legend Figure 2c-h). AE (red) or AE^{+} and other AML^{CEBPA+} or AE^{-} using student t-Test ($p<0.05$).



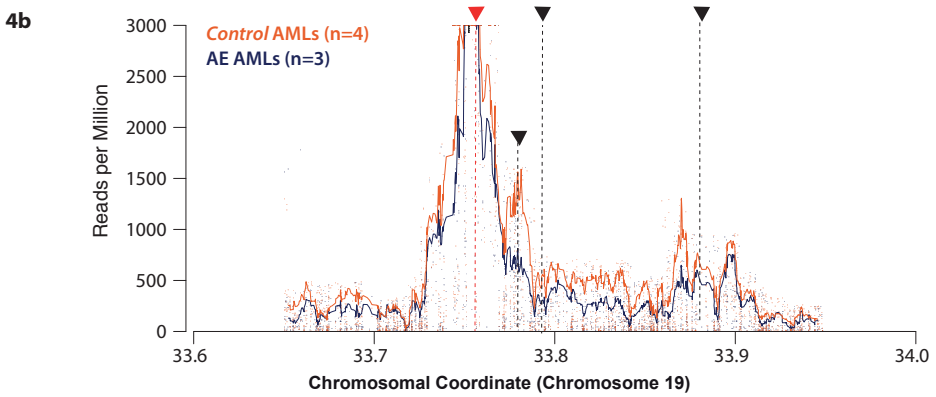


***Figure 4**. Loss of enhancer to CEBPA interaction in AE AML patients as determined by 4C-Sequencing.* (a) Contact profile of the active *CEBPA* locus in an MLL-AF9 AML patient as opposed to an inactive *CEBPA* locus in an AE AML patient. Viewpoint is at the +42kb enhancer (Viewpoint) showing interaction with +9kb enhancer in the MLL-AF9 patient cells, whereas reduced interaction between the +42kb (Viewpoint) and the +9kb enhancer was observed in the AE patient. (b) A semi-quantitative approach to investigate for chromatin interactions in the *CEBPA* locus by comparing 4 control AML patients against 3 AE AML patients, using the +42kb as a viewpoint (Red Triangle). Significant interaction is observed between the +42kb and the +9kb enhancer in control AMLs (Red line)(Green triangle. This interaction is not seen in AE AMLs. The control AMLs represent from top to bottom a case with an inversion 16, DNMT3A mutant, NPM1 mutant, MLL-AF9 leukemia and another NPM1 mutant case.

DISCUSSION

Cell lineage commitment and differentiation in the bone marrow are driven by the interplay of TF networks that establish chromatin states of differentiation related genes. In neutrophilic differentiation, TFs physically bind to critical gene loci, modulate their chromatin state and regulate expression of target genes. The *CEBPA* locus is highly occupied by a TF network in CD34+ HSPCs, which binds predominantly to its autonomous +42kb enhancer. Some of the major players in this network are commonly deregulated by oncogenic mechanisms in AML, including FLI1[25], ERG[26, 27], GATA2[28, 29], RUNX1[13, 15, 30], LYL1[31], and LMO2[32]. In myeloid differentiation, RUNX1 binds to the +42kb enhancer to form a feed-forward loop to to drive differentiation of committed myeloid progenitors[6]. In AML, RUNX1 exhibits point mutations in its coding sequence[12, 14, 33] and is recurrently involved in chromosomal translocations such as t(8;21) and t(3;21), leading to the generation of RUNX1 fusion genes and haploinsufficiency of the wild-type RUNX1 protein[15].

The reciprocal translocation t(8;21) generates a fusion oncoprotein RUNX1-RUNX1T1 (RUNX1-ETO or AML1-ETO (AE), which hijacks the myeloid differentiation program by attenuating the expression of early myeloid genes, including *CEBPA*[22-24]. In our cohort of patients, we observed that the AE subgroup exhibits low *CEBPA* expression levels compared to other AML patients (Fig.1). Moreover, unsupervised clustering of gene expression analysis of AML patients shows that all AE patients are clustered within the same group, indicating that AE is the common abnormality that drives the disease in the AE AML subtype. A study by Dan Tenen and colleagues reported similar observations, and they proposed a mechanism stating that AE interferes with the binding of the C/EBPα protein to its own promoter by disrupting its auto-regulatory loop required for sustaining *CEBPA* expression levels[24]. However, the absence of genome wide technologies at the time the study was conducted, limited the authors to reach optimal conclusions. Using ChIP-seq to identify the genome wide profile of AE binding to DNA, we show that the fusion oncoprotein binds predominantly at the +42kb enhancer within the *CEBPA* locus in AE patients and in cell line models, rather than the *CEBPA* promoter.

We hypothesized that AE binding to the enhancer caused changes in the chromatin state and conformation of the *CEBPA* locus leading to the observed decreased expression of the gene. Changes in the chromatin state takes into consideration several aspects of chromatin structure and function, which are involved in transcriptional control. These include, the post-translational modifications at N-terminal sites of histone tails [34, 35], the recruitment of PolII at enhancer regions[36], the production of bidirectional enhancer RNAs [37-39] and the participation of chromatin factors that write, maintain and erase the histone code[40]. Acetylation of histones on lysines such as H3K27ac and H3K9ac are associated with active transcription[34]. These histone marks are catalyzed by two different histone acetyl-transferases (HATs) protein complexes. The GCN5/PCAF complex acetylates

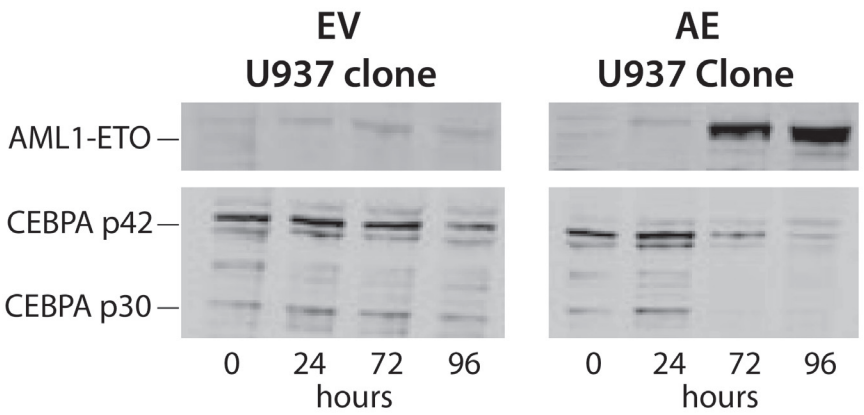
H3K9 whereas CBP/p300 catalyzes the H3K27[41, 42]. Both the GCN5/PCAF complex and the CBP/p300 complexes are chromatin writers of acetyl residues to lysines. By conducting genome-wide histone profiling of both marks in AE patients (low *CEBPA* expression) and a control cohort with higher *CEBPA* expression levels, we found that H3K9ac as well as H3K27ac levels were relatively low at the +42kb enhancer in AE expressing cells. This can be explained by differences in protein complex recruitment by AE to specific gene loci. In fact, upon knockdown of AE, H3K9ac [22-24] and H3K27ac are severely increased in a genome-wide fashion in Kasumi-1 cells. This is a striking observation because it indicates that there is a selective recruitment capacity of specific protein complexes by AE that alters specific histone modifications to deregulate genes. Our results suggest that AE might compete with specific HATs for acetylated regions occurring at H3K9 as well as on H3K27, genome-wide. We hypothesize that this can occur by the oncogenic recruitment of histone deacetylase (HDACs) complexes via AE that will overcome the normal physiological function of PCAF/GCN5 or CBP/p300. The family of NAD+ dependent histone deacetylases known as sirtuins, particularly SIRT1 and SIRT6, are responsible for deactylation of H3K9 residues and are also expressed in myeloid cells [43-45]. Whether SIRT1 and SIRT6 are found in the AE protein complex and recruited to the +42kb enhancer is yet to be determined. Further functional experiments are recommended to study the interplay between PCAF/GCN5, CBP/p300, HDACs (in this case SIRT1, SIRT6) and AE to investigate how they act in the same protein complex to hijack the *CEBPA*-enhancer and its expression levels in AML.

An active chromatin state correlates with the engagement of genes to corresponding enhancers in a chromatin loop. We previously showed that *CEBPA* is located in a topological associated domain (TAD) and interacts with its +42kb enhancer mainly in *CEBPA* expressing myeloid cell lines. The *CEBPA* TAD contains two genes, *CEBPA* and *CEBPG*. We and other groups have shown that the +42kb enhancer is critical for *CEBPA* regulation but not for *CEBPG* [7, 46, 47]. We show by 4C-seq that in the presence of AE the +42kb enhancer dissociates from the *CEBPA* gene in AML patients, potentially causing its down-regulation. We observed that the the +42kb enhancer forms a loop with the +9kb enhancer of *CEBPA* and that this interaction is absent in AE patient cells. (Fig.4b). This suggests that the +9kb enhancer plays an important role in *CEBPA* gene activation. This observation is supported by findings in this study (Fig.2b and Fig.3a), demonstrating that the +9kb enhancer shows the highest and most significant loss of H3K9ac and H3K27ac enrichment in the presence of AE, and from a previous study in which we showed that all hematopoietic and non-hematopoietic cells expressing *CEBPA* have an active +9kb enhancer[7].

Does AE protein on its own cause loss of interaction between the +42kb and the +9kb enhancer and is this the main cause of histone deacetylation and subsequent downregulation of *CEBPA* mRNA levels? Tethering experiments to investigate the gene deregulation effect

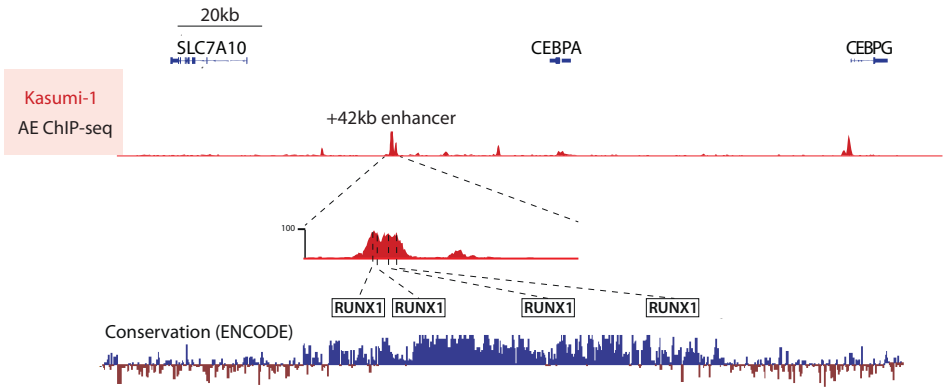
of AE can be applied using genome editing systems to target specifically the enhancer by transcriptional repressing protein complexes. For instance, fusing dCas9 with ETO or any of its repressive protein complexes, such as HDACs or SIRT de-acetyl transferases, will allow us to study in a time-wise fashion how the loss of acetylation marks at the *CEBPA* locus deregulate gene expression and chromatin topology. In a recent study, tethering HDACs to a specific locus showed negative effects on H3K27ac levels and on the expression level of the gene of interest[48]. Whether the chromatin changes occur sequentially or as independent events has still to be determined [49]. In conclusion, the results described in this study report how the AE oncoprotein inactivates gene expression by reversing the active chromatin state of the *CEBPA* TAD in myeloid progenitors and disturbs chromatin interactions that are thought to support transcriptional activation. This mechanism of how oncoproteins deregulate genes can be used as a paradigm for other oncoproteins in AML and might serve as a platform for therapeutic implications.

SUPPLEMENTARY FIGURES

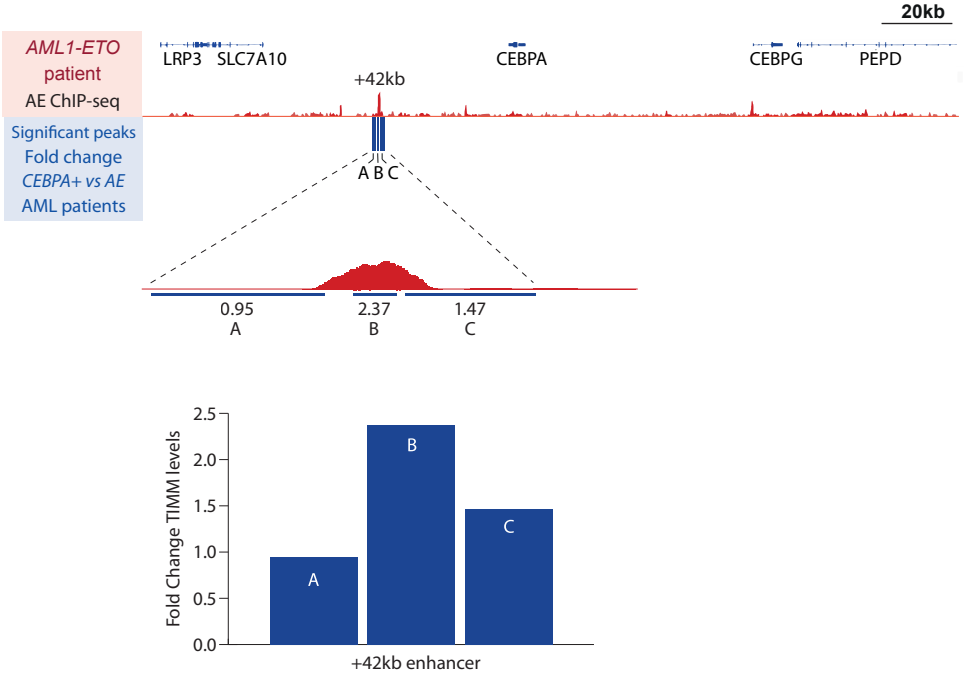


Supp. Figure1. Western blot using antibodies directed to ETO or CEBPα showing increased ectopic expression of AE in U937 cell line and reduction of C/EBPα expression across different time points, post tetracycline.

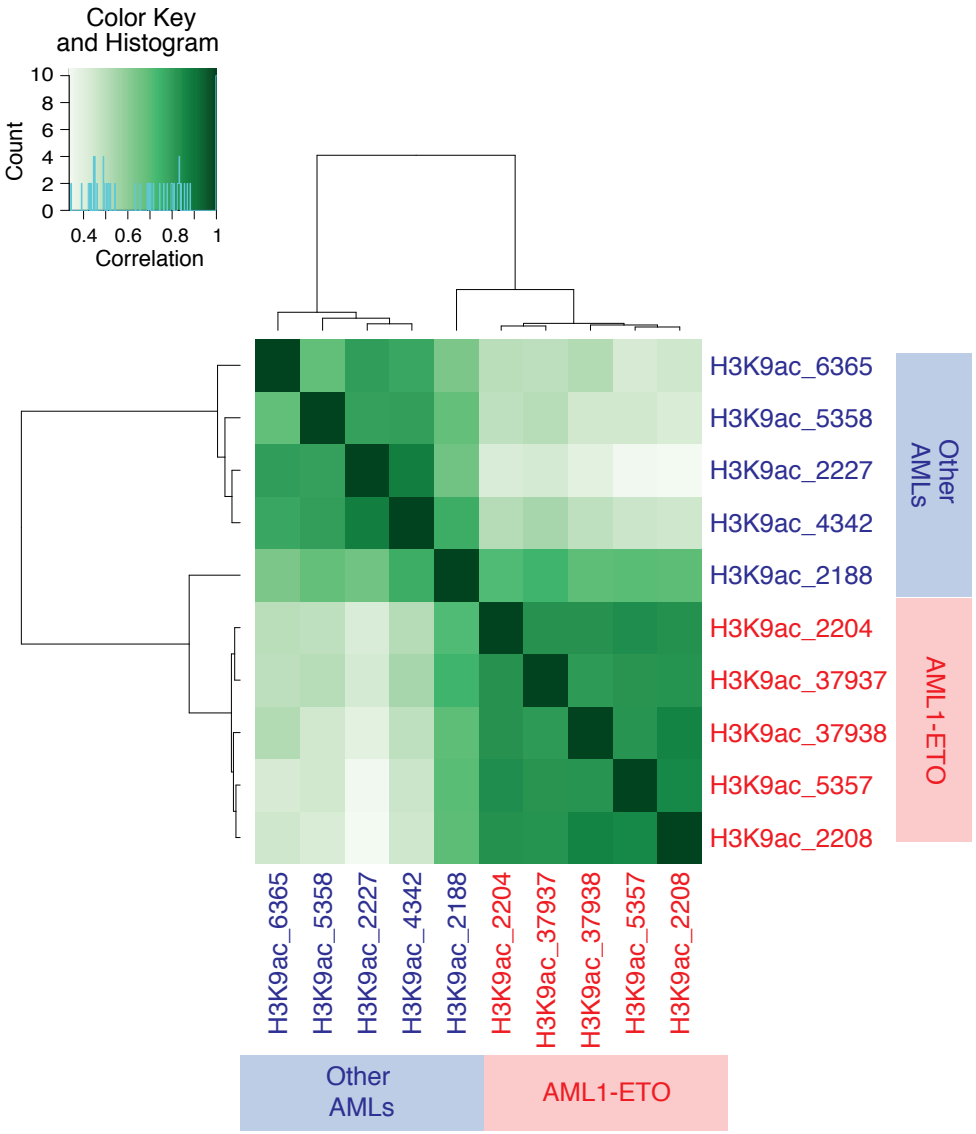
Supplemetary Fig.2a



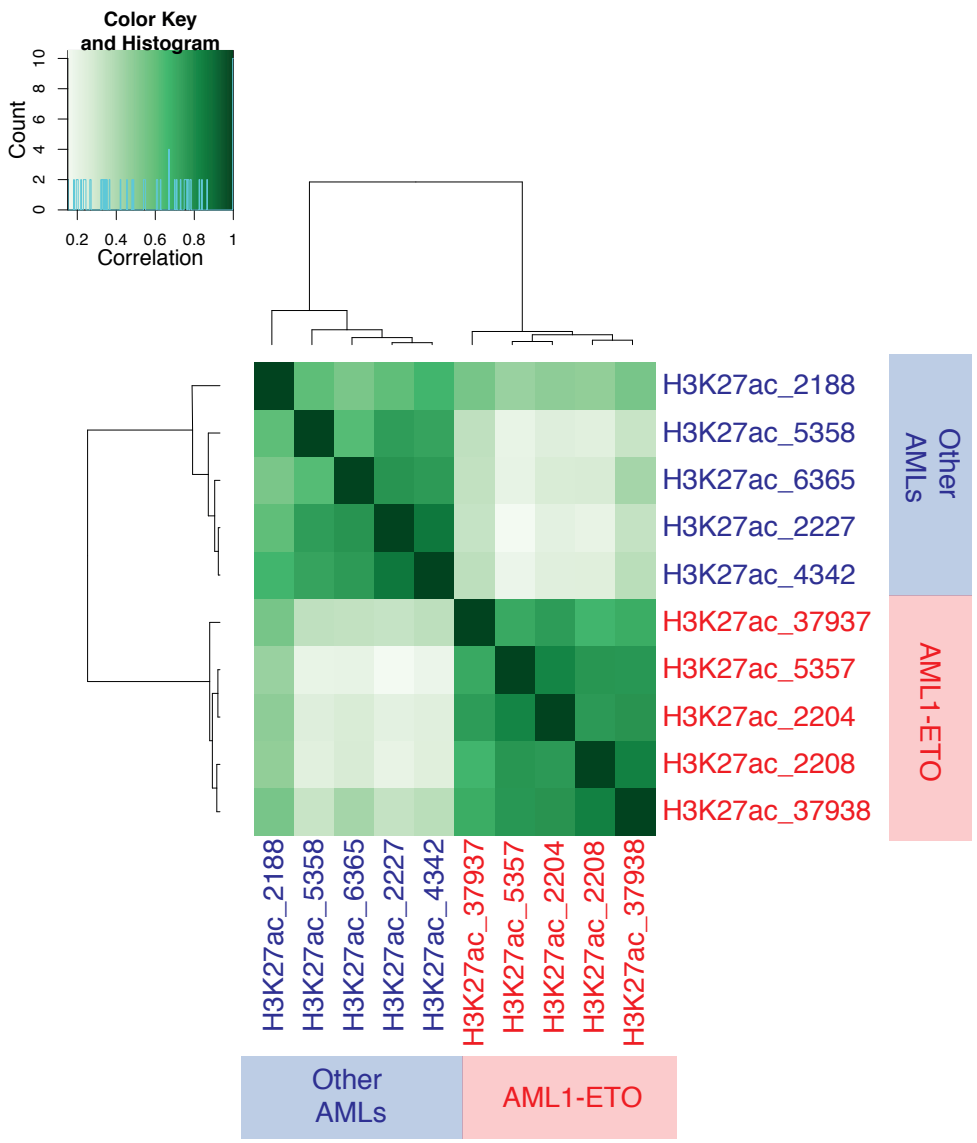
Supplemetary Fig.2b



Supp. Figure 2. (a) Chip-seq of AE in Kasumi-1 showing binding at the +42kb enhancer and RUNX1 conserved motifs (b) Chip-seq for AE using ETO specific antibody at the +42kb enhancer showing three significant (FDR<0.05) differentially enriched H3K9ac regions, A,B and C in AE and non-AE patients. The B region is where AE binds to DNA, being the most significantly differentially enriched H3K9ac region within the +42kb enhancer.

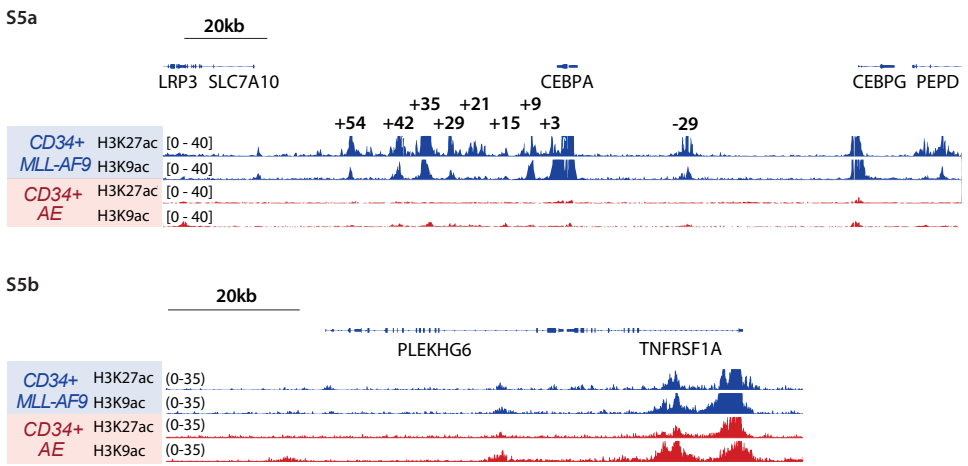


Supp.Figure 3. Pearson correlation of 10 AML samples, i.e. 5AE AML samples and 5 control AMLs. Clustering was based on genome-wide H3K9ac patterns.

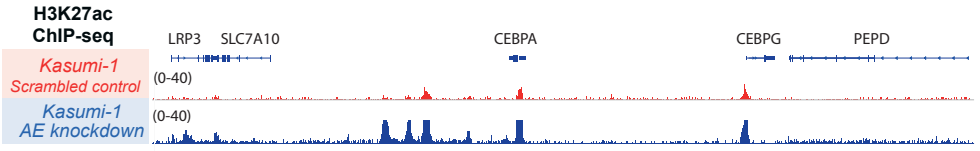


Supp. Figure 4. Pearson correlation of 10 AML samples based H3K27ac patterns (See legend Supp. Figure 3).

Supplementary Figure 5



Supp.Figure 5. Chip-seq of H3K9ac and H3K27ac of MLL-AF9 and AE expressing CD34+ cord blood cells. Acetylation patterns of the CEBPA locus (S5a) and a control region (S5b) is shown



Supp.Figure 6. Chip-seq of H3K27ac at the *CEBPA* locus (S6a) and a control in Kasumi cells. Acetylation patterns between AE knock down cells are compared to scrambled control treated cells.

table 1

Patient Number	Material	AML Subtype
37937	BM3	AE
2204	BL2	AE
2208	BL4	AE
5357	BM3	AE
37938	BM3	AE
6365	BL3	inv16
2188	BL3	DNMT3A
2227	BM4	NPM1
5358	BM3	MLL
4342	BL3	NPM1

REFERENCES

1. Ye, M., et al., *Hematopoietic Differentiation Is Required for Initiation of Acute Myeloid Leukemia.* Cell Stem Cell, 2015. **17**(5): p. 611-23.
2. Ohlsson, E., et al., *Initiation of MLL-rearranged AML is dependent on C/EBPalpha.* J Exp Med, 2014. **211**(1): p. 5-13.
3. Frohling, S. and H. Dohner, *Disruption of C/EBPalpha function in acute myeloid leukemia.* N Engl J Med, 2004. **351**(23): p. 2370-2.
4. Nerlov, C., *C/EBPalpha mutations in acute myeloid leukaemias.* Nat Rev Cancer, 2004. **4**(5): p. 394-400.
5. Paz-Priel, I. and A. Friedman, *C/EBPalpha dysregulation in AML and ALL.* Crit Rev Oncog, 2011. **16**(1-2): p. 93-102.
6. Guo, H., et al., *Runx1 deletion or dominant inhibition reduces Cebpa transcription via conserved promoter and distal enhancer sites to favor monopoiesis over granulopoiesis.* Blood, 2012. **119**(19): p. 4408-18.
7. Avellino, R., et al., *An autonomous CEBPA-enhancer specific for myeloid-lineage priming and neutrophilic differentiation.* Blood, 2016. **127**(24): p. 2991-3003.
8. Lancrin, C., et al., *The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage.* Nature, 2009. **457**(7231): p. 892-5.
9. Lie, A.L.M., et al., *RUNX1 positively regulates a cell adhesion and migration program in murine hemogenic endothelium prior to blood emergence.* Blood, 2014. **124**(11): p. e11-20.
10. Sood, R., Y. Kamikubo, and P. Liu, *Role of RUNX1 in hematological malignancies.* Blood, 2017. **129**(15): p. 2070-2082.
11. Roumier, C., et al., *New mechanisms of AML1 gene alteration in hematological malignancies.* Leukemia, 2003. **17**(1): p. 9-16.
12. Migas, A., et al., *AML1/RUNX1 gene point mutations in childhood myeloid malignancies.* Pediatr Blood Cancer, 2011. **57**(4): p. 583-7.
13. Osato, M., *Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia.* Oncogene, 2004. **23**(24): p. 4284-96.
14. Taketani, T., et al., *AML1/RUNX1 mutations are infrequent, but related to AML-M0, acquired trisomy 21, and leukemic transformation in pediatric hematologic malignancies.* Genes Chromosomes Cancer, 2003. **38**(1): p. 1-7.
15. Loke, J., et al., *RUNX1-ETO and RUNX1-EVI1 Differentially Reprogram the Chromatin Landscape in t(8;21) and t(3;21) AML.* Cell Rep, 2017. **19**(8): p. 1654-1668.
16. Sun, X.J., et al., *A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis.* Nature, 2013. **500**(7460): p. 93-7.
17. Valk, P.J., et al., *Prognostically useful gene-expression profiles in acute myeloid leukemia.* N Engl J Med, 2004. **350**(16): p. 1617-28.
18. Verhaak, R.G., et al., *Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling.* Haematologica, 2009. **94**(1): p. 131-4.
19. Matheny, C.J., et al., *Disease mutations in RUNX1 and RUNX2 create nonfunctional, dominant-negative, or hypomorphic alleles.* EMBO J, 2007. **26**(4): p. 1163-75.
20. Burel, S.A., et al., *Dichotomy of AML1-ETO functions: growth arrest versus block of differentiation.* Mol Cell Biol, 2001. **21**(16): p. 5577-90.
21. Link, K.A., et al., *Supraphysiologic levels of the AML1-ETO isoform AE9a are essential for transformation.* Proc Natl Acad Sci U S A, 2016. **113**(32): p. 9075-80.

22. Ptasinska, A., et al., *Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding.* Leukemia, 2012. **26**(8): p. 1829-41.
23. Ptasinska, A., et al., *Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal.* Cell Rep, 2014. **8**(6): p. 1974-88.
24. Pabst, T., et al., *AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia.* Nat Med, 2001. **7**(4): p. 444-51.
25. Kornblau, S.M., et al., *Abnormal expression of FLI1 protein is an adverse prognostic factor in acute myeloid leukemia.* Blood, 2011. **118**(20): p. 5604-12.
26. Panagopoulos, I., et al., *Fusion of the FUS gene with ERG in acute myeloid leukemia with t(16;21)(p11;q22).* Genes Chromosomes Cancer, 1994. **11**(4): p. 256-62.
27. Kong, X.T., et al., *Consistent detection of TLS/FUS-ERG chimeric transcripts in acute myeloid leukemia with t(16;21)(p11;q22) and identification of a novel transcript.* Blood, 1997. **90**(3): p. 1192-9.
28. Niimi, K., et al., *GATA2 zinc finger 2 mutation found in acute myeloid leukemia impairs myeloid differentiation.* Leuk Res Rep, 2013. **2**(1): p. 21-5.
29. Groschel, S., et al., *A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia.* Cell, 2014. **157**(2): p. 369-381.
30. Rodriguez-Perales, S., et al., *Truncated RUNX1 protein generated by a novel t(1;21)(p32;q22) chromosomal translocation impairs the proliferation and differentiation of human hematopoietic progenitors.* Oncogene, 2016. **35**(1): p. 125-34.
31. Meng, Y.S., et al., *Oncogenic potential of the transcription factor LYL1 in acute myeloblastic leukemia.* Leukemia, 2005. **19**(11): p. 1941-7.
32. Patel, J.L., et al., *LMO2 (LIM domain only 2) is expressed in a subset of acute myeloid leukaemia and correlates with normal karyotype.* Histopathology, 2014. **64**(2): p. 226-33.
33. Roumier, C., et al., *M0 AML, clinical and biologic features of the disease, including AML1 gene mutations: a report of 59 cases by the Groupe Francais d'Hematologie Cellulaire (GFHC) and the Groupe Francais de Cytogenetique Hematologique (GFCH).* Blood, 2003. **101**(4): p. 1277-83.
34. Lara-Astiaso, D., et al., *Immunogenetics. Chromatin state dynamics during blood formation.* Science, 2014. **345**(6199): p. 943-9.
35. Winter, D.R. and I. Amit, *The role of chromatin dynamics in immune cell development.* Immunol Rev, 2014. **261**(1): p. 9-22.
36. Stees, J.S., et al., *Recruitment of transcription complexes to enhancers and the role of enhancer transcription.* Biology (Basel), 2012. **1**(3): p. 778-93.
37. Flynn, R.A. and H.Y. Chang, *Active chromatin and noncoding RNAs: an intimate relationship.* Curr Opin Genet Dev, 2012. **22**(2): p. 172-8.
38. Stavreva, D.A. and G.L. Hager, *Chromatin structure and gene regulation: a dynamic view of enhancer function.* Nucleus, 2015. **6**(6): p. 442-8.
39. Eidem, T.M., J.F. Kugel, and J.A. Goodrich, *Noncoding RNAs: Regulators of the Mammalian Transcription Machinery.* J Mol Biol, 2016. **428**(12): p. 2652-9.
40. Brien, G.L., D.G. Valerio, and S.A. Armstrong, *Exploiting the Epigenome to Control Cancer-Promoting Gene-Expression Programs.* Cancer Cell, 2016. **29**(4): p. 464-476.
41. Jin, Q., et al., *Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation.* EMBO J, 2011. **30**(2): p. 249-62.

42. Sun, X.J., et al., *The Role of Histone Acetyltransferases in Normal and Malignant Hematopoiesis.* Front Oncol, 2015. **5**: p. 108.
43. Kawahara, T.L., et al., *SIRT6 links histone H3 lysine 9 deacetylation to NF-kappaB-dependent gene expression and organismal life span.* Cell, 2009. **136**(1): p. 62-74.
44. Michishita, E., et al., *SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin.* Nature, 2008. **452**(7186): p. 492-6.
45. Chen, C.W., et al., *DOT1L inhibits SIRT1-mediated epigenetic silencing to maintain leukemic gene expression in MLL-rearranged leukemia.* Nat Med, 2015. **21**(4): p. 335-43.
46. Guo, H., S. Cooper, and A.D. Friedman, *In Vivo Deletion of the Cebpa +37 kb Enhancer Markedly Reduces Cebpa mRNA in Myeloid Progenitors but Not in Non-Hematopoietic Tissues to Impair Granulopoiesis.* PLoS One, 2016. **11**(3): p. e0150809.
47. Cooper, S., H. Guo, and A.D. Friedman, *The +37 kb CEBPA-enhancer Is Critical for Cebpa Myeloid Gene Expression and Contains Functional Sites that Bind SCL, GATA2, C/EBPalpha, PU.1, and Additional Ets Factors.* PLoS One, 2015. **10**(5): p. e0126385.
48. Kwon, D.Y., et al., *Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC.* Nat Commun, 2017. **8**: p. 15315.
49. Hug, C.B., et al., *Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription.* Cell, 2017. **169**(2): p. 216-228 e19.
40. Jin, Q., et al., *Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation.* EMBO J, 2011. **30**(2): p. 249-62.
41. Sun, X.J., et al., *The Role of Histone Acetyltransferases in Normal and Malignant Hematopoiesis.* Front Oncol, 2015. **5**: p. 108.
42. Kawahara, T.L., et al., *SIRT6 links histone H3 lysine 9 deacetylation to NF-kappaB-dependent gene expression and organismal life span.* Cell, 2009. **136**(1): p. 62-74.
43. Michishita, E., et al., *SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin.* Nature, 2008. **452**(7186): p. 492-6.
44. Chen, C.W., et al., *DOT1L inhibits SIRT1-mediated epigenetic silencing to maintain leukemic gene expression in MLL-rearranged leukemia.* Nat Med, 2015. **21**(4): p. 335-43.
45. Guo, H., S. Cooper, and A.D. Friedman, *In Vivo Deletion of the Cebpa +37 kb Enhancer Markedly Reduces Cebpa mRNA in Myeloid Progenitors but Not in Non-Hematopoietic Tissues to Impair Granulopoiesis.* PLoS One, 2016. **11**(3): p. e0150809.
46. Cooper, S., H. Guo, and A.D. Friedman, *The +37 kb CEBPA-enhancer Is Critical for Cebpa Myeloid Gene Expression and Contains Functional Sites that Bind SCL, GATA2, C/EBPalpha, PU.1, and Additional Ets Factors.* PLoS One, 2015. **10**(5): p. e0126385.
47. Kwon, D.Y., et al., *Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC.* Nat Commun, 2017. **8**: p. 15315.
48. Hug, C.B., et al., *Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription.* Cell, 2017. **169**(2): p. 216-228 e19.

THESIS_Roberto_Avellino.indd 143 07-05-18 10:18

6

THESIS_Roberto_Avellino.indd 144 07-05-18 10:18

NEUTROPENIA CAUSED BY LOSS OF *CEBPA* EXPRESSION REDUCES HEMATOPOIETIC STEM CELL NUMBERS DRIVEN BY A NON-CELL AUTONOMOUS MECHANISM

Roberto Avellino[1], Marije Havermans[1], Remco Hoogenboezem[1], Leonie Smeenk[1], Zheng Ping[1], Nathan Salomonis[2], Rebekka Schneider[1], Eric Bindels[1], Lee Grimes[2], Ruud Delwel[1]

[1]Department of Hematology, Erasmus University Medical Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio
[2]Center, Rotterdam, The Netherlands; Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio

Work in progress

THESIS_Roberto_Avellino.indd 145 07-05-18 10:18

ABSTRACT

A constant daily production of blood cells requires tight regulation of cell differentiation to accomplish steady-state hematopoiesis. Hematopoiesis is mainly governed by transcription factors (TFs) that prime gene expression programs and commit the long-term hematopoietic stem cells (LT-HSCs) into hematopoietic progenitor cells (HPCs) and towards terminally differentiated cells. The neutrophil gene expression program is activated by C/EBPa, a leucine zipper TF indispensable for neutrophil development. Knockout of *Cebpa* or its myeloid specific +37kb *Cebpa*-enhancer in mouse models have two major effects: (1) neutropenia in bone marrow and peripheral blood; (2) decrease in LT-HSC numbers. Whether the latter finding is cell autonomous (intrinsic) to the LT-HSCs or an extrinsic event exerted on the stem cell compartment is not clear yet. We investigated this in a knockout mouse model of the +37kb *CEBPA*-enhancer generated in embryos by CRISPR/Cas9. Transcriptomics using RNA-seq on bulk HPCs and single LT-HSCs show that *Cebpa* is expressed in HPCs but not in LT-HSCs. In line with these results, FACS analysis in the *Cebpa*-enhancer knockout model showed that the reduction in LT-HSC numbers observed was proportional to the degree of neutropenia. These findings suggest that (1) *Cebpa* primes the neutrophilic lineage in HPCs but not LT-HSCs, and that (2) the negative effect on the stem cell numbers is more of an extrinsic event caused by neutropenic HPCs. To test this hypothesis, we used a sub-lethal transplantation model and tail-injected wild type bone marrow cells as controls (n=8) and homozygous *Cebpa*-enhancer deleted neutropenic bone marrow cell (n=16) in healthy recipient mice. The neutropenic bone marrow cells did not show any evidence of leukemic out-growth at any time post-transplantation. However, recipient mice transplanted with neutropenic bone marrow cells showed physical weakness accompanied by hypocellularity, dysplasia, and eventually a complete block in differentiation of recipient bone marrow cells, which occurred between 10-15 months post-transplantation. These findings indicate that *Cebpa*-enhancer deletion causes a cell autonomous neutropenia, which later imposes a negative effect on LT-HSCs causing latent severe bone marrow dysfunction as an extrinsic event.

INTRODUCTION

Differentiation and proliferation of hematopoietic stem and progenitors (HSPCs) requires tight regulation to maintain a constant production of blood cells in the bone marrow. HSPCs are highly heterogeneous and constitute a mixture of lineage precursors cells that make up the myeloid and the lymphoid lineage[1-3]. General HSPC transcription factors (gTFs) such as RUNX1, PU.1 or GATA2 control myeloid fate progenitors and activate lineage specific TFs (LTFs) to drive differentiation into various myeloid lineages including neutrophils, eosinophils, basophils, monocytes, macrophages and dendritic cells[4, 5].

C/EBPα is the major LTF that drives the neutrophilic differentiation program in HSPCs and myeloid fate progenitors [6-8]. General TFs in HSPCs bind and activate the *Cebpa* +37kb enhancer (+42kb in humans) to modulate *Cebpa* expression levels in neutrophilic differentiation. Knock out of *Cebpa* or its +37kb enhancer in mice show a striking neutropenia in both the peripheral blood (PB) and BM, followed by the expansion of myeloid fate progenitors [9-13]. *In vivo* reporter studies for *Cebpa* and its +37kb enhancer show that their activity increases upon differentiation i.e. 10-20% in HSPCs to 60-80% in myeloid progenitors. The most primitive long-term HSC population, defined by Lineage- Sca1+ cKit+ and CD48- CD150+ immuo-phenotypic markers, exhibits a detectable low 4% [13, 14]. Whether *Cebpa* is intrinsically critical for the function of LT-HSCs requires further investigation

Previous studies linked the function of *Cebpa* to neutrophil lineage priming in HSCs. Haseman et al. claimed that *Cebpa* binds to genes in HSPCs, which are expressed at later stages of neutrophilic differentiation[15]. In addition to its priming function, *Cebpa* is also critical for the function and maintenance of HSCs in the bone marrow. Conditional knockout of *Cebpa* increase the rate of apoptosis in HSCs, hence clarifying the observed decrease in numbers [15]. Previous studies showed that loss of *Cebpa* expression in HSPCs leads to an increase in self-renewal capacity via the upregulation of Bmi1 and *n-Myc* [10, 16], explaining the increase in proliferative potential in HSPCs. There are two reasons that jeopardize the claims raised by these studies: First, although claiming a biological function for *Cebpa* in HSCs, these studies were conducted in the LSK or HSPC BM cell populations, possibly because the number of stem cells required to conduct these experiments poses a challenge to obtain the required amounts. Second, the global effect on the stem cell population number observed is relatively more than the number of *Cebpa*-expressing HSCs claimed in the reporter studies[13,14]. This poses the question whether the reduced LT-HSC numbers observed in these models are cell intrinsic (cell autonomous) or caused by extrinsic events.

These results suggest that consequences of *Cebpa* knockout might lead to defects in differentiated lineages which can affect the HSC in a cell-extrinsic manner. A recent perspective review supports the hypothesis of HSC extrinsic effects in the *CEBPA* knockout phenotype: neutropenia caused by *Cebpa* knockout or enhancer deletion induces a feedback mechanism to activate HSCs and recover neutrophil differentiation. Prolonged neutropenia

will ultimately lead to the consumption and exhaustion of HSCs [15]. Another hypothesis addressed was that circulating neutropenic progenitor cells exert negative influence on the bone marrow, which ultimately leads to failure[17].

To investigate whether reduced HSC numbers are caused by neutropenia, we used the +37kb enhancer deletion model characterized by loss of *Cebpa* expression, neutropenia and decreased CD48-CD150+ LT-HSC numbers [18-21]. We found that *Cebpa* is expressed in LSKs but not in the more primitive LT-HSC population. The degree of neutropenia correlated with loss of LT-HSC number, which appeared to be enhancer-dosage and consequently *Cebpa*-level dependent. To further investigate the association between neutropenia and the loss of HSCs over a prolonged time-course, we conducted transplantation experiments of +37kb enhancer deleted BM cells into sub-lethally irradiated recipient animals. After 10-13 months post-transplantation, the mice transplanted with enhancer-deleted neutropenic cells exhibited BM defects characterized by hypo-cellularity, dysplasia and defects in cell lineage differentiation of recipient bone marrow progenitors. We conclude that an intrinsic deletion of +37kb enhancer causes persistent neutropenia that translates into a latent and extrinsic event that negatively influences hematopoiesis of the host.

RESULTS

Deficiency of the neutrophil gene expression program is cell autonomous in enhancer deleted HSPCs

The hematopoietic stem and progenitor cell (HSPC) population is composed of a heterogeneous multi-lineage pool of precursor cells [22]. To understand the role of the +37kb-*Cebpa*-enhancer in neutrophilic lineage priming in HSPCs, we investigated the transcriptome differences of FACS sorted bone marrow LSK fractions obtained from +37kb enhancer deleted (+37kb^{HOM}) (n=3) and wild type (+37kb^{WT}) (n=3) mice as control. Based on FPKM values, *Cebpa* is expressed at relatively low levels (Fig.S1) in wild type (+37kb^{WT}) sorted LSK fractions, indicating that only a small proportion of HSPCs that express *Cebpa* are primed for neutrophilic differentiation [13, 14]. In the +37kb^{HOM} LSK cells, *Cebpa* expression was significantly reduced (Fig.S1), confirming the regulation of *Cebpa* transcription by the +37kb enhancer occurs in immature HSPCs. In addition, other genes involved in neutrophilic priming and maturation were significantly lower in the +37kb^{HOM} LSK fraction, including *Camp*, *S100A8/9*, *Cish*, *Ccl3*, *Ngp* and *Ebi3* (Table1; Figure 1a and b)[23]. Surprisingly, the *Mpo* gene is increased in the +37kb^{HOM} LSKs. Ectopic reduction of *Cebpa* expression causes a stoichiometric cell lineage imbalance and a cell-fate bias towards other cell lineages [24-27]. In fact, *Tcf7* and *Bcl11b* T-cell related genes (Figure 1a) as well as the LTFs critical for pre-monoyctic/dendritic cells, *Irf8* and *Id2* (FigS1), were also increased in levels [28, 29]. These gene expression datasets demonstrate that deletion of the +37kb enhancer inactivates *Cebpa* and abrogates priming of neutrophil differentiation in LSK cells.

CEBPA-enhancer dosage correlates with neutropenia in bone marrow and peripheral blood

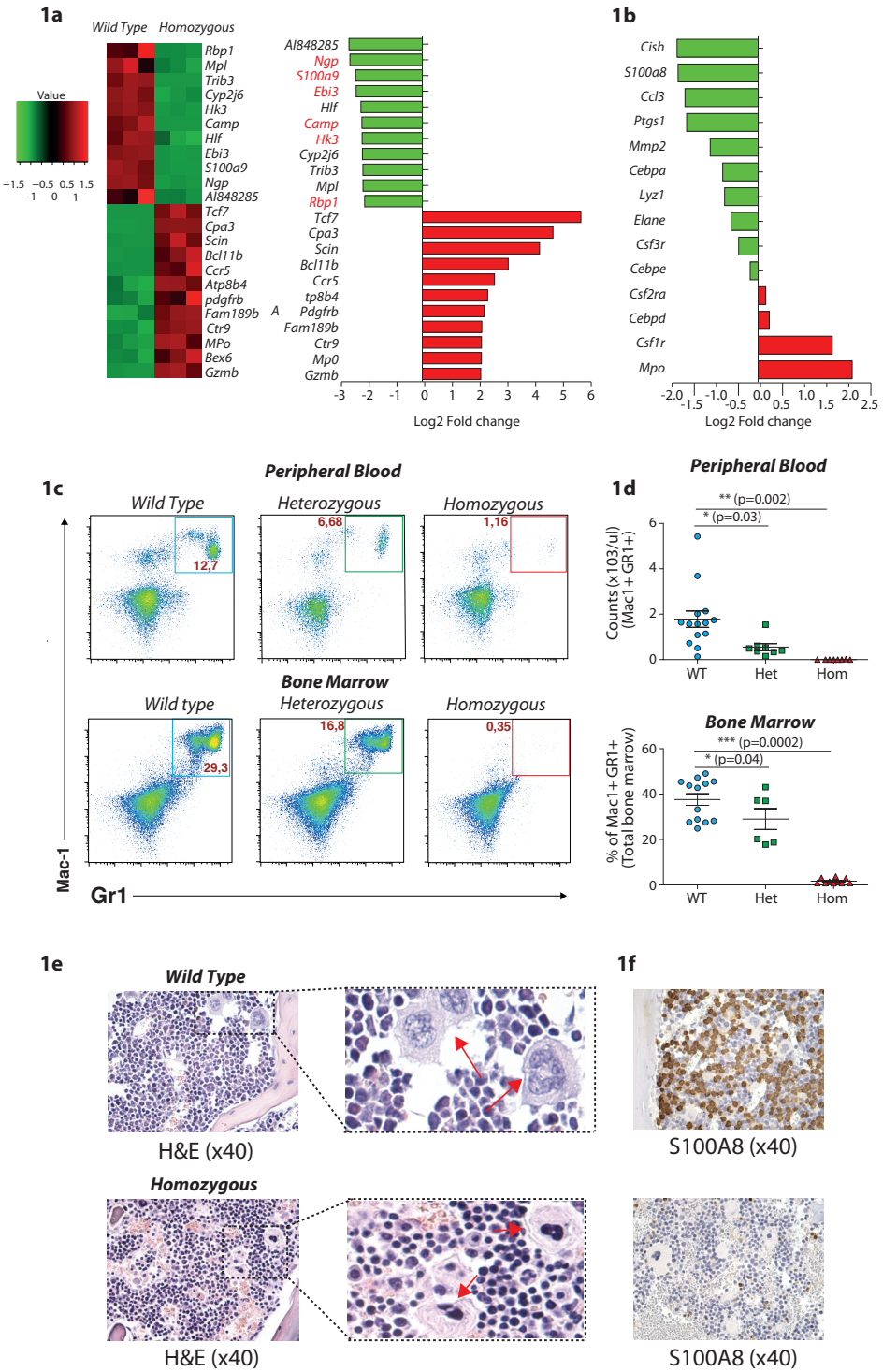
The loss of a neutrophil associated gene expression signature in the LSK population of +37kb^{HOM} mice led us to assess the downstream consequences on neutrophil maturation. We hypothesized that the *Cebpa*-enhancer dosage corresponds to the degree of neutrophil maturation. We therefore studied the effects of +37kb enhancer deletion on peripheral neutrophil counts in 6-8 weeks old +37kb^{HOM}(n=8), +37kb^{HET}(n=7), and +37kb^{WT} mice (n=14) (Figure 1c). The +37kb^{HET} mice, which are predicted to exhibit 50% *Cebpa* expression levels, had a significant decrease of Mac1+Gr1+ neutrophils when compared to controls (Figure 1c and 1d). The +37kb^{HOM} mice do not express *Cebpa* and showed low to no Mac1+/Gr1+ neutrophils in BM and PB (Figure 1c and 1d). Histopathology of the +37kb^{WT} and +37kb^{HOM} mice using hematoxylin and eosin (H&E) staining confirmed the marked loss of mature neutrophils in the BM (Figure 1e). In complement with these findings, we stained BM sections with the myeloid marker S100A8 and found 80-90% positivity in +37kb^{WT} mice, compared to 3-7% positivity in +37kb^{HOM} mice (Figure 1f). These data are in line with RNA-seq data, showing decreased *S100A8* mRNA levels in +37kb^{HOM} LSK cells (Figure 1a). Signs of dysplasia

were detected in the +37kb^{HOM} mice in the myeloid lineage and more predominantly in megakaryocytes, which appeared hypo-lobulated and in microcytic forms (red arrows in Figure 1e). The cellularity and bone architecture were normal with regular bone aligning cells and no changes in bone architecture. We conclude that a gradual reduction of enhancer dosage between heterozygous and homozygous enhancer knockouts, correlates with the degree of neutropenia and dysplasia occurring in the +37kb^{HOM} mice.

Perturbed LT-HSC signature in the +37kb-*CEBPA*-enhancer deleted HSPCs

We next investigated the gene expression profile of WT compared to +37kb^{HOM} LSKs and looked specifically into genes that constitute the HSC gene signature [30] (Figure 2a). In particular, *Mecom and Meis1* were severely down regulated in +37kb^{HOM} LSKs. It is well-established that *Mecom* is expressed in LT-HSCs and its expression is reduced upon further differentiation [31-33]. To determine whether the change of the HSC expression signature is an intrinsic consequence of enhancer deletion, we investigated whether *Cebpa* and *Mecom* are co-expressed in the same cells or whether the two genes are expressed in separate populations. Single cell RNA-sequencing (scRNAseq) data [34] from wild type LSKs and LT-HSCs (Figure 2b) showed that *Cebpa* is expressed in a small subset of LSK cells, but not in LT-HSCs, whereas *Mecom* is only expressed in LT-HSCs. Because *Cebpa* is not expressed in LT-HSCs but in a subset of LSKs, the reduction in *Mecom* expression in *the Cebpa-e*nhancer deleted LSKs is more likely to be caused by extrinsic effects on the LT-HSCs. Moreover, based on scRNA seq, the strong downregulation of *Meis1, Egr1, Gata2, Cxcr4* and *Stil* in +37kb enhancer deleted progenitors (Figure 2a), can neither be explained by intrinsic *Cebpa* loss, since the majority of cells expressing these genes in the LSK-fraction are *Cebpa* negative (Figure 2b).

Figure 1 : ***Cebpa-enhancer deletion inactivates neutrophil lineage priming in LSKs***. (1a left): Heat-map showing the 11 most downregulated genes (green) and the 11 most upregulated genes (red). (1a right) Fold change of the 11 most downregulated genes in green bars (neutrophilic associated genes are shown in red) and the 11 most upregulated genes are shown in red bars from an RNA-seq analysis of LSK +37kb^{HOM} (n=3) in comparison/relative to LSK wild type mice (n=3). (1b) Graph showing 14 selected neutrophil associated genes from RNA-seq data which are differentially regulated in LSK of +37kb^{HOM} mice when compared to wild type mice. The genes were selected from a previous publication[22]. (1c) FACS plots of peripheral blood and bone marrow analysis from wild type, +37kb^{HET} and +37kb^{HOM} using neutrophil markers Mac1 and GR1. (1d) Absolute numbers of Mac1 GR1 positive cells in peripheral blood and bone marrow of wild type, heterozygous +37kb enhancer deleted and homozygous +37kb enhancer deleted mice. (1e) Hematoxylin and Eosin stain of bone marrow sections (left) and S100A8 immuno-staining (right) used as a neutrophil differentiation marker on bone marrow sections from wild type and homozygous +37kb enhancer deleted mice. Red arrows showing normal megakaryocytes in wild type and abnormal small and hypo-lobulated megakaryocytes in homozygous+37kb enhancer deleted bone marrow. LSK: lineage, Sca-1, c-Kit; +37kb^{HOM}: +37kb homozygous deleted mice; +37kb^{HET}: +37kb heterozygous deleted mice. A non-parametric student t-test was used: pvalue= <0.05 (*); pvalue= <0.005 (**).


1a
Wild Type
Homozygous
Value
Rbp1
Mpl
Trib3
Cyp2j6
Hk3
Camp
Hlf
Ebi3
S100a9
Ngp
Al848285
Tcf7
Cpa3
Scin
Bcl11b
Ccr5
Atp8b4
pdgfrb
Fam189b
Ctr9
MPo
Bex6
Gzmb
Log2 Fold change
1b
Cish
S100a8
Ccl3
Ptgs1
Mmp2
Cebpa
Lyz1
Elane
Csf3r
Cebpe
Csf2ra
Cebpd
Csf1r
Mpo
Log2 Fold change
1c
Peripheral Blood
Wild Type
Heterozygous
Homozygous
12,7
6,68
1,16
Bone Marrow
Wild type
Heterozygous
Homozygous
29,3
16,8
0,35
Mac-1
Gr1
1d
Peripheral Blood
** (p=0.002)
* (p=0.03)
Counts (x103/ul) (Mac1+GR1+)
WT
Het
Hom
Bone Marrow
*** (p=0.0002)
* (p=0.04)
% of Mac1+GR1+ (Total bone marrow)
1e
Wild Type
H&E (x40)
Homozygous
H&E (x40)
1f
S100A8 (x40)
S100A8 (x40)

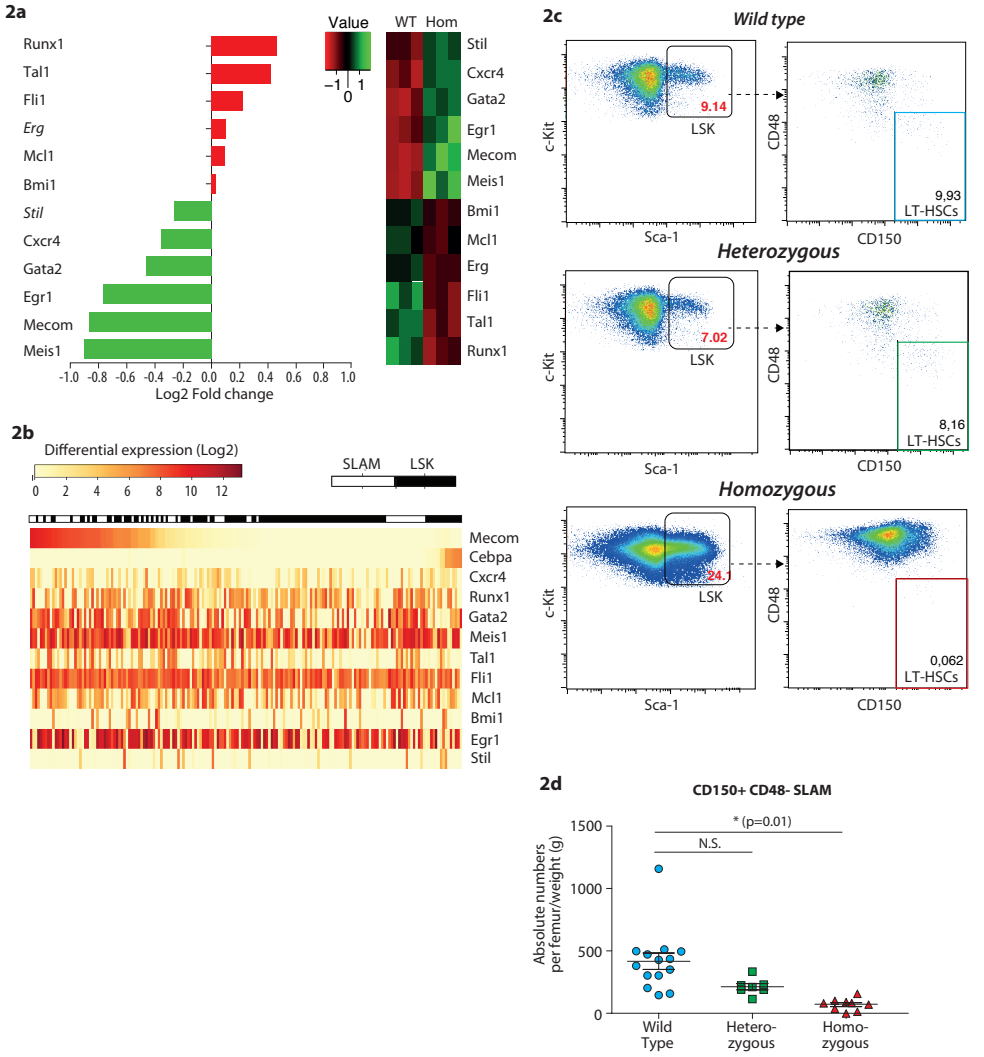


Figure 2. ***SLAM CD150+ HSC population reduction in +37kb^{HOM} mice is independent of the CEBPA-enhancer deletion***. Fold change of HSC related genes shown in a bar graph (left) and represented as heat-map (right) from RNA-seq analysis of LSK populations from homozygous deleted mice compared to wild type mice. (2b) Single cell analysis of wild type LSKs and wild type SLAM HSCs showing distinct and independent cell populations based on the expression of *Mecom* and *Cebpa*. (2c) FACS plots showing CD150+CD48- SLAM HSCs gated from LSK populations (lineage negative, Sca-1 positive, c-Kit positive) from bone marrows of wild type, +37kb^{HET} and +37kb^{HOM} mice. (2d) Absolute numbers of SLAM CD150+ HSCs from bone marrows of wild type, +37kb^{HET} and +37kb^{HOM} mice. A non-parametric student t-test was used: N.S.= not significant; pvalue= <0.05 (*); pvalue= <0.005 (**).

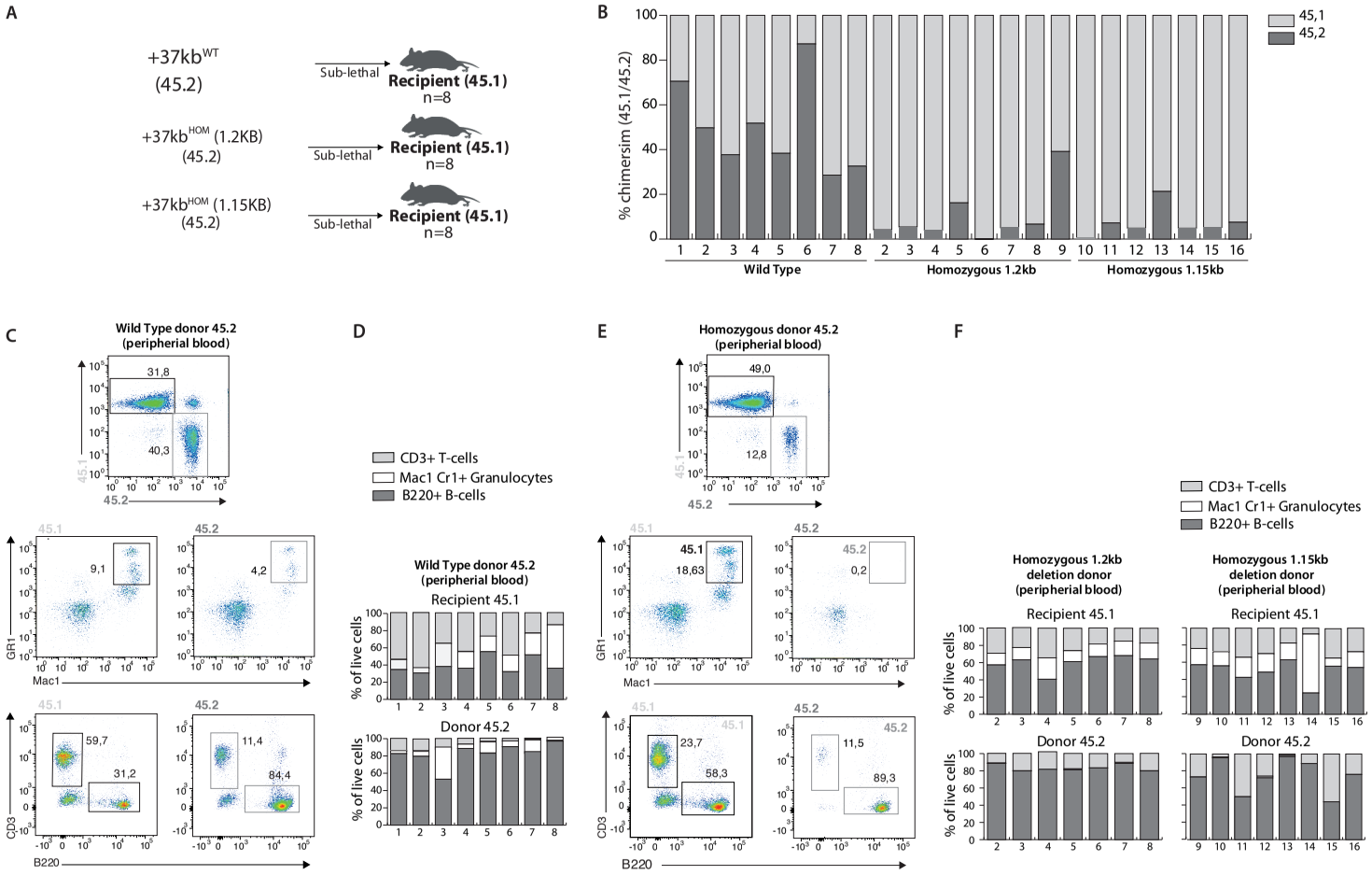
Reduction of CD150+ HSCs correlates with the degree of neutropenia in +37kb^{HET} and +37kb^{HOM} mice

The down-regulation of HSC-related genes in non-*Cebpa* expressing LT-HSCs of the +37kb^{HOM} mice led us to investigate the LT-HSCs numbers in +37kb^{HOM} and +37kb^{HET} BM relative to the BM from +37kb^{WT} mice, using SLAM CD48 and CD150 markers. Expansion of MPPs (CD48+ CD150-) was noticed as previously reported [11], probably compensating for loss of neutrophilic differentiation. This results in a left-shifted hematopoiesis that does not reach HSC levels but only leads to expansion of progenitors. The LT-HSCs were significantly reduced (p=<0.005) in the +37kb^{HOM} mice (figure 2c and d), coherent with enhancer dosage and the degree of neutropenia. In line with our RNA-seq results, the absolute numbers of the LT-HSC population are significantly reduced upon enhancer deletion, which explains the perturbed HSC gene expression signature and come to the conclusion that this occurs as an independent and extrinsic event.

Mice transplanted +37kb^{HOM} BM cells are neutropenic and exhibit low chimerism

To test the hypothesis whether neutropenia has a negative influence on the hematopoietic system as an extrinsic event, we conducted transplantation experiments. Three cohorts of sub-lethally irradiated recipient CD45.1 mice (8 mice per cohort) were transplanted with comparable numbers (See Materials and Methods) of Lin- cells of CD45.2 bone marrow cells derived from either +37kb^{WT}(N=8) +37$^{1.2}$kb (N=8) or the +37$^{1.15}$kb (N=7) homozygous deleted mice [11] (Figure 3a). These two different strains were generated via a random recombination event induced by CRISPR/Cas9, which resulted in two mouse strains with various enhancer deletion sizes; the expected 1.2kb (+37$^{1.2}$kb) deleted size and a 50bp shorter enhancer deletion of 1.15kb (+37$^{1.15}$kb) (Figure S2 and [11]). PB samples drawn twelve weeks after transplantation showed that mice transplanted with WT BM had comparable donor (45.2) to recipient (45.1) chimerism in the PB (Figure 3b) .The wild type mice showed reconstitution of both

the myeloid and the lymphoid lineages (Figure 3c and 3d). The two mouse cohorts transplanted with +37kb^{HOM} BM showed a weak donor-to-recipient chimerism (Figure 3c), with only 2-40% (Median: 3.4%) of blood cells derived from the 45.2 donor cell origin. Similar to the wild type transplanted host, mice transplanted with +37kb^{HOM} bone marrow cells showed a marked increase in CD3+ T-cells. The Mac1+Gr1+ cells were derived from the host only, with similar values compared to transplanted wild type controls. The absence of +37kb^{HOM} donor-derived Mac1+Gr1+ cells confirms the intrinsic and cell autonomous defect in neutrophilic differentiation caused by the +37kb *CEBPA*-enhancer deletion.



THESIS_Roberto_Avellino.indd 154 07-05-18 10:18

Figure 3. Transplanted enhancer deleted bone marrow cells show low chimerism and are neutropenic. (3a) Transplantation scheme representing donor (45.2) wild type, donor (45.2) homozygous 1.2kb deleted and donor (45.2) homozygous 1.15kb deleted (n=8) bone marrow cells transplanted in (45.1) recipients. Eight recipient mice per condition (n=24) were used in the experiment. (3b) Bar chart showing percentage of 45.1 and 45.2 cell chimerism in peripheral blood twelve weeks after transplantation of wild type and homozygous. (3c and 3e) Peripheral blood withdrawn from recipient mice transplanted with wild type (3e) and homozygous enhancer deleted (3d) bone marrow cells after twelve weeks of transplantation. FACS plots showing Mac1Gr1+ neutrophils, CD3 and B220 lymphocytes gated from 45,1 and 45.2 populations. (3d and 3f) Bar charts showing percentages of CD3+ T-cells, Mac1Gr1 Granulocytes and B220 B-cells calculated from peripheral blood of recipient mice transplanted with wild type (3d) and homozygous enhancer deleted (3f left and right) bone marrow cells after twelve weeks of transplantation.

Long exposure of circulating neutropenic cells causes cytopenia and dysplasia in recipient bone marrow

Despite the pre-leukemic properties of +37kb^{HOM} BM cells in vitro[11], we did not find any signs of leukemia in the recipient mice. Chimerism was low upon transplantation of the +37kb^{HOM} BM cells until the 10th month post-transplantation (Figure S4). Three mice transplanted with +37kb^{HOM} bone marrow cells died from unknown causes, with extensive palpable growth in the breast (n=2) and colon (n=1) (Table 2). Of the mice transplanted with WT BM, one had a CD34+ CD16/32 clonal population that was also Mac1Gr1 positive in bone marrow and spleen (data not shown). Follow up analyses on total white blood cell count, from the 10th to the 14th month post-transplantation, were conducted for both wild type and +37kb^{HOM} transplanted mice. For the remaining +37kb^{HOM} (13/16) and wild type (7/8) transplants, PB cell counts were analyzed monthly. Six of the +37kb^{HOM} showed physical weakness and a severe loss of wild blood cell counts was observed (Figure S5a). Five of these mice were analyzed further. We classified these mice as +37kb$^{HOM\text{-}severe}$, based on cellularity and degree of qualitative changes (dysplasia) in the BM. Cytopenia was only seen in the +37kb^{HOM} transplanted mice but not in mice transplanted with wild-type bone marrow cells (Figure 4a and 4b). We compared the results of the +37kb$^{HOM\text{-}severe}$ mice with wild-type sacrificed animals and mice transplanted with +37kb^{HOM} marrow which did not (yet) develop a cytopenia, which we classified as +37kb$^{HOM\text{-}N/I}$ (N/I = normal to intermediate) (n=8). White blood cell counts were comparable between the +37kb$^{HOM\text{-}N/I}$ mice and +37kb^{WT} controls. No differences were observed in hemoglobin levels, but peripheral blood platelet counts decreased significantly in the +37kb$^{HOM\text{-}severe}$ subgroup, and to a lesser extent in the +37kb$^{HOM\text{-}N/I}$ subgroup (Figure S5b and c). Given that on the day we sacrificed the mice the percentage chimerism of homozygous +37kb^{HOM} donor cells was close to 0% (median = 0.89%), all the cells analyzed were predominantly derived from the recipient (45.1). Bone marrow cellularity of the +37kb$^{HOM\text{-}severe}$ mice was significantly decreased more than two-fold compared to the other mice (Figure 5a). These findings comply with the reduced absolute numbers of the 45.1 LSK (Figure 5b) and LT-HSC population (Figure 5c) observed in the host. The histopathological examination results are in line with bone marrow cellular counts; +37kb$^{HOM\text{-}severe}$ bone marrows were characterized by dilated vessels filled with erythrocytes surrounded by sparsely interconnected hematopoietic cells with left-shifted hematopoiesis

as compared to wild type (Figure 5d) (Wild type; Severe). Bone marrow cellularity in the +37kb$^{HOM-N/I}$ mice varied from normocellular to hypocellular. Both the +37kb$^{HOM-severe}$ and the +37kb$^{HOM-N/I}$ cohorts showed microcytic hypo-lobulated megakaryocytes with dysplastic features (Figure 5b; Normal to Intermediate; Severe). This is in line with the decreased peripheral blood platelet counts (Figure S5c). These findings confirm that in the presence of circulating transplanted neutropenic bone marrow cells, the recipient hosts undergo bone marrow phenotypic changes with reduced HSC numbers, increased dysplastic features and a marked shift to the left hematopoiesis, suggesting cell differentiation failure.

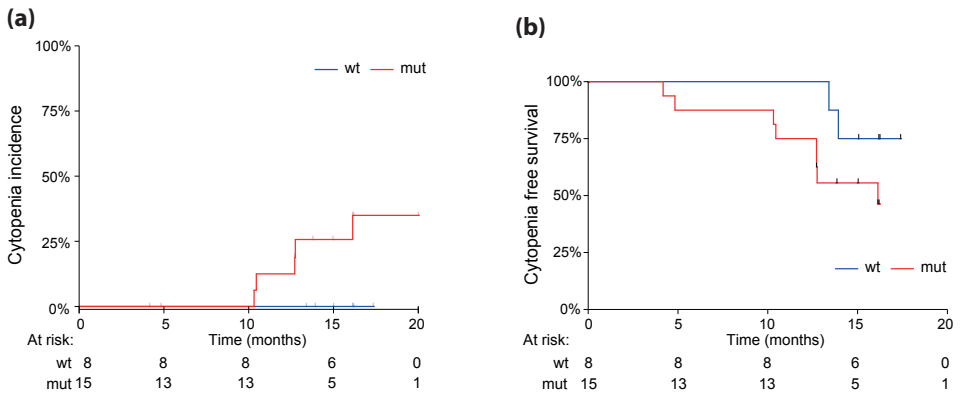
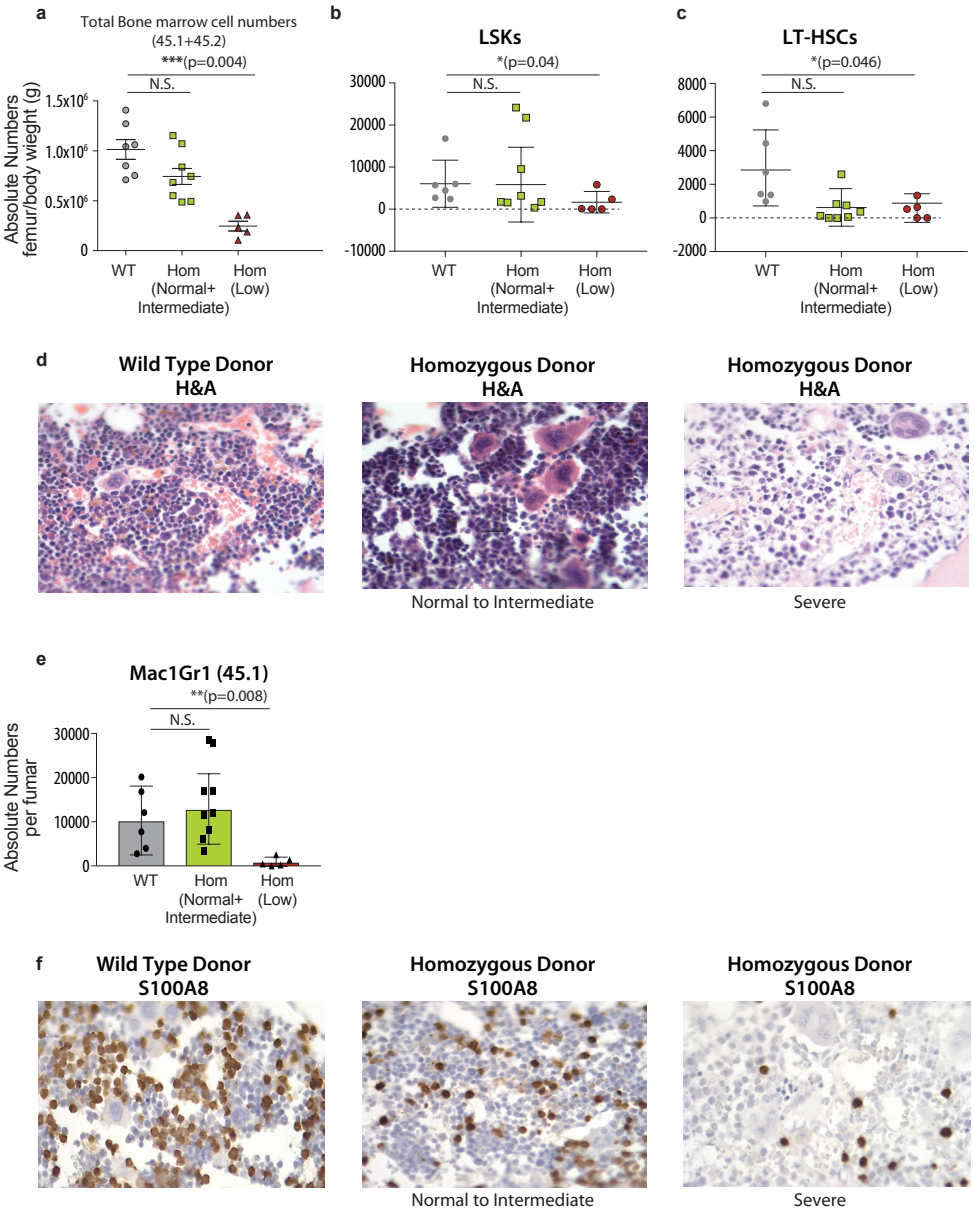


Figure 4. Cytopenia incidence and survival of transplanted mice. (a) Cytopenia incidence of mice transplanted with bone marrow of +37kb^{HOM} mice (red line) versus wild type mice (blue line). (b) Cytopenia-free survival of +37kb^{HOM} mice (red line) versus wild type mice (blue line).

Figure 5. Prolonged neutropenia exerts negative effects on HSPC numbers and myeloid cell differentiation. (4a) Total bone marrow cellularity per femur and corrected for body weight in grams of each mouse. The mice transplanted with homozygous bone marrow cells were divided into two subgroups: homozygous normal to intermediate and homozygous severe, based on bone marrow cellularity. (4b) Histological examination using hematoxylin and eosin on processed paraffin bone marrow sections of recipient mice: wild type control, homozygous donor with normal to intermediate phenotype and homozygous donor with severe phenotype. (4c) Lineage^{-ve}Sca-1^{pos}c-Kit^{pos} (LSK) absolute numbers calculated from LSK/lineage negative/live cells (7AAD) and corrected for total cellularity per femur and body weight (in grams) of each mouse analyzed. (4d-4f) Absolute numbers of Mac1 GR1, B-220 B-cells and CD3 T-cells from bone marrows of 45.1 recipient mice. (4g) LT-HSC absolute numbers calculated from CD48-CD150+/LSK/lineage negative/live cells (7AAD) and corrected for total cellularity per femur and body weight in grams of each mouse analyzed. (4h) Immunohistochemistry of S100A8 protein expression on histological bone marrow sections of recipient 45.1 mice transplanted with wild type and homozygous recipient. A non-parametric unpaired T-test was used: N.S.= not significant; pvalue= <0.05 (*).


a
Total Bone marrow cell numbers
(45.1+45.2)
***(p=0.004)
N.S.
Absolute Numbers
femur/body wieght (g)
1.5x10^6
1.0x10^6
0.5x10^6
0
WT
Hom
(Normal+
Intermediate)
Hom
(Low)
b
LSKs
*(p=0.04)
N.S.
30000
20000
10000
0
-10000
WT
Hom
(Normal+
Intermediate)
Hom
(Low)
c
LT-HSCs
*(p=0.046)
N.S.
8000
6000
4000
2000
0
-2000
WT
Hom
(Normal+
Intermediate)
Hom
(Low)
d
Wild Type Donor
H&A
Homozygous Donor
H&A
Homozygous Donor
H&A
Normal to Intermediate
Severe
e
Mac1Gr1 (45.1)
**(p=0.008)
N.S.
Absolute Numbers
per fumar
30000
20000
10000
0
WT
Hom
(Normal+
Intermediate)
Hom
(Low)
f
Wild Type Donor
S100A8
Homozygous Donor
S100A8
Homozygous Donor
S100A8
Normal to Intermediate
Severe

Neutropenia prevails as a major lineage-differentiation host-derived defect in the +37kb$^{HOM\text{-}severe}$ transplanted mice

Next, we assessed neutrophil differentiation in the bone marrow. Given that the chimerism on the day of sacrificing the mice was close to 0%, the neutrophils investigated in these transplantation experiments are all derived from the host. Thus, any changes in neutrophil numbers reflect abnormal myelopoiesis coming from the bone marrow of the recipient mouse. Mac1Gr1 markers on selected host bone marrow cells showed a marked reduction of the neutrophil absolute counts in the +37kb$^{HOM\text{-}severe}$, whereas the +37kb$^{HOM\text{-}N/I}$ mice showed normal to low counts, within a range comparable to the wild type controls (Figure 5e). In adjunct with FACS analysis, we stained histological BM sections for protein expression of S100A8 as a marker for myeloid differentiation [34, 35]. BM from+37kb$^{HOM\text{-}severe}$ mice was almost devoid of S100A8 expression, which correlates with the loss of myeloid cells observed in the FACS analysis (Figure 5f). These results also comply with the reduction in *S100*A8/*A9* gene expression from our RNA-seq data in the *Cebpa*-enhancer deleted LSKs (Figure 1a and b). These findings suggest that transplanted +37kb$^{HOM\text{-}severe}$ cells induce a negative systemic influence on the bone marrow of the host resulting in bone marrow failure.

DISCUSSION

In this study, we show that the role of *Cebpa* and its enhancer are responsible for neutrophil lineage priming in the LSK progenitors (or hematopoietic and stem cell progenitors - HSPCs), earlier than it was proposed. A recent study reported that *Cebpa* drives unilineage neutrophilic differentiation in the LK bone marrow population, a population that mainly constitutes progenitors of the myeloid lineage[23]. This is supported with our RNA-seq data showing that the expression of neutrophil associated genes in the LSK population is strongly down regulated upon enhancer deletion, indicating that *Cebpa* activates the neutrophil gene expression program and prime LSK progenitors for neutrophil differentiation[16]. The LSK population is heterogeneous and, by using SLAM markers, one can distinguish between the early primed HSPCs from long-term HSCs (LT-HSCs) that are able to re-populate the bone marrow throughout the lifetime of an organism[30]. Our findings support this emerging concept of unilineage hematopoiesis by showing that neutrophil lineage priming occurs at more primitive stages, i.e. in the LSKs rather than in the LK progenitors[23].

Based on previous studies [10-12, 15], the main question in this study was whether *Cebpa* and its enhancer are linked to an intrinsic function in HSCs. We showed that knockout of *Cebpa* or the +37kb Cebpa-enhancer has a dual intrinsic influence on neutrophil lineage priming and HSC numbers. *Cebpa*-enhancer knockout causes neutropenia in blood and bone marrow and lacks a significant proportion of LT-HSCs. In our study, we dissected these two events in separate models, i.e., a *Cebpa*-enhancer knockout model to study the biology of neutrophil lineage priming in HSPCs and a sub-lethal transplantation model to investigate the effects of prolonged donor neutropenia on the bone marrow integrity of the host as a HSC functional output.

LT-HSCs constitute a very small population of the LSK bone marrow fraction and constantly undergo different kinetics such as dormancy, quiescence, self-renewal and cell division, thus making it difficult to delineate a standard gene expression program to define them[36]. Deletion of the enhancer resulted in low LT-HSC numbers and downregulated HSC related genes. We used *Mecom* expression as a HSC marker in order to delineate any *Cebpa* expressing LT-HSCs. *Mecom* encodes a transcription factor known as EVI1 or PRDM3 that is specifically expressed in approximately 60% of LT-HSC population. Based on functional experiments, *Mecom* is critical for the development, maintenance and proliferation of HSCs [32]. Unlike previously reported [16], *Cebpa* was not expressed in any of the LT-HSCs analysed by single-cell RNA-seq and all *Mecom*+ LT-HSCs were *Cebpa* negative for expression. *Cebpa* was expressed in a subset of LSKs, which verifies that the negative influence on LT-HSCs upon enhancer deletion is not intrinsically related.

Cebpa-enhancer and *Cebpa* knockout mice have very low LT-HSC numbers, which engraft poorly in lethally irradiated hosts[15, 37], possibly due to the low LT-HSC numbers. We used a sub-lethal irradiation setup to allow for reconstitution to take place from the host bone

marrow because of the low chimerism derived from the transplanted donor bone marrow cells. Even though the enhancer deleted bone marrow donor cells exhibit pre-leukemic characteristics, including a differentiation block and increased clonogenic capacity in vitro[11], the donor bone marrow cells did not outcompete the host cells but retained their differentiation block characteristics.

How does a small population of donor cells blocked in differentiation can have such a strong negative impact on the hematopoietic system of the host? The lineage differentiation output of Mac1Gr1 neutrophils derived from the host was reduced to very low levels on the day we sacrificed the mice. This phenomenon can be explained by a negative influence the donor cells might exert on specific bone marrow niches that support myeloid development. Block in differentiation might fairly lead to aberrant production and secretion of molecules in the bone marrow that compromise its integrity and normal function [35]. In addition, there is an increased body of evidence that myeloid fate progenitors tend to form patches or clusters in pre-malignant and malignant conditions to influence normal hematopoiesis [38]. This is in line with recent studies that investigated how pre-leukemic and leukemic cells attenuate normal hematopoiesis by the secretion of cytokines, growth factors or exosomes [39-41].

Investigating different niche populations such as the mesenchymal stromal progenitors and osteoblasts would also provide a deeper insight whether such extrinsic event caused by neutropenia involves the niche directly. Our preliminary data suggest bone marrow niche involvement since we observe changes in cell numbers of bone marrow niche cells in the recipient mice transplanted with +37kbHOM (Fig.S7). We observed that megakaryocytes, which 6are also an important part of the HSC niche, exhibited dysplastic features thus, providing a slight indication that distinct niche constituents are also targets for deregulation in this experimental setup.

We conclude that the *CEBPA*-enhancer mediates neutrophil lineage priming in HSPCs by activating C/EBP**α** expression and neutrophilic differentiation. Deletion of the enhancer causes neutropenia as a cell intrinsic and autonomous event. Prolonged neutropenia is later translated into an extrinsic event that suppresses normal hematopoiesis of the host. Practically, in humans this phenomenon is hardly seen since patients with severe neutropenia are treated immediately with intra-cutaneous GCSF (granulocytic colony stimulating factor) to boost up their neutrophil counts. Although the clinical relevance of this phenomenon is yet to be discovered, it is becoming more evident that aberrant circulating bone marrow cells exert effects on localized bone marrow areas to negatively influence normal hematopoiesis. They may also circulate to other peripheral organs where they trigger the onset of complex non-hematological disorders such as neurodegenerative diseases[42].

MATERIALS AND METHODS

(1) RNA sequencing

Total sample RNA was extracted using Trizol with Genelute LPA (Sigma) as a carrier and SMARTer Ultra Low RNA kit for Illumina Sequencing (Clontech) was used for cDNA synthesis according to the manufacturer's protocol. The cDNA was sheared with the Covaris device and further processed according to the TruSeq RNA Sample Preparation v2 Guide (Illumina). The amplified sample libraries were subjected to paired-end sequencing (2 x 75 bp) and aligned against mm10 using TopHat v256. All reads were aligned against genes annotated in the RefSeq Transcriptome database and remaining non-aligned reads were aligned against the full genome. Gene expression levels were quantified by the fragments per kilobase of exon per million fragments mapped (FPKM) statistic as calculated by Cufflinks. Hierarchical clustering analysis was performed on the FPKM values using complete linkage as clustering methodology and Euclidean distance as distance measure using the G-plots package in the R environment. Read counts were determined with HTSeq-count and subsequently used for differential expression analysis in DESeq2, with default parameters, in the R environment. Multiple testing correction was achieved by performing the Benjamini-Hochberg procedure on the calculated p-values to control the False Discovery Rate (FDR).

(2) Single cell RNA sequencing

Existing murine lineage-negative Sca1+ CD117+ cells (LSK) bone marrow progenitor single-cell RNA-Seq data from the Fluidigm microfluidics platform (GSE70244) aligned using the RSEM algorithm [34]. SLAM captured HSCs were enriched, as previously described and profiled at by the Cincinnati Children's Hospital Single-Cell Genomics Core as previously described [34] and co-analyzed with this existing LSK data. For expression visualization, log2 RSEM TPM values are displayed for selected genes.

(3) Mice and transplantation procedures

The strains of +37kb enhancer$^{1.2kb}$ and +37kb enhancer$^{1.15kb}$ deleted mice generated by CRISPR/Cas9 have been previously described. Animals were maintained in specific pathogen free conditions in the Experimental Animal Center of ErasmusMC (EDC). All mice are C57BL/6. Non-transplanted mice were sacrificed for FACS analysis between 4-8 weeks of age. For transplantation experiments, 24 mice were sub-lethally irradiated. On the same day, pooled bone marrows of wild type (CRISPR/Cas9 treated) mice, +37kb enhancer$^{1.2kb}$ and +37kb enhancer$^{1.15kb}$ deleted homozygous mice were harvested in PBS/5%FCS and injected intravenously in tails of (45.1) female mice. The 45.2 mice sacrificed for bone marrow harvesting were four weeks of age. Three cohorts of female 45.1 mice, were transplanted with wild type(n=8), 37kb enhancer$^{1.2kb}$(n=8) and +37kb enhancer$^{1.15kb}$(n=8) bone marrow cells. The number of transplanted bone marrow cells was corrected based on the LSK

population which was 5 times bigger in the homozygous compared to the wild type i.e. wild type = $1x10^6$ and homozygous = $2x10^5$ bone marrow cells per mouse/condition. All mice were sacrificed in a CO_2 chamber. Animal studies were approved by the Animal Welfare/ Ethics Committee of the EDC in accordance with legislation in the Netherlands (approval No. EMC 2067, 2714, 2892, 3062).

(4) Flowcytometry and sorting

Flowcytometry was carried out on the LSRII and the FACSCanto II (Becton Dickinson) was used for cell sorting, using the following fluorescent antibodies: CD11B- APC/GR1-FITC/ B220-PE/CD45 PerCP CY5/LIN bio-cocktail streptavidin-pacific orange/cKIT-APC/SCA1-PB/CD48-FITC/CD150-PE-CY7/. All antibodies were purchased from BD Biosciences and Biologend. Sorted LSK fractions were collected in 500µl PBS with 5% FCS, spun down and re-suspended in 800µl of Trizol and used for RNA-seq.

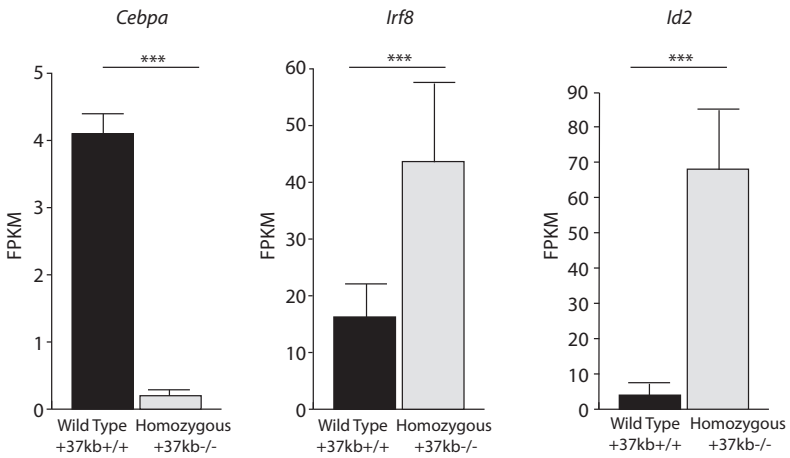
(5) Histology sections and staining

Bone marrows from mice were treated with 4% formaldehyde overnight, decalcified (EDTA), dehydrated and embedded in paraffin. Hematoxylin-eosin (H&E) staining was done according to routine protocols. For immunohistochemical staining, antigen retrieval of mice tissues was performed by pressure cooker treatment in citrate buffer (pH 6.0; Zytomed) and by microwave-heating for human samples. After Peroxidase blocking (DAKO for mice tissues; 3%, AppliChem Panreac for human tissue) and incubation in blocking solution (5% Normal Goat serum, DAKO), sections were incubated using anti-S100a8 (Abcam, ab92331, 1:200). Biotinylated goat-anti-mouse/rabbit (DAKO, K5001) was used as secondary antibody. The avidin-biotin complex (ABC)–horseradish peroxidase (HRP) kit (DAKO, K5001) was applied for color development. Images were obtained using a Zeiss Axioplan microscope equipped with a Zeiss AxioCam ERc5s.

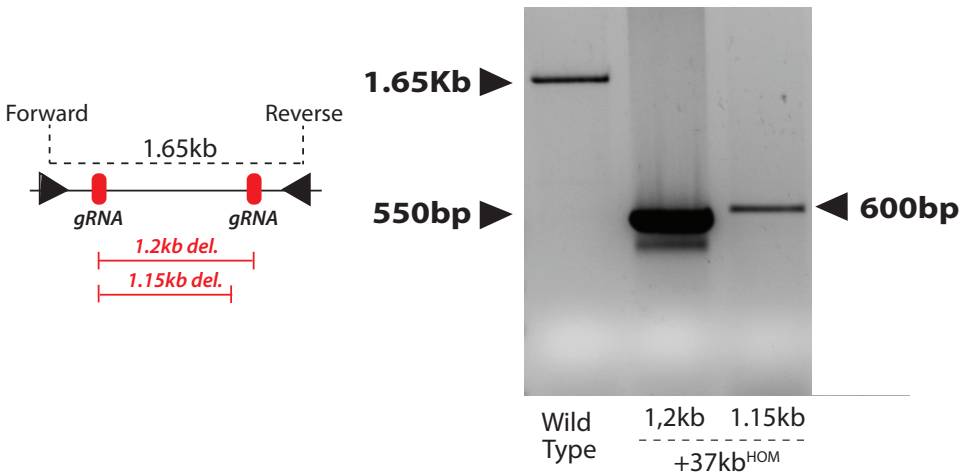
(6) Statistics

Statistical analysis was performed using Prism 7 (GraphPad Software). Unpaired, two-tailed Student's t test (single test) was used to evaluate statistical significance, defined as $p < 0.05$. All results in graphs are mean value ± SEM. Survival analysis: Cytopenia-free survival was defined as the start date (please describe t=0 when the experiment started) to cytopenia or death from any cause, whichever came first. Survival estimates were graphically represented with the Kaplan-Meier method. The log-rank test was used to compare the survival distributions of the groups. P-values are two sided and p-values <0.05 were considered statistically significant. Statistical analysis were performed in STATA statistical software, Release 15.1 (Stata, College Station, TX, USA).

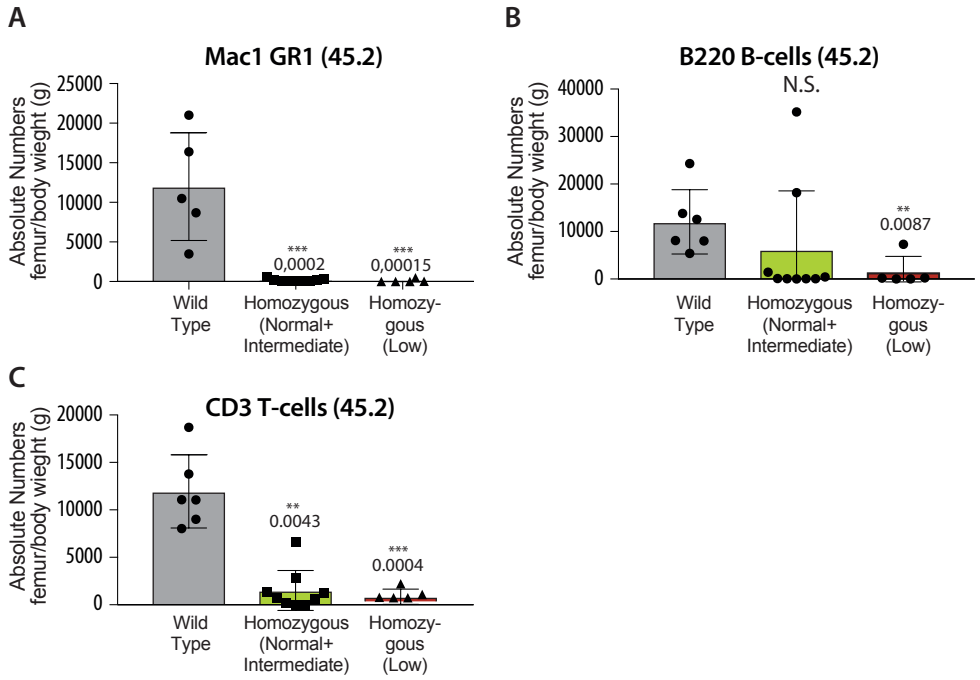
SUPPLEMENTARY FIGURES



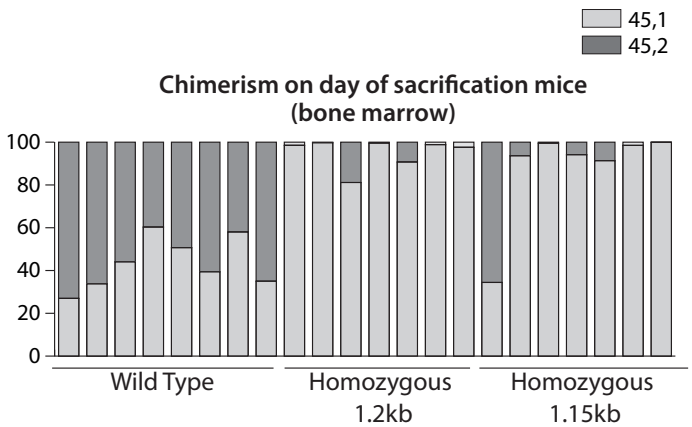
Supplementary Figure1. FPKM values of Cebpa, Irf8 and Id2, from RNA-seq experiments conducted on LSK sorted fractions of wild type (n=3) and homozygous (n=3) mice.



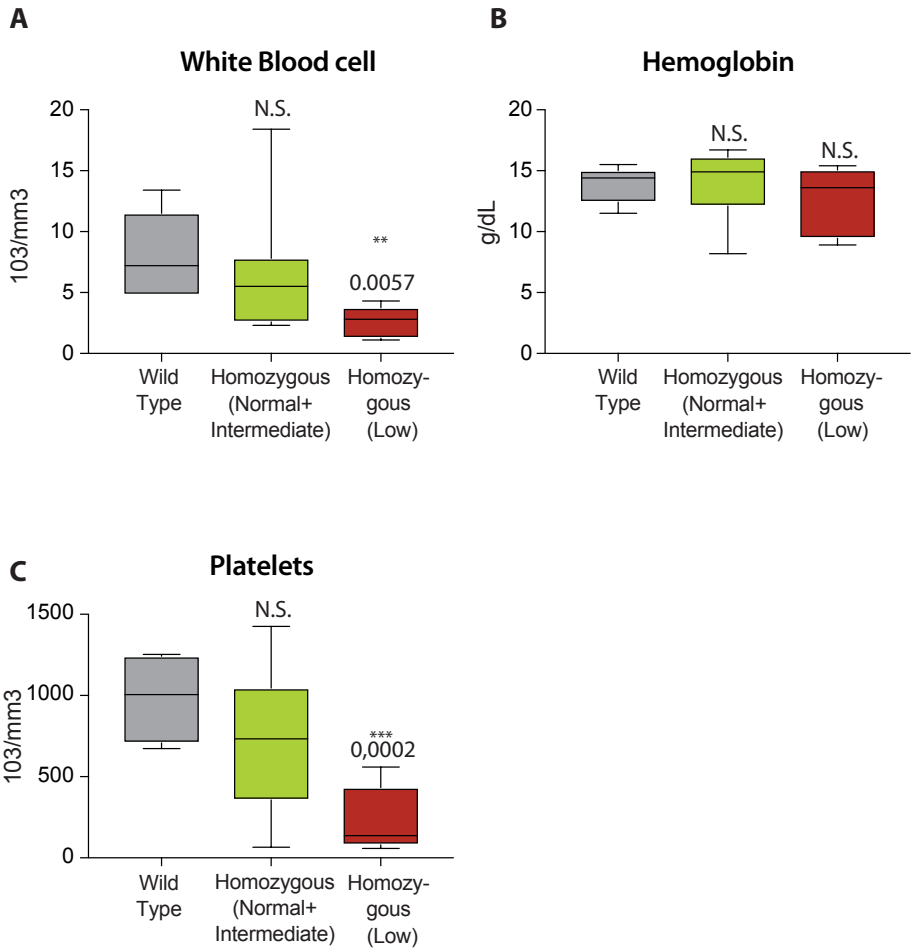
Supplementary Figure 2. Genotyping of tail genomic DNA by PCR of wild type mice and the enhancer deleted homozygous strains, +37kb$^{1.2kb}$ and +37kb$^{1.15kb}$ homozygous mice.



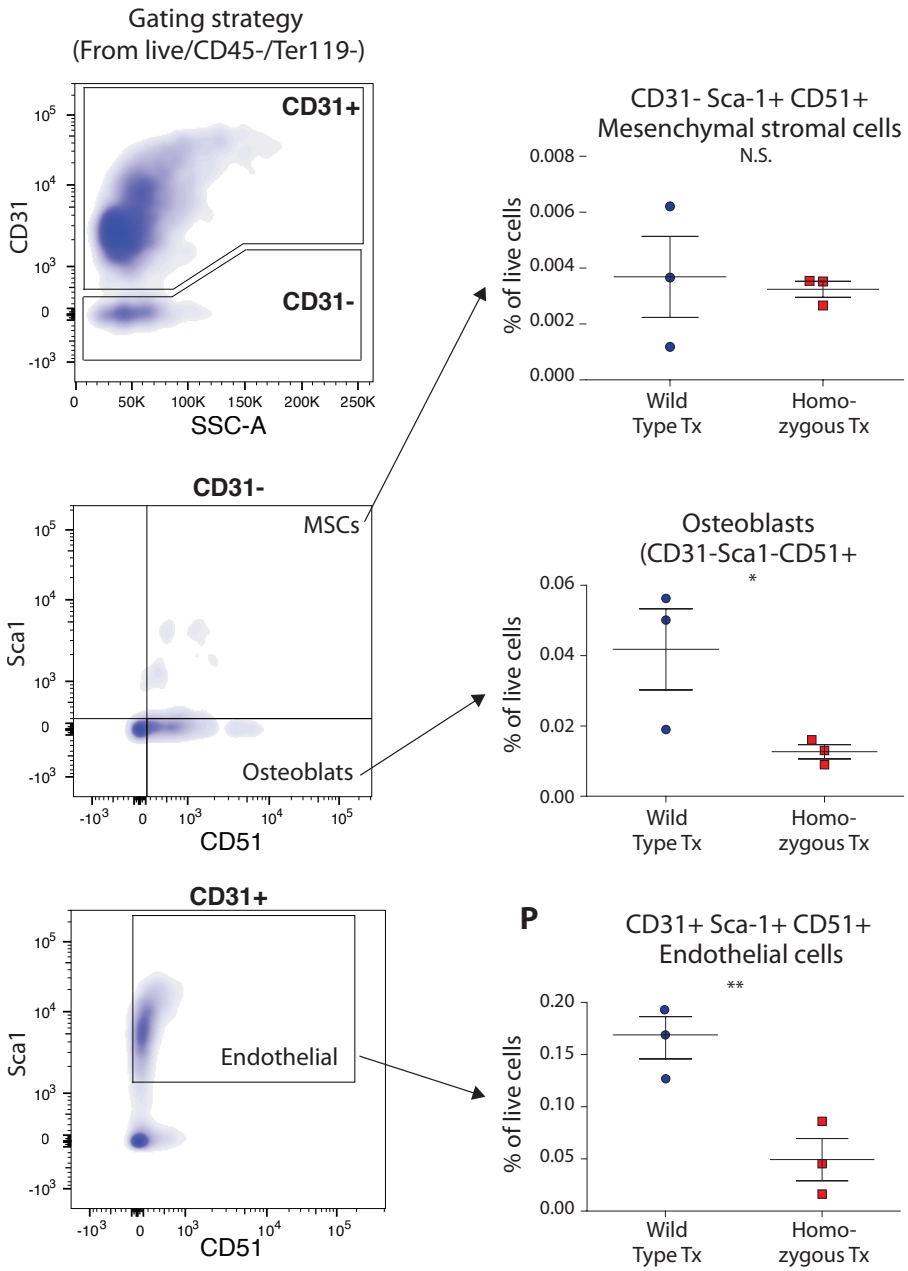
Supplementary Figure 3. Absolute numbers of Mac1Gr1 neutrophils, B220 B-cells and CD3 T-cells derived from donor 45.2 mice on the day the mice were sacrificed.



Supplementary Figure 4. Percentage chimerism between donor (45.1) and (45.2) mice on the day the mice were sacrificed.



Supplementary Figure 5. Peripheral blood counts of mice on the day of sacrifice made between wild type, homozygous (normal+intermediate), homozygous (low). A non-parametric unpaired T-test was used: N.S.= not significant.



Supplementary Figure 6. Flow cytometry analysis of on non-hematopoietic bone marrow stromal cells gated from live cells (7AAD), excluding CD45+ hematopoietic cells and Ter119+ red cells. CD31 marker was used to differentiate CD31+ endothelial cells (CD31+ CD51+ Sca-1+) from CD31- mesenchymal stromal cells (CD31- CD51+ Sca-1+) and CD31- osteoblast progenitors (CD31- CD51+ Sca-1-). Analysis was done in mice transplanted with wild type bone marrow cells (Wild Type Tx: n=3) and mice transplanted with homozygous enhancer deleted bone marrow cells (Homozygous Tx: n=3). The homozygous TX mice were derived from the+37kb$^{HOM\text{-}N/I}$ (normal to intermediate) subgroup. Analysis for statistical significance was carried out using student T test (p=<0.05*; p=<0.01**).

table 1

Gene	Protein (full name)	Fold change (down)	Function
AI848285	Shisa Family Member 8	-2,644665132	Unknown
Ngp	Neutrophilic granule protein	-2,609837162	Neutrophil cytoplasmic granule with cysteine-type endopeptidase inhibitor activity
S100a9	S100 calcium binding protein A9	-2,405284626	Regulation of inflammatory processes; increase neutrophilic bactericidal activity
Ebi3	IL-27 subunit Beta	-2,385582175	A heterodimeric cytokine which functions in innate immunity
Hlf	Hepatic leukemia Factor	-2,21435241	Leucine zipper DNA binding factor
Camp	cathelicidin antimicrobial peptide	-2,179885084	Neutrophil anti-bacterial peptide; Neutrophil degranulation
Hk3	Hexokinase 3	-2,171658723	Glycolysis; Neutrophil degranulation
Cyp2j6	cytochrome P450, family 2, subfamily j, polypeptide 6	-2,161487127	Monoxygenase; metal binding
Trib3	tribbles pseudokinase 3	-2,142603408	Protein kinase inhibitor
Mpl	MPL proto-oncogene, thrombopoietin receptor	-2,140558486	Thrombopoietin receptor activity
Rbp1	Retinol Binding protein 1	-2,080710243	Regulation of granulocyte differentiation

Gene	Protein (full name)	Fold change (up)	Function
Tcf7	Transcription factor 7	5,715925951	Transcriptional activator involved in T-cell lymphocyte differentiation
Cpa3	Mast cell Carboxypeptidase 3	4,710997644	Carboxypeptidase; metabolism of Angiotensinogen to Angiotensins
Scin	Adseverin	4,22659805	Ca2+-dependent actin filament-severing protein; required for megakaryocytic differentiation
Bcl11b	B-cell lymphoma/leukemia	3,096578829	Key regulator of both differentiation and survival during thymocyte development in mammals
Ccr5	Chemokine Receptor 5	2,601170803	Receptor for a number of inflammatory CC-chemokines
Atp8b4	Phospholipid transporting ATPase	2,355680865	Transport molecule
Pdgfrb	Platelet derived growth factor Beta	2,229151017	Tyrosine-protein kinase that acts as cell-surface receptor
Fam189b	Protein FAM189B	2,145325689	lysosomal enzyme glucosylceramidase;
Ctr9	RNA polymerase-associated protein CTR9 homolog	2,133128665	Associates with RNA polymerase II; Chromatin states and transcriptional control with MLL proteins
Mpo	Myeloperoxidase	2,130607807	Microbicidal activity released by polymorphonuclear leucocytes
Gzmb	Granzyme B	2,107236859	Enzyme necessary for target cell lysis in cell-mediated immune responses

REFERENCES

1. Zhong, C. and J. Zhu, *Transcriptional regulators dictate innate lymphoid cell fates.* Protein Cell, 2017. **8**(4): p. 242-254.
2. Perie, L., et al., *The Branching Point in Erythro-Myeloid Differentiation.* Cell, 2015. **163**(7): p. 1655-62.
3. Collombet, S., et al., *Logical modeling of lymphoid and myeloid cell specification and transdifferentiation.* Proc Natl Acad Sci U S A, 2017. **114**(23): p. 5792-5799.
4. Mercier, F.E. and D.T. Scadden, *Not All Created Equal: Lineage Hard-Wiring in the Production of Blood.* Cell, 2015. **163**(7): p. 1568-70.
5. Naik, S.H., et al., *Diverse and heritable lineage imprinting of early haematopoietic progenitors.* Nature, 2013. **496**(7444): p. 229-32.
6. Diffner, E., et al., *Activity of a heptad of transcription factors is associated with stem cell programs and clinical outcome in acute myeloid leukemia.* Blood, 2013. **121**(12): p. 2289-300.
7. Beck, D., et al., *Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes.* Blood, 2013. **122**(14): p. e12-22.
8. Wilson, N.K., et al., *Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators.* Cell Stem Cell, 2010. **7**(4): p. 532-44.
9. Zhang, D.E., et al., *Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice.* Proc Natl Acad Sci U S A, 1997. **94**(2): p. 569-74.
10. Zhang, P., et al., *Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha.* Immunity, 2004. **21**(6): p. 853-63.
11. Avellino, R., et al., *An autonomous CEBPA-enhancer specific for myeloid-lineage priming and neutrophilic differentiation.* Blood, 2016. **127**(24): p. 2991-3003.
12. Cooper, S., H. Guo, and A.D. Friedman, *The +37 kb CEBPA-enhancer Is Critical for Cebpa Myeloid Gene Expression and Contains Functional Sites that Bind SCL, GATA2, C/EBPalpha, PU.1, and Additional Ets Factors.* PLoS One, 2015. **10**(5): p. e0126385.
13. Guo, H., O. Ma, and A.D. Friedman, *The Cebpa +37-kb enhancer directs transgene expression to myeloid progenitors and to long-term hematopoietic stem cells.* J Leukoc Biol, 2014. **96**(3): p. 419-26.
14. Wolfler, A., et al., *Lineage-instructive function of C/EBPalpha in multipotent hematopoietic cells and early thymic progenitors.* Blood, 2010. **116**(20): p. 4116-25.
15. Hasemann, M.S., et al., *C/EBPalpha is required for long-term self-renewal and lineage priming of hematopoietic stem cells and for the maintenance of epigenetic configurations in multipotent progenitors.* PLoS Genet, 2014. **10**(1): p. e1004079.
16. Ye, M., et al., *C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence.* Nat Cell Biol, 2013. **15**(4): p. 385-94.
17. Avellino, R. and R. Delwel, *Expression and regulation of C/EBPalpha in normal myelopoiesis and in malignant transformation.* Blood, 2017. **129**(15): p. 2083-2091.
18. Oguro, H., L. Ding, and S.J. Morrison, *SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors.* Cell Stem Cell, 2013. **13**(1): p. 102-16.
19. Yilmaz, O.H., M.J. Kiel, and S.J. Morrison, *SLAM family markers are conserved among hematopoietic stem cells from old and reconstituted mice and markedly increase their purity.* Blood, 2006. **107**(3): p. 924-30.
20. Kim, I., et al., *Enhanced purification of fetal liver hematopoietic stem cells using SLAM family receptors.* Blood, 2006. **108**(2): p. 737-44.
21. Kiel, M.J., et al., *SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells.* Cell, 2005. **121**(7): p. 1109-21.

22. Paul, F., et al., *Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors.* Cell, 2015. **163**(7): p. 1663-77.
23. Paul, F., et al., *Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors.* Cell, 2016. **164**(1-2): p. 325.
24. De Obaldia, M.E., et al., *T cell development requires constraint of the myeloid regulator C/EBP-alpha by the Notch target and transcriptional repressor Hes1.* Nat Immunol, 2013. **14**(12): p. 1277-84.
25. Feng, R., et al., *PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells.* Proc Natl Acad Sci U S A, 2008. **105**(16): p. 6057-62.
26. Di Tullio, A., et al., *CCAAT/enhancer binding protein alpha (C/EBP(alpha))-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation.* Proc Natl Acad Sci U S A, 2011. **108**(41): p. 17016-21.
27. van Oevelen, C., et al., *C/EBPalpha Activates Pre-existing and De Novo Macrophage Enhancers during Induced Pre-B Cell Transdifferentiation and Myelopoiesis.* Stem Cell Reports, 2015. **5**(2): p. 232-47.
28. Kurotaki, D., et al., *IRF8 inhibits C/EBPalpha activity to restrain mononuclear phagocyte progenitors from differentiating into neutrophils.* Nat Commun, 2014. **5**: p. 4978.
29. Welner, R.S., et al., *C/EBPalpha is required for development of dendritic cell progenitors.* Blood, 2013. **121**(20): p. 4073-81.
30. Pietras, E.M., et al., *Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions.* Cell Stem Cell, 2015. **17**(1): p. 35-46.
31. Goyama, S., et al., *Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells.* Cell Stem Cell, 2008. **3**(2): p. 207-20.
32. Kataoka, K., et al., *Evi1 is essential for hematopoietic stem cell self-renewal, and its expression marks hematopoietic cells with long-term multilineage repopulating activity.* J Exp Med, 2011. **208**(12): p. 2403-16.
33. Kustikova, O.S., et al., *Activation of Evi1 inhibits cell cycle progression and differentiation of hematopoietic progenitor cells.* Leukemia, 2013. **27**(5): p. 1127-38.
34. Olsson, A., et al., *Single-cell analysis of mixed-lineage states leading to a binary cell fate choice.* Nature, 2016. **537**(7622): p. 698-702.
35. Zambetti, N.A., et al., *Mesenchymal Inflammation Drives Genotoxic Stress in Hematopoietic Stem Cells and Predicts Disease Evolution in Human Pre-leukemia.* Cell Stem Cell, 2016. **19**(5): p. 613-627.
36. Kowalczyk, M.S., et al., *Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells.* Genome Res, 2015. **25**(12): p. 1860-72.
37. Guo, H., S. Cooper, and A.D. Friedman, *In Vivo Deletion of the Cebpa +37 kb Enhancer Markedly Reduces Cebpa mRNA in Myeloid Progenitors but Not in Non-Hematopoietic Tissues to Impair Granulopoiesis.* PLoS One, 2016. **11**(3): p. e0150809.
38. Herault, A., et al., *Myeloid progenitor cluster formation drives emergency and leukaemic myelopoiesis.* Nature, 2017. **544**(7648): p. 53-58.
39. Hornick, N.I., et al., *AML suppresses hematopoiesis by releasing exosomes that contain microRNAs targeting c-MYB.* Sci Signal, 2016. **9**(444): p. ra88.
40. Kumar, B., et al., *Acute myeloid leukemia transforms the bone marrow niche into a leukemia-permissive microenvironment through exosome secretion.* Leukemia, 2017.
41. Boyd, A.L., et al., *Acute myeloid leukaemia disrupts endogenous myelo-erythropoiesis by compromising the adipocyte bone marrow niche.* Nat Cell Biol, 2017. **19**(11): p. 1336-1347.
42. Mass, E., et al., *A somatic mutation in erythro-myeloid progenitors causes neurodegenerative disease.* Nature, 2017. **549**(7672): p. 389-393.

7

THESIS_Roberto_Avellino.indd 170 07-05-18 10:18

SUMMARY OF THE THESIS
AND GENERAL DISCUSSION

THESIS_Roberto_Avellino.indd 171 07-05-18 10:18

THESIS_Roberto_Avellino.indd 172 07-05-18 10:18

SUMMARY OF THE THESIS

In this thesis the transcriptional control of the *CEBPA* gene is studied in human and in murine hematopoiesis. This thesis focuses on the discovery of enhancers that are responsible for *CEBPA* transcriptional control in the bone marrow in health and disease.

The **second chapter** of this thesis comprises a review that discusses novel concepts about the role of *CEBPA* in normal hematopoiesis and in disease. It focuses mainly on the many layers of transcriptional regulation of *CEBPA,* with particular interest on the function of the +42kb *CEBPA*-enhancer in hematopoiesis. Hypothetical insights were raised to address the importance of the enhancer to maintain *CEBPA* levels at a steady-state balance to protect the bone marrow from undergoing HSC consumption, exhaustion and later bone marrow failure. Moreover, the +42kb enhancer can potentially be targeted by oncogenic mechanisms to minimize *CEBPA* levels down to a threshold that it's insufficient to drive the myeloid differentiation program and predispose myeloid committed progenitors to acute myeloid leukemia. In line with the importance of *CEBPA* levels in AML, the emerging role of *CEBPA* in the development of AML has been given particular attention since such findings confirm that *CEBPA* expression levels require transcriptional control in bone marrow progenitors throughout the lifetime of an organism to protect from onset of the disease.

The **third chapter** describes thoroughly the *CEBPA* locus in humans and mice to identify potential enhancers that are responsible for *CEBPA* transcriptional control in myelopoiesis. The use of multi-technical approaches including active histone modifications, chromosome conformation capture and genome editing technology (CRISPR/Cas9 approach) revealed that a primary responsive enhancer in the *Cebpa* locus, located at +42kb and +37kb in humans and mice, respectively, activates *Cebpa* and the neutrophilic gene expression program. This enhancer engages with *CEBPA* in myeloid cells only, suggesting tissue specificity. Germ-line knockout of the enhancer in murine models compromises *Cebpa* expression in the bone marrow only and halts the neutrophilic differentiation program causing neutropenia. In addition, germ line enhancer knockout influences the HSC population, which leaves it an open end for further investigations. In conclusion, this study characterized an important *CEBPA*-enhancer that has crucial implications in neutrophilic development.

The **fourth** and the **fifth chapter** focus on the role of the non-coding region of the *CEBPA* locus in AML. A substantial number of AML patients exhibit relative low *CEBPA* expression levels compared to other AMLs. Within these subsets are AML patient subsets with known hematological abnormalities such as the fusion oncoprotein AML1-ETO, ectopic high EVI1, FLT3-ITD, DNMT3, and a subset of AMLs with no recurrent abnormalities. For the latter group, the link to low *CEBPA* expression levels was hypothesized to occur via mutations in the non-coding genome of the *CEBPA* locus. However, screening 200 AML patients including controls, did not find any recurrent point mutations or chromosomal abnormalities,

THESIS_Roberto_Avellino.indd 173 07-05-18 10:18

except for one patient that harbored a bi-allelic deletion of *CEBPA*. These findings lead to conclusions that the *CEBPA* locus in AML is not targeted by mutations but potentially deregulated at the chromatin level. The mechanism by which AML1-ETO, one of the most common recurrent abnormality in AML, causes low *CEBPA* expression levels is investigated in Chapter 5. We demonstrate that AML1-ETO binds to the *CEBPA* locus via the +42kb enhancer, reverses its active chromatin state, disengage enhancer-gene interactions, down-regulate *CEBPA* expression and predispose bone marrow progenitors for the onset of acute myeloid leukemia.

In **Chapter six**, we partly tackled the unanswered question from **Chapter two** about the drop in HSC numbers in *CEBPA*-enhancer-deleted and neutropenic mice. The main question addressed is two-way; (1) whether deleting the enhancer leads to the loss of the HSC sub-population that express *Cebpa* and shuts down the neutrophil differentiation program at the HSC stage, thus explained as a cell-autonomous effect; or (2) the neutropenia negatively influences the HSCs as an extrinsic effect in the absence of *Cebpa*. Using single-cell RNA-seq in combination with bulk RNA-seq in hematopoietic stem progenitor cells, the most primitive HSC population does not express *Cebpa*, thus excluding that enhancer deletion influence HSCs in a cell autonomous way. The second hypothesis was tested using a transplantation model. Transplanted enhancer-deleted (neutropenic) bone marrow cells induced HSC loss and latent bone marrow failure in recipient mice, suggesting that neutropenic bone marrow cells compromise hematopoiesis by using, undefined, extrinsic mechanisms.

In conclusion, this thesis reveals new mechanisms of how a *CEBPA*-enhancer acts as a primary switch to control neutrophilic differentiation and maintain bone marrow integrity. This enhancer is also a target for oncogenic deregulation of *CEBPA* expression in human AML, thus studying its function serves as a paradigm for epigenetic therapeutic targeting in the field of leukemia.

GENERAL DISCUSSION

8.1 Introduction

High throughput genome-wide sequencing technologies have revolutionized our knowledge of the structure and function of the genome [1, 2]. To-date we know that the non-coding genome, which was previously considered as non-relevant "junk" DNA[3, 4], regulates coding genes and comprises conserved DNA sequences termed enhancers that bind transcription factors and other nuclear proteins for transcription regulation[5-7].

Joined efforts from different consortia, such as encyclopedia of DNA elements ENCODE [5] and functional annotation of mouse (FANTOM) [8-10], are constantly providing and updating publicly available datasets to study and characterize chromatin states of enhancers in different tissues and species. These datasets offer a detailed representation of the regulatory genome based on histone modifications and binding of co-regulators involved in transcriptional control. In the last decade, high throughput sequencing was combined with other technologies to study chromatin composition, structure and function[11-14]. For instance, chromatin immunoprecipitation combined with high throughput sequencing (ChIP-seq) is majorly used to identify histone modifications and binding of transcription factors (or other chromatin-related proteins) genome wide to define the current state of the enhancer(s) under investigation. Additional technologies that contribute to chromatin profiling include DNase-seq and ATAC-seq to study chromatin accessibility[15, 16]. Moreover, chromatin conformation capture (3C) based technologies reveal chromatin interactions between regulatory elements and corresponding gene promoters, which is a substantial hallmark of active enhancers[17-19]. *In vitro* and *in vivo* models are needed conduct functional experiments and determine the role of these elements normal physiology and disease. Conventional methods used to generate these models are labor intensive and time consuming. However, the advent of genome editing technologies such as transcription activator-like effector nucleases (TALENs) and clustered repeats of interspersed elements (CRISPR) increased the efficiency of DNA manipulation in primary tissues as well as in cell lines, and decreased the turnaround time to generate genetically manipulated *in vivo* models[20-23]. The main concept of this thesis is to reveal transcriptional mechanisms that regulate expression of *CEBPA* in neutrophilic differentiation and disease. We used the above approaches to characterize enhancers in the *CEBPA* locus to unravel their function using human cell lines and in vivo mouse models. We found that a myeloid specific enhancer located 42kb downstream of *CEBPA* is conserved between humans and mice (+37kb in mice) and acts as a primary switch to activate *Cebpa* expression during development and in adult life. Based on recent data in our models, we revised our hypothesis about *CEBPA*-*CEBPG* regulation and provided new insights stating that *CEBPA* and *CEBPG* are regulated independently from each other and that both genes are engaged in two separate chromatin loops. Then we questioned the role of the *CEBPA* locus in transcriptional deregulation

in AML patients. We studied mechanisms of how *CEBPA* is deregulated in acute myeloid leukemia by oncoprotein enhancer hijacking and screened the *CEBPA* locus in AML patients for point mutations or chromosomal abnormalities that could be cryptic and undetected by conventional karyotype analysis. In Chapter 6 we further studied our finding showing a correlation between severe neutropenia in *Cebpa*-enhancer knockout models and a drastic loss of long-term hematopoietic stem cells (LT-HSCs). Using transplantation models we observed that neutropenic bone marrow progenitor cells cause a stoichiometric imbalance in the bone marrow and induces a latent and extrinsic negative influence on the bone marrow of the host leading to bone marrow failure. In conclusion, this thesis reveals transcriptional mechanisms to drive neutrophilic differentiation, how these mechanisms are deregulated in AML and how LT-HSC integrity depends on the stoichiometry of lineage differentiation in the bone marrow.

8.2 CEBPA regulation in hematopoiesis: The Enhancer-rich CEBPA locus

8.2.1 Regulation inside the *CEBPA* TAD

The human *CEBPA* locus on chromosome 19 encompasses a 220kb topological associated domain, containing *CEBPA* and *CEBPG* genes that belong to the *CEBP* family of leucine zipper transcription factors (**Chapter 3** Fig.1a). *CEBPA* is located 70kb upstream from *CEBPG*, and the intergenic DNA in between the two genes constitutes many conserved DNA sequences between mouse and humans, one of which is marked by the histone acetylation mark H3K27ac, particularly in myeloid cells (**Chapter 3** Fig.2). Downstream of *CEBPA* there is a stretch of more than 50kb of conserved potential regulatory elements marked by active histone acetylation, termed here as the myeloid regulatory hub (MRH) cluster of enhancers (**Chapter 7** Fig.1). The MRH cluster is active only in *CEBPA* expressing cell lines and tissues (**Chapter 3** Fig.2) which exhibit a particular pattern of MRH cluster activity, depending on the tissue and cell type, that varies in (1) the number of enhancers active, and (2) which of the 8 enhancers are active. From 12 different tissues investigated, only the neutrophils and monocytes have all 8 enhancers active, including the myeloid-specific enhancers located at +35kb and +42kb downstream of *CEBPA*.

The architectural protein CTCF insulates the *CEBPA* TAD borders and also compart-mentalizes the *CEBPA* 3' site that form the *CEBPA* extrusion loop, from the *CEBPG* 5' site, i.e. the *CEBPG* extrusion loop. Loop extrusions are formed when two CTCF-bound DNA regions (motifs) are in a converging position and ready to move towards each other while extruding out a chromatin loop, which engages promoters, their genes and their corresponding enhancers[24].

CTCF binding at the *CEBPA* TSS and at 1.5kb of the *CEBPA* promoter are most probably the sites that generate the two independent extrusions loops for *CEBPA* and *CEBPG*, suggesting that they are regulated independently. The *CEBPA* extrusion loop contains the *CEBPA* gene and the MRH cluster of enhancers. In this section, we focus on the *CEBPA* extrusion loop and discuss the role of the MRH cluster of enhancers on the transcriptional control of *CEBPA* in hematopoiesis.

8.2.2 The +42kb enhancer: the primary switch of *CEBPA* expression and the MRH cluster

In **chapter 3** we thoroughly studied the function of the +42kb enhancer in mice and humans. The active histone-mark H3K27ac in HSPCs and the engagement of the +42kb enhancer with *CEBPA* in myeloid cells prompted us to study its role in hematopoiesis. We generated a knockout mouse model and found that the +42kb enhancer (+37kb enhancer in mice) acts as the primary responsive genetic element to initiate *Cebpa* expression in hematopoiesis, possibly already in fetal life. This particularly comes from a previous studies using *LacZ* reporter transgenic expression driven by this enhancer in mice, showing LacZ-activity in the dorsal aorta and in the fetal liver [25]. The +42kb enhancer acts autonomously, as previously stated, since germ line knockout of the enhancer shuts down the neutrophilic transcriptional program in early LSK progenitors by down-regulating *Cebpa* expression (**Chapter 3** and **6**). This concludes that the +42kb enhancer works independently from the rest of the MRH cluster of enhancers to initiate and regulate *CEBPA* expression, hence it functions as the primary genetic element for neutrophil lineage priming (**Chapter 7** Fig.1).

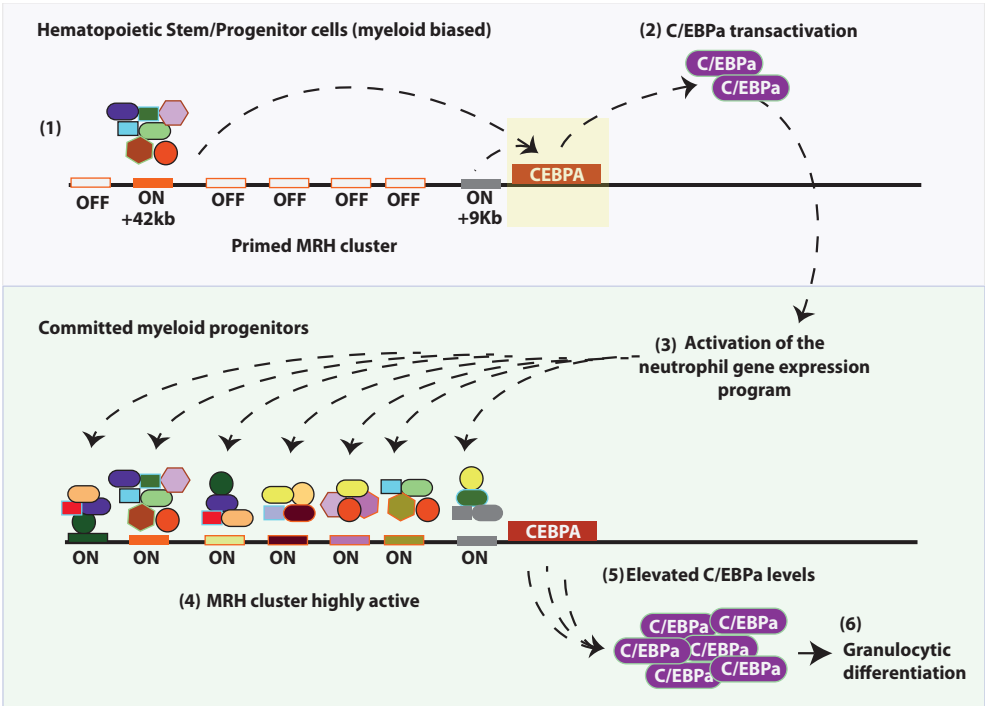


Figure1. A schematic overview representing the hierarchical transcriptional activation of *CEBPA* in neutrophil differentiation. (1) Upstream transcription factors bind the accessible +42kb enhancer that acts as the primary responsive element to initiate *CEBPA* expression. (2) Transcriptional and translational expression of C/EBPα activates the gene expression program of neutrophilic differentiation. (3) The neutrophilic differentiation program activates the expression of neutrophilic associated factors that in turn (4) activates the whole MRH (myeloid regulatory hub) cluster of enhancers. (5) Nuclear C/EBPα dosage reaches threshold to (6) induce granulocytic differentiation.

8.2.3 3' *CEBPA* Intergenic region: The MRH cluster and its role in early hematopoiesis and adult myelopoiesis

Revealing the function of the MRH cluster of enhancers in hematopoiesis is our next milestone. The independent and autonomous function of the +42kb enhancer does not exclude or underestimate the importance of other enhancers within the MRH cluster during myeloid differentiation. We hypothesize that the other enhancers within the MRH cluster have (1) an additive and/or redundant function during neutrophilic differentiation, (2) an important role in other lineages of the myeloid fate in the bone marrow and (3) a role in the development of other *CEBPA* dependent organs.

Enhancers occur in clusters to fine-tune the transcriptional output of each and every enhancer in any given locus and increase the expression levels of the corresponding gene(s). Examples of genes exhibiting clusters of enhancers with additive function, termed additive enhancers, include the Indian Hedgehog (*Ihh*) gene locus during development[26]. The additive function of the 9 enhancers located upstream of the *Ihh* gene was demonstrated using a *Sleeping Beauty* cassette, with a transgenic LacZ gene inserted in the intron of

Ihh. The increase in numbers of active enhancers within the MRH cluster in the HSPC CD34+ cells to mature neutrophils correlates with the increase in *Cebpa* expression levels (Chapter 2 Fig.3a). We hypothesize that the number of active enhancers will increase with differentiation and that the enhancers at the MRH cluster constitute a combined additive function to increase *Cebpa* output. In addition, multiple active enhancers in a given locus might act as shadow enhancers to exert redundancy and protect from potential disease-prone mutations at enhancer sequences [27-29]. Whether the MRH enhancers have redundant functions, requires further investigations that are discussed in the next section.

In addition to neutrophilic differentiation, any of these enhancers might also confer cell type specificity within the myeloid lineage. Recently it has been reported that an enhancer located at +39kb regulates *CEBPA* expression for basophilic differentiation in common myeloid progenitor cell populations [30]. Another enhancer located at +34kb of *CEBPA* is active in mature neutrophils and monocytes, but not in any other tissue. This implies that the enhancers within the MRH cluster have cell type specificity also between myeloid cells in the bone marrow. The function of the other enhancers located at +9kb, +15kb, +21kb, +29kb, +34kb, +42kb, +50kb and +55kb still needs to be further investigated. Interestingly, the +9kb enhancer is active in murine LSK populations and in human $CD34^+$ HSPCs (**Chapter 2** Fig.2). This element draws a special interest for further investigations since its enrichment for H3K27ac histone mark always correlates with *CEBPA* expression, independent of cell type and tissue specificity. We believe that the +9kb enhancer is active in early fetal life since *CEBPA* expressing tissues are derived from all the three germ line layers, mesoderm, ectoderm and endoderm. In line with this, the +9kb enhancer contains motif clusters for TFs that are both cell lineage and tissue specific, as well as for TFs that are generic or ubiquitously expressed. . Based on previous findings and current hypothetical presumptions, the *CEBPA* locus has a major role in defining the epigenetic landscape of the myeloid fate, thus, a clear experimental setup is required to further investigate and reveal the function of these enhancers in hematopoiesis.

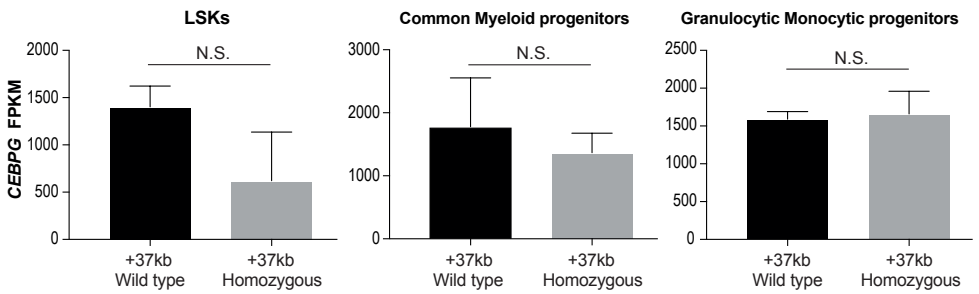


Figure2. *Cebpg* expression levels expressed in RPKM normalized values from sorted bone marrows of murine wild type and +37kb enhancer deleted LSKs, CMPs and GMPs. Non-parametric Student T-test was used to test for statistical significance.

THESIS_Roberto_Avellino.indd 179 07-05-18 10:18

8.2.4 Experimental approach: Characterizing the MRH cluster in Hematopoiesis

All the enhancer sequences mentioned above are of human origin and are all conserved in the mouse genome. Given that low to no *Cebpa* expression levels acquire clonogenic potential with infinite re-plating of murine total bone marrow cells, we plan to set up an enhancer-knockout screening method using CRISPR/Cas9 genome editing and infinite re-plating as a readout of *Cebpa* deregulation in murine bone marrow progenitors. We plan to start with single enhancer deletions and characterize immune-phenotypically the cells that undergo clonogenic potential and also investigate their chromatin status using H3K27ac and DNA/chromatin accessibility using DNase-seq or ATAC-seq.

Enhancers that are selected from the screening method will be interrogated *in vivo* in mouse models. There are several ways and approaches to investigate their function *in vivo*. A reporter model to reveal their role as regulators of myelopoieisis during fetal development or in adult bone marrows can be generated by introducing either a LacZ or a fluorescent protein cassette 3' of *Cebpa* using a sophisticated system such as a *Sleeping Beauty* transposon cassette that minimally interferes with surrounding gene regulation[31]. By generating this model, deletion of single or combinations of enhancers allows for testing whether these regulatory elements have an additive or redundant function. It will also allow us to study whether different individual enhancers have distinct roles in regulating *Cebpa* expression in different cell types of the myeloid lineage. Another possibility is to delete the enhancer in zygotes to generate a germ-line deletion model. However, this might impose problems on the viability of the mice since different enhancers in the MRH cluster might also regulate *Cebpa* in other tissues[32]. For instance, a germ line knockout of the +9kb potential enhancer (active in all *CEBPA* expressing tissues), hypothetically, should give a similar phenotype as the germ-line *Cebpa* gene knockout. Thus, to investigate its function, a conditional enhancer knockout model is preferred over a germ-line deletion model.

These approaches altogether will unveil the function of the MRH cluster of enhancers in *CEBPA* expressing myeloid cell types. These studies may provide a paradigm to understand how other similarly-structured gene loci that have strong implications in development, differentiation and disease, are regulated in the genome.

8.3 Transcriptional control of CEBPA and CEBPG

8.3.1 *Mechanistic insights in CEBPG regulation:* Transcriptional control in trans?

The CEBP family members, including the single exon genes *CEBPA*, *CEBPB*, *CEBPD* and *CEBPE* are mostly known for their differentiation mediated potential in the bone marrow. Less is known about *CEBPG*, which is located 70kb upstream of *CEBPA*. In the discussion section of **Chapter 3,** we discuss that, unlike previously reported in Alberich-Jorda et al, the absence of *Cebpa* expression by *Cebpa*-enhancer deletion, does not influence *Cebpg* expression levels. Alberich-Jorda et al show that *CEBPA* and *CEBPG* in a human AML subset exhibit an inverse

correlation in expression ((See Chapter 3). DNA methylation data in combination with gene expression profiling shows that in a subset of AML patients, the *CEBPA* promoter exhibits DNA methylation with low/no *CEBPA* expression levels, whereas the *CEBPG* promoter is free of methylation and exhibits strongly increased *CEBPG* levels when compared to other AMLs. Such findings led to the hypothesis that *CEBPA* negatively influences the transcription of *CEBPG*, and that *CEBPA* silencing leads to high *CEBPG* expression levels. C/EBPα binds to the promoter of *Cebpg* in both mouse and human cell lines and negatively influences the E2F transcription factor activity, known for its transcriptional activating potential of *CEBPG*. Furthermore, *Cebpa* and *Cebpg* expression in mouse bone marrow cells follow an inverse pattern of expression, indicating that both genes are inversely regulated. In the *Cebpa* knockout mouse model, the lineage^{-ve} c-kit^{+ve} Sca-1^{+ve} fraction, the population of bone marrow cells where stem and progenitor cells reside, *Cebpg* is significantly increased. Reintroduction of a *Cebpa* cDNA into the *Cebpa* knock out LSK cells resulted in partial repression of *Cebpg* mRNA levels. In addition, *Cebpg* seems to have negative implications on emergency granulopoiesis, which is overtaken by CEBPb, since knockdown of *CEBPG* shows a release from the differentiation block.

8.3.2 CTCF may play a major role in *CEBPA* and *CEBPG* gene expression

From the study by Alberich-Jorda et al., it was concluded that the CEBPα protein is responsible for the down modulation of *Cebpg* by binding to the promoter and repressing its expression *Cebpg* transcription in *trans*. However, our studies with *in vivo* +37kb enhancer deleted mouse model and the +42kb enhancer (human homolog of the murine +37kb enhancer) deleted THP-1 cells suggest a different mechanism. We observed that loss or decrease of *Cebpa* does not affect *Cebpg* expression. (Chapter 2). In the murine model, upon enhancer deletion, *Cebpa* levels by RNA-seq were significantly decreased in different immunophenotypically-sorted bone marrow populations. These populations include the lineage^{-ve} c-kit^{+ve} Sca-1^{+ve} (LSK) hematopoietic stem and progenitor population the common myeloid progenitor (CMP) and the granulocytic-monocytic progenitor (GMP) cell populations. However, despite the loss of *Cebpa* expression, we did not see any effects on the expression of *Cebpg* in LSKs, CMPs or GMPs, which contradicts the findings observed in the *Cebpa* knockout mouse model (**Chapter 2** Fig.7). In addition, knockdown of *Cebpa* in lineage negative bone marrow cells and deletion of the 42kb enhancer in THP-1 cells did not show any changes in *Cebpg/CEBPG* expression, respectively.

Generating gene knockouts in mouse models poses challenges, given that deletion of critical DNA sequences (in addition to the gene under investigation) that belong to regulatory elements or DNA motifs recognized by nuclear proteins might have detrimental consequences on the transcriptional control of nearby genes. The TAD domain that contains *Cebpa* and *Cebpg* has defined insulated boarders, bound by the TF CTCF

(**Chapter 3** Fig.1). Within the TAD itself, there are conserved sites enriched for CTCF binding, which are close to the *Cebpa* gene. These include the *Cebpa* TSS, *Cebpa* -1.8kb and *Cebpa* -9kb. Excision of the *Cebpa* gene as reported in the study by Zhang et al[32], caused deletion of additional nearby sequences upstream of *Cebpa*, including the DNA binding sites for CTCF and perhaps for other potential unknown factors responsible for the chromatin folding and looping. This raises questions about the underlying mechanism that increase *Cebpg* expression levels upon the excision of *Cebpa*. We hypothesize that the *CEBPA* TAD is sub-divided into two loops: 5' *CEBPG* extrusion loop contains the DNA sequences located upstream of the *Cebpa* promoter to *Cebpg*, and a *CEBPA* loop extrusion containing DNA sequences from the *Cebpa* promoter to the promoter of *Slc7a10* gene including MRH cluster of enhancers located 3' of *Cebpa*. Studies have shown that deletion of CTCF sites within a TAD, but not at borders, has detrimental effects on the expression of genes within the TAD. We hypothesize that upon *Cebpa* excision, deleted CTCF sequences located upstream of *Cebpa* form a new single loop that brings at close proximity *Cebpg* with the MRH cluster, leading to a steady increase in *Cebpg* expression levels (**Chapter 7** Fig.3).

A subset of human AML patients showed elevated levels of *CEBPG*, associated with DNA hyper-methylation at the *CEBPA* gene, its promoter, and upstream sequences including the CTCF binding sites. CTCF was the first DNA binding factor to be recognized as DNA methylation sensitive i.e. it doesn't bind to methylated DNA[33]. Abnormal DNA methylation patterns caused by mutations in the de-methylation pathway, such as in genes encoding enzymes IDH1 and IDH2, alter gene expression programs via loss of CTCF binding and decreased insulation of neighboring topological associated domains (TADs) [34]. Such a mechanism might also explain the increase in *CEBPG* levels in the $CEBPA^{silenced}$ AML subgroup. Our unpublished data of $CEBPA^{silenced}$ patients and other AMLs show that by applying enhanced reduced representation bisulfite sequencing (eRRBS), the *CEBPA* gene and its promoter are hyper-methylated[35]. whereas the +42kb enhancer is hypo-methylated in all AMLs (**Chapter 7** Fig.3). Putting these findings into perspective, (1) DNA methylation occurring at the *CEBPA* promoter and its TSS inhibits the CTCF binding to DNA, (2) loss of CTCF binding reduces border strength and insulation between *CEBPG* and *CEBPA* extrusions loops, (3) decreased insulation leads to exposure of *CEBPG* to the MRH cluster of enhancers, (4) *CEBPG* transcription becomes under the control of the +42kb enhancer.

CEBPG deregulation in AML requires a deeper understanding based on further experimental investigations. To mimic the DNA methylation effect of the CTCF binding sites located upstream of *CEBPA*, fusing the methyl transferase domain of DNMT enzymes to nuclease-deficient Cas9 (dCas9) and target the fusion protein by gRNAs to the CTCF binding sites upstream of *CEBPA*, would methylate cytosines in the CTCF motifs and prevents its DNA binding. This approach would mimic, at least in part, the observed DNA methylation findings in the $CEBPA^{silenced}$ AML subset of patients. Alternatively, the CTCF motifs can be deleted or mutated using gRNAs flanking the CTCF binding sites to cut and replace the

sequence with a mutant-corrected donor template. This experimental approach could be conducted in human myeloid cell lines; one expressing *CEBPG* only (K562) and another one expressing *CEBPG* and *CEBPA* (THP-1, U937, HL-60). Next, to investigate whether by altering the CTCF binding site reduces the insulation between the *CEBPA* and the *CEBPG* loop, and perhaps increases the interaction between the MRH cluster of enhancers with *CEBPG*, applying 4C-seq and capture HI-C will (1) capture novel potential interactions or (2) detect an increase in the interaction frequency between enhancers and promoters that ultimately becomes one whole loop.

8.3.3 Yet so close but distinctively regulated: Different regulatory mechanisms within the same TAD

Based on the outcome of these experiments, it may be concluded that *CEBPA* and *CEBPG* are regulated separately and independently. A possible reason to explain this assumption is because *CEBPA* and *CEBPG* encode TFs that have a different function in different biological pathways. *CEBPG* is expressed in all tissues thus, acting more like a housekeeping gene that is constitutively activated (Unpublished observations). On the other hand, *CEBPA* is expressed in specific tissues and at specific stages of differentiation. *CEBPA* requires regulatory mechanisms that modulate its expression in a spatiotemporal manner. This means that the two genes within the same TAD require a different transcriptional regulatory mechanism, hence the formation of two independent and separate extrusion loops (**Chapter 7** Fig.3). Ubiquitous genes or housekeeping genes, in general, do not engage with enhancers for their transcriptional control[36]. Instead, their promoters contain TF binding motifs that can be recognized, bound and activated in any cell type and the activation is strong enough to drive enough to drive constitutive expression of the corresponding house keeping gene{36].

In conclusion, Jorda et al studied the regulatory mechanisms occurring in trans between *CEBPA* and *CEBPG* within the *CEBPA* TAD. In this section we hereby discussed other potential mechanisms that go beyond the *CEBPA-CEBPG* regulation in trans by protein to DNA interaction. We propose that regulation between the two genes occur at a higher ordinary structure of the genome. We hypothesize that loop anchors set by CTCF at borders and at the *CEBPA* promoter, divides the *CEBPA* TAD into two separate loop extrusions, and allow *CEBPA* to be transcriptionally regulated by its 3' enhancers. Whereas in AML, CTCF anchors are disturbed and the *CEBPA* and *CEBPG* extrusion loops become one single loop and deregulate *CEBPG* levels, which might contribute to AML.

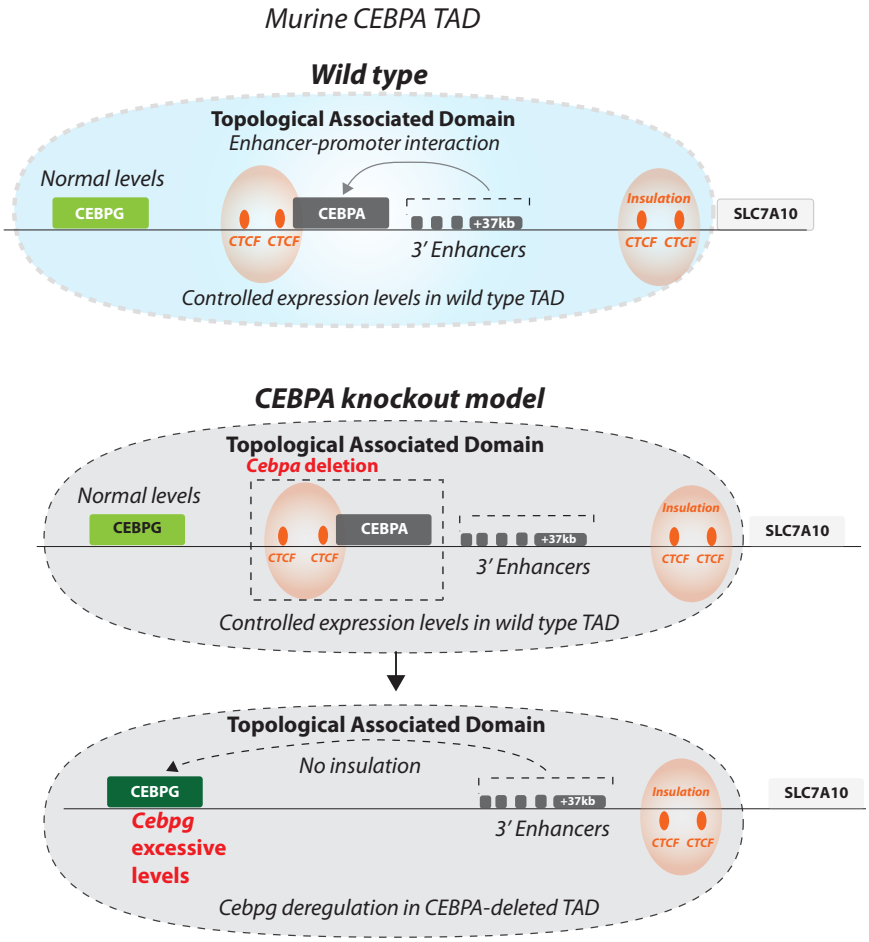
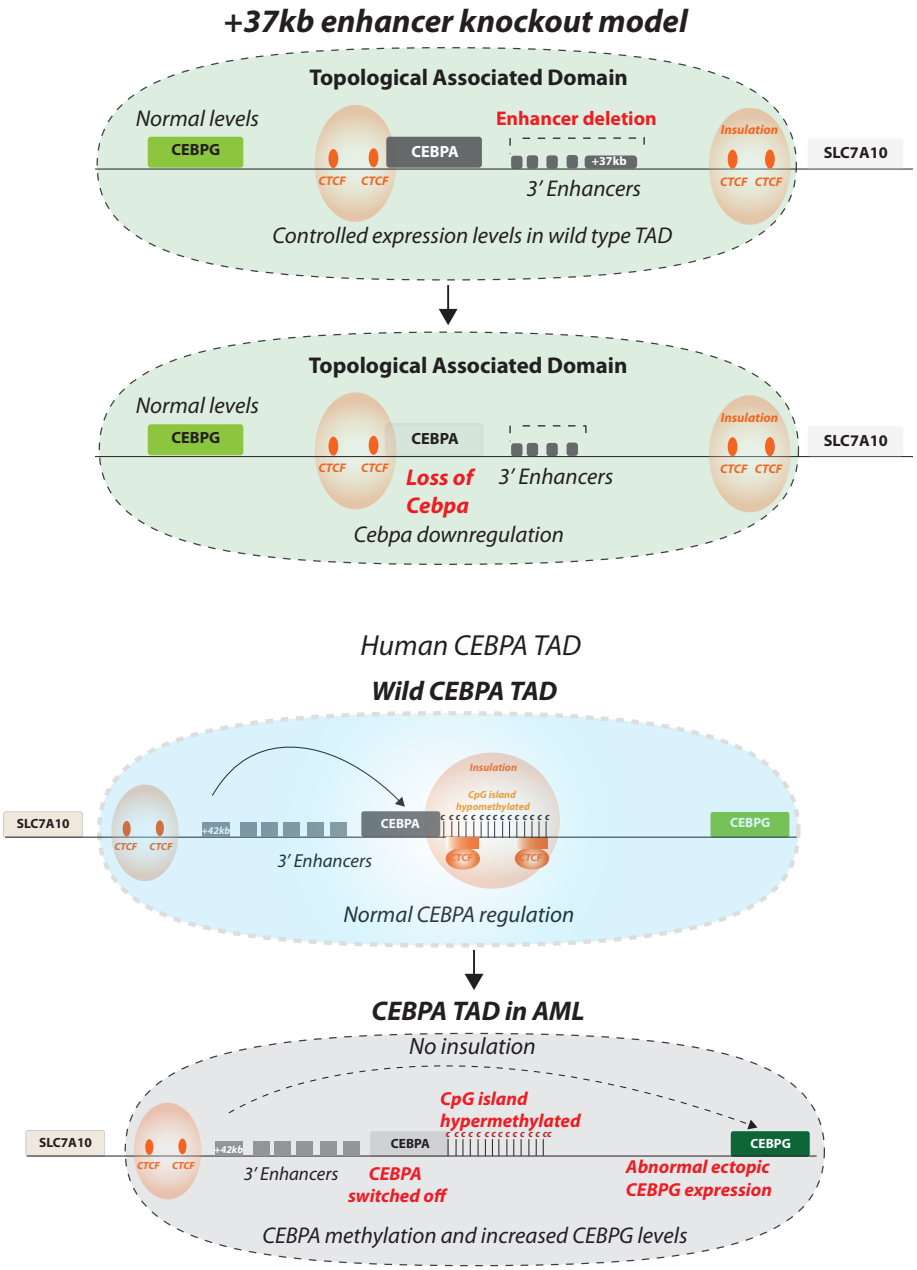


Figure3. A schematic overview representing the topological associated domain of *CEBPA* and *CEBPG* in mouse and humans. (a) Wild type condition (light blue) showing that *Cebpa* and its 3' enhancers engage in interaction independent from any interaction with *Cebpg*. *Cebpa* knockout model (grey) showing the deletion of *Cebpa* gene together with CTCF motifs creates a new interaction between *Cebpg* and the enhancers, leading to increased *Cebpg* expression levels. In the +37kb enhancer model, the +37kb enhancer is deleted, *Cebpa* expression is lost, but not *Cebpg*. (b) Human *CEBPA* and its 3' enhancers engage in interaction independent from any interaction with *CEBPG*. DNA methylation at *CEBPA* promoter and at CTCF sites, disturb the insulation between *CEBPA* and *CEBPG* leading to ectopic interaction of *CEBPA*-enhancers with *CEBPG* resulting in an increase in *CEBPG* expression levels.


+37kb enhancer knockout model
Topological Associated Domain
Normal levels
CEBPG
CTCF CTCF
CEBPA
Enhancer deletion
+37kb
3' Enhancers
Insulation
CTCF CTCF
SLC7A10
Controlled expression levels in wild type TAD
Topological Associated Domain
Normal levels
CEBPG
CTCF CTCF
CEBPA
Loss of Cebpa
3' Enhancers
Insulation
CTCF CTCF
SLC7A10
Cebpa downregulation
Human CEBPA TAD
Wild CEBPA TAD
SLC7A10
CTCF CTCF
+42kb
CEBPA
3' Enhancers
Insulation
CpG island hypomethylated
CTCF CTCF
CEBPG
Normal CEBPA regulation
CEBPA TAD in AML
No insulation
SLC7A10
CTCF CTCF
+42kb
CEBPA
3' Enhancers
CEBPA switched off
CpG island hypermethylated
CEBPG
Abnormal ectopic CEBPG expression
CEBPA methylation and increased CEBPG levels

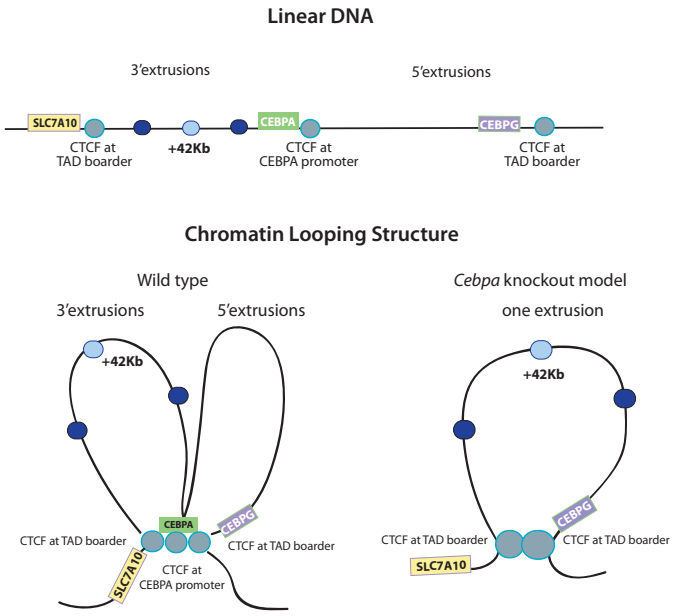


Figure4. A schematic representing the *Cebpa* (3') and *Cebpg* (5') extrusion loops based on CTCF binding at the *Cebpa* promoter. The two extrusions loops show that both genes are regulated independent from each other. In the *Cebpa* knockout model, CTCF sites at the *Cebpa* promoter are disrupted which further disrupts the two extrusion loops and create a new extrusion loop that disturbs *Cebpg* expression.

8.4 *The role of the CEBPA TAD in disease*.

In **Chapter 5** and **Chapter 6** we studied the role of the *CEBPA* TAD in disease. Given the tumor suppressor function of *CEBPA*, we studied how oncogenic mechanisms in AML target the transcriptional regulatory network of *CEBPA* to halt the neutrophil differentiation program. We show that onco-proteins, such as AML1-ETO, are addicted to +42kb enhancer to downregulate *CEBPA* expression. Surprisingly, the +42kb enhancer and the rest of the *CEBPA* TAD is not a target for DNA mutations, except in one patient, which harbours a bi-allelic deletion of *CEBPA* in AML. The findings reported in this thesis will be discussed thoroughly in the next sections.

8.4.1 Hijacking tumor-supressor enhancers: predisposing bone marrow progenitors to leukemia

Oncoproteins use diverse mechanisms to deregulate transcription of target genes involved in cell cycle regulation, DNA repair and differentiation pathways. Perturbation in any of these pathways compromise normal hematopoiesis and triggers the onset of leukemia. Oncoproteins are commonly generated either by chromosomal abnormalities in coding sequences (translocations, inversions and deletions), by point mutations in coding sequences, or by mutations in regulatory elements that lead to their ectopic activation[37-43]. One of

the most studied mutations in AML is the chromosomal translocation t(8;21) which generates the AML1-ETO fusion oncoprotein accompanied by very low expression levels of *CEBPA*[44]. The oncogenic function of AML1-ETO in AML was studied extensively for the last two decades, however mechanisms of how this oncoprotein deregulate the neutrophilic gene expression program is still poorly understood. Attenuation of AML1-ETO in the Kasumi-1 cell line by siRNA activates a gene expression program driven by C/EBPα, indicating that *CEBPA* is an early target of AML1-ETO to halt the neutrophilic differentiation program[45, 46].

The oncogenic domains of the AML1-ETO protein were previously studied in AML [47-49]. The N-terminal part of AML1-ETO binds DNA motif sequences via AML1 (also known as RUNX1), whereas the C-terminal part of the oncoprotein contains the 4-nervy homologous (NRH1-4) domains of ETO and recruits different transcriptional co-regulators including transcriptional repressors[50-53]. The scope of our study was to reveal how AML1-ETO down-regulates *CEBPA* expression. Transcriptional regulation involves the combination of different transcriptional agents that act together to activate or repress genes via a series of events. Combined studies suggest that, for transcription to occur, a gene requires to be engaged into an active chromatin loop together with regulatory elements located in a topological associated domain and form complexes of DNA, RNA and protein [7, 54, 55].

Applying a set of experiments using AE-models as well as AE-patient samples we defined four unique events occurring in the *CEBPA*-TAD. Such events could be discriminated according to how they happen sequentially or independently upon AE expression: 1) loss of *CEBPA* expression; , 2) physical binding of the AML1-ETO oncoprotein in the *CEBPA* locus, predominantly to the +42kb enhancer; 3) histone deacetylation of H3K9 and H3K27 at the MRH enhancers; 4) and disengagement of chromatin interactions between *CEBPA* and its enhancers.

We asked whether these events are (1) AML1-ETO dependent and (2) whether they occur sequentially or simultaneously. We used inducible retroviral overexpression and induced lentiviral shRNA knockdown systems for AML1-ETO, in U937(*CEBPA*$^{+ve}$; AML1-ETO^{-ve}) and Kasumi-1 (*CEBPA*^{low}; AML1-ETO^{pos}) cell lines, respectively. Our hypothesis states that (1) binding of AML1-ETO to the enhancer recruits transcriptional repressors by the ETO sequentially leading to reversal of the active chromatin state of the locus, and disruption of the looping between *CEBPA* and its enhancers, ultimately leading to loss of *CEBPA* expression (**Chapter 7** Fig.5). The findings in our models overlapped with observations in the AML1-ETO patient specimens.

An additional approach to address the question whether binding of AML1-ETO to the *CEBPA* locus is the main driver of leukemic transformation, is to tether ETO or ETO-bound co-repressors to the *CEBPA* locus. Tethering a protein of interest to a specific locus allows the investigation for rapid and early transcriptional responses. The tethering methodologies developed so far include TALENs and CRISPR/Cas9 approaches[23, 56-58]. To reveal whether the binding to the enhancer and deacetylation are the two initial events that lead

to reduced *CEBPA* expression levels, we propose to use CRISPR/Cas9 and fuse the catalytic domain of a transcriptional repressor to dCas9 (dead Cas9 i.e. binds but doesn't cleave the DNA) and tether it to the +42kb enhancer using guide RNAs with sequences complementary to the enhancer[59]. Fusing the NHR domains of ETO to dCas9 would potentially mimic the AML1-ETO repressive function on genomic target sites. An additional approach is to fuse the catalytic domains of HDACs to dCas9, which are part of the AML1-ETO repressing complex. Histone de-acetylation by dCas9-HDACs showed adequate repression on different target genes [60]. Following the introduction of fused dCas9 proteins in the model of interest, *CEBPA* expression will be monitored followed by ChIP-seq experiments to test for tethering of dCas9 to the locus and the chromatin status of the enhancer. (**Chapter 7** Fig.5).

The next step will be to study whether the loss of the chromatin interaction between *CEBPA* and the enhancer is a consequence of enhancer deacetylation or whether it is an independent event. Methods that can be applied to detect changes in chromatin interactions include 4C-seq and capture-HI-C[19, 61]. Changes in chromatin interactions can explain whether deacetylated enhancers disrupt looping with *CEBPA*. Deacetylation of lysine residues deactivates functional domains of non-histone proteins[48]. Therefore, one other possibility is that AML1-ETO deacetylates and deactivate unknown potential architectural proteins bound to the enhancer that form loop complexes with the *CEBPA* gene. An important statement to highlight here is that 20% of AML1-ETO patients have mutations in genes that encode for the genome architectural protein complex of cohesion. Together with CTCF, cohesion forms ring structures and secure chromatin loops. Loss of function in cohesion is associated with loss of chromatin loop formation and reorganization of whole TADs[62]. Thus, such mutated genes may cooperate with the oncogenic function of AML1-ETO to deregulate downstream target genes.

8.4.2 Genomic instability causing TAD deletions and *CEBPA* null mutations

Bi-allelic mutations in the *CEBPA* coding sequence account for 10% of AML patients. Screening for *CEBPA* mutations has become a part of the diagnostic service since these mutations carry clinical significance and provide prognostic knowledge on whether to devise bone marrow transplantations for leukemic patients[63]. Given that the regulatory regions in the non-coding genome of the *CEBPA* locus might be a target for mutations as well, DNA sequencing was conducted in AML patients. It was hypothesized that point mutations in motif binding site of transcription factors within enhancer sequences or other mutations such as deletions may lead to low affinity or loss of TF binding and interfering with transcriptional activation. Attenuation of enhancer activity and transcriptional output would lead to the low *CEBPA* expression levels. Thus, a specific AML subgroup with relative low expression levels was selected for investigation.

Of the 200 AML cases sequenced, only one patient presented with a mutation in the whole TAD. Nonetheless, without excluding the importance of single nucleotide variants

(SNVs) in relationship to transcriptional regulation, enhancers harbored SNVs without any evidence of DNA mutations occurring in the *CEBPA* locus. This one patient harbored a bi-allelic deletion of *CEBPA* and *CEBPG*. One allele had a gross deletion of the long arm of chromosome 19 whereas the other allele contained a deletion of the whole *CEBPA* TAD, with deletion breakpoints occurring at or close to CTCF sites. Other regions in the genome exhibited deletions, indicating that gross genomic instability that highly predisposed the patient for these deletions. The patient's history showed that the AML was secondary related to therapy of a primary tumor that occurred in the small intestine.

It is of interest to note that this particular patient was previously recognized as an AML case with a unique mixed myeloid/lymphoid phenotype without *CEBPA* expression [35]. These cases were designated *CEBPA*-silenced leukemias. Almost all patients in this AML subgroup were reported to have full CpG methylation of the *CEBPA* gene, explaining the absence of *CEBPA* transcripts in the leukemia cells by epigenetic silencing. However, patient 7120 exhibited loss of both alleles at the DNA level and could not be investigated. Since patients with *CEBPA* silencing by methylation express a very comparable mixed myeloid/T-lymphoid gene expression signature, frequently carry mutations in NOTCH1, we argue that further analysis of bi-phenotypic leukemias should be screened for abnormalities in the *CEBPA* locus[35]. Furthermore, patients with chromosome 19q deletions should be included as well in such an analysis, as well as AML patients with genomic instability resulting in very complex karyotypes and frequent chromosomal deletions.

Another finding that was reported separately is that the *CEBPA*^{silenced} leukemias had a very strong CpG methylation signature, with many genes silenced by CpG methylation. A question that has remained unanswered is whether *CEBPA* is just one out of many genes being silenced by methylation, or whether silencing of *CEBPA* leads to the loss of *CEBPA* protein binding to its many target genes and whether this consequently leads to methylation and silencing of target genes. In the latter case, the patient with the deleted *CEBPA* gene (7120) should present exhibit a similar methylation signature as the other bi-phenotypic *CEBPA*-methylated cases.

8.5 *In vivo Cebpa-Enhancer-deletion and bone marrow failure*

The role of *Cebpa* and the +37kb enhancer in murine neutrophilic differentiation has been established in previous studies[32, 64-66]. Germ-line knockout of *Cebpa* impairs neutrophilic differentiation and also compromises organogenesis, majorly the liver, adipose system, lungs in feta life. Deletion of the +37kb enhancer Given its myeloid specificity, the +37kb enhancer deletion jeopardizes the development of neutrophils only (**Chapter 3**).

Deletion of the +37kb enhancer led to the following observations: (1) inactivation of the neutrophilic differentiation program occurs in HSPCs and causes severe neutropenia in bone marrow and peripheral blood; (2) Prolonged neutropenia negatively influence LT-HSCs and cause bone marrow failure. The experiments carried out to reach these conclusions were conducted in two separate mouse models; a germ-line deletion of the +37kb enhancer mouse model and in a transplantation model (enhancer deleted bone marrow cells transplanted in sub-lethal irradiated mice). Reduction in LT-HSC numbers raised questions whether it is caused by (1) lack of production of *Cebpa*-expressing LT-HSCs due to germ-line +37kb enhancer deletion, or (2) extrinsic events caused by neutropenia.

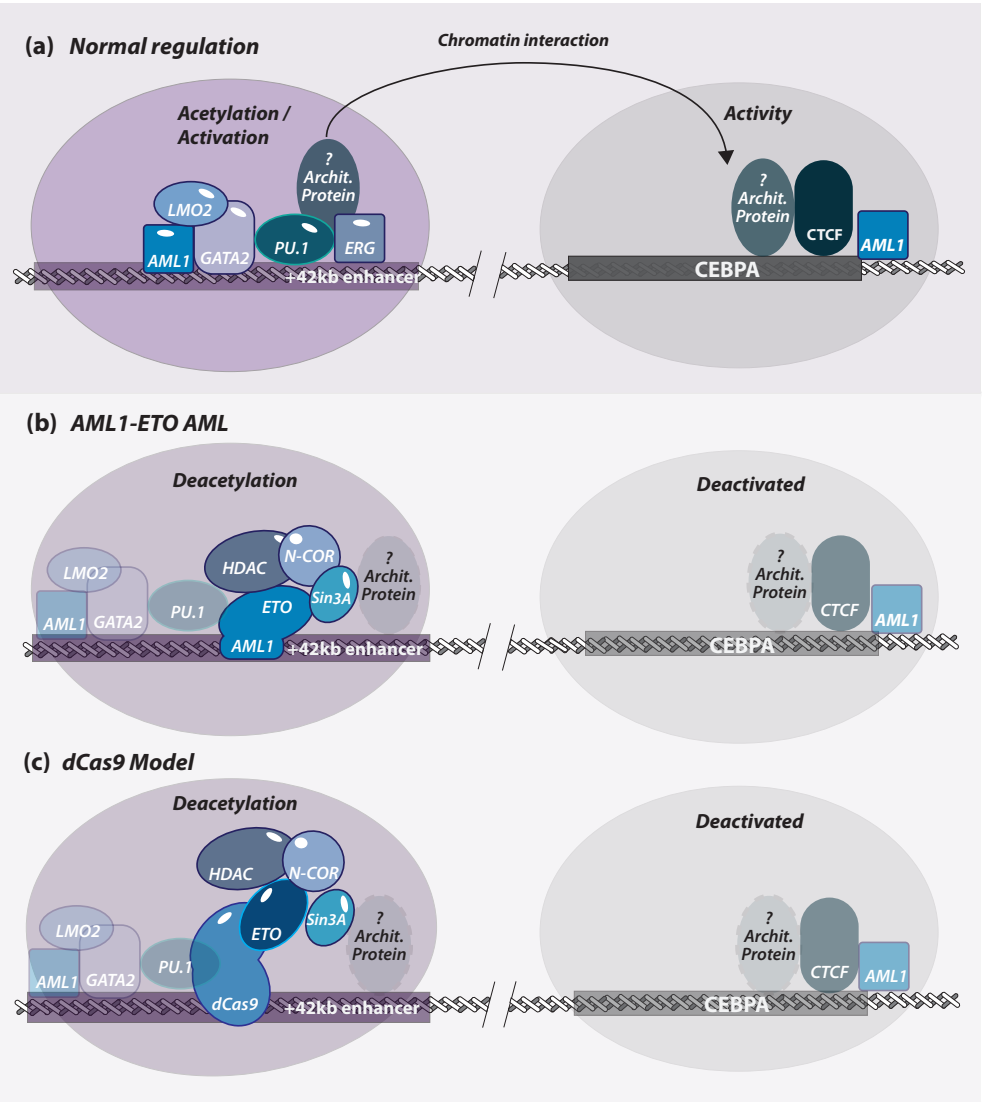


Figure5. A schematic representing (a) the binding of TFs known to transactivate *Cebpa* expression via architectural proteins that promote DNA looping together with CTCF. (b) The presence of AML1-ETO binding to the enhancer recruits co-repressor complexes and either displaces or deactivates the TF complex required to activate *CEBPA* expression. (c) Modeling of this mechanism by using dCas9-ETO tethering system.

8.5.1 LT-HSCs do not express *Cebpa* : Insights from single cell RNA-sequencing

To ask whether the +37kb enhancer-deletion is intrinsically related to the loss of LT-HSCs, we applied single cell RNA sequencing. From all single LT-HSCs investigated, we found that LT-HSCs do not express *Cebpa*, contradicting conclusions reported previously[67]. In this previous report [67], *Cebpa* was investigated from a heterogeneous pool of hematopoietic stem progenitor cells. Measuring mRNA levels from a bulk population can impose several problems caused by bone marrow heterogeneity. The application of single-cell RNA sequencing dissects cell populations based on their expression program and distinguishes between different populations that share a similar immune-phenotype. It also offers a plausible identification of transcriptional networks to characterize single cell-lineages at a higher resolution than immune-phenotyping[68, 69].

By dissecting bone marrow heterogeneity into wild type single cells we showed that the integrity of LT-HSCs is not intrinsically dependent on *Cebpa* activity. Losing *Cebpa* expression deactivates the neutrophilic differentiation program, which leads to an unidentified mechanism that compromises the integrity of the LT-HSC population. Like *Cebpa*, *Spi1* (PU.1) has also been considered to be a stem cell factor, however single-cell studies showed that LT-HSCs do not harbor PU.1 expression and any influence on LT-HSC number and function appears extrinsic but not cell autonomous[70,71].

8.5.2 Upregulation of genes in enhancer-deleted LSKs: Bias towards specific cell populations upon loss of the neutrophilic lineage?

The changes that we observe in *Tcf7*, *Bcl11b* and *Irf8* genes by conventional RNA-seq conducted on bulk +37kb^{HOM} LSKs, defines changes in cell population numbers. The increase in expression of each of these factors is most probably representing a particular cell lineage or lineages that becomes predominant in the bone marrow upon loss of the neutrophilic differentiation line. Loss of the neutrophilic lineage causes a stoichiometric imbalance in the myeloid fate, which alters the numbers of myeloid fate progenitors (increase in IRF8 expression levels)[72-74] and possibly, influencing also lymphoid fate progenitors (increase in TCF7 and BCL11b)[75, 76].

8.5.3 Aberrant extrinsic mechanisms influence normal hematopoiesis of the host: The role of the niche.

A two-way hypothesis is addressed in **Chapter 6** to explain the LT-HSC phenotype. We previously showed that enhancer deleted bone marrow cells are blocked in neutrophilic differentiation and have clonogenic capacity *in vitro* (**Chapter 3**). These pre-leukemic characteristics are reflected in the abnormal gene expression program of the LSKs mentioned in previous sections. Although at low chimerism, prolonged circulation of transplanted pre-leukemic bone marrow cells might explain the latent negative effect on the hematopoiesis of the host.

It is becoming more clear how pre-malignant and malignant cells influence the bone marrow to suppress normal hematopoiesis as well as the function of external non-hematopoietic organs [77-79]. The communication between hematopoietic stem and progenitor cells and the bone marrow microenvironment has been the focus of many scientific studies for many decades. It is now well established that HSCs are regulated by non-hematopoietic bone marrow stromal cells including osteoblasts, adipocytes, endothelial cells and mesenchymal stem cells [80-84]. These cellular compartments make up the environmental niche of bone and vascular origins. A current dogma states that an abnormal microenvironment induces genotoxic stress onto HSPCs leading to the onset of clonal hematological malignancies, which is the underlying mechanism for many bone marrow failure disorders. In fact, as a preliminary experiment, we studied the distribution of these distinct bone marrow stromal cell types in three mice transplanted with +37kb *CEBPA*-enhancer deleted marrow and three wild type controls, which did not develop a bone marrow failure yet (the +37kb$^{normal/intermediate}$ homozygous mice). Compared to controls, the +37kb homozygous mice showed abnormal absolute numbers of osteoblasts and endothelial cells (Supplementary Figure 6 in **Chapter 6**). A larger mouse cohort would suffice to investigate whether these findings are associated and complementary to the bone marrow failure phenotype.

Our findings suggest that bone marrow cell progenitors defective in neutrophilic differentiation induce hematopoietic deficiency on the host. Several examples of extrinsic mechanisms exerted on the bone marrow environment leading to differentiation defects have been reported. Leukemic cells secrete transcripts of miRNAs encapsulated in exosomes to silence critical stem cell factors such as c-MYB in the wild-type environment and compromise normal hematopoiesis [78, 85, 86]. Transplantation of leukemic HL-60 cells in a normal host suppress the normal hematopoietic environment providing a growth advantage to the leukemic HL-60 cells [78]. Immune suppression of the host contributes to immune evasion by the tumor cells, which allows them to grow further against minimized suppression. Such phenomenon is contained within the bone marrow in leukemogenesis, however, abnormal progenitors of the myeloid-erythroid lineage defective in BRAF signaling have been demonstrated to bypass the blood brain barrier and migrate into the central nervous system. There it disturbs neuronal function and contribute to neurodegenerative diseases[87]. All these examples underline the possibility of abnormal HSPCs and their progeny circulating in the bone marrow and other organs, which can negatively influence their niche and alter or suppress organ function.

THESIS_Roberto_Avellino.indd 193 07-05-18 10:18

REFERENCES

1. Baryawno, N. and D.T. Scadden, *Blood loses it when nerves go bad.* Cell Res, 2014. **24**(10): p. 1151-2.
2. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860 921.
3. Pennisi, E., *Genomics. ENCODE project writes eulogy for junk DNA.* Science, 2012. **337**(6099): p. 1159, 1161.
4. Ling, H., et al., *Junk DNA and the long non-coding RNA twist in cancer genetics.* Oncogene, 2015. **34**(39): p. 5003-11.
5. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.
6. Shapiro, J.A., *Genome system architecture and natural genetic engineering in evolution.* Ann N Y Acad Sci, 1999. **870**: p. 23-35.
7. Heinz, S., et al., *The selection and function of cell type-specific enhancers.* Nat Rev Mol Cell Biol, 2015. **16**(3): p. 144-54.
8. Kawaji, H., et al., *The FANTOM5 collection, a data series underpinning mammalian transcriptome atlases in diverse cell types.* Sci Data, 2017. **4**: p. 170113.
9. Lizio, M., et al., *Gateways to the FANTOM5 promoter level mammalian expression atlas.* Genome Biol, 2015. **16**: p. 22.
10. Maeda, N., et al., *Transcript annotation in FANTOM3: mouse gene catalog based on physical cDNAs.* PLoS Genet, 2006. **2**(4): p. e62.
11. Ansorge, W.J., *Next-generation DNA sequencing techniques.* N Biotechnol, 2009. **25**(4): p. 195-203.
12. Marguerat, S., B.T. Wilhelm, and J. Bahler, *Next-generation sequencing: applications beyond genomes.* Biochem Soc Trans, 2008. **36**(Pt 5): p. 1091-6.
13. Morozova, O. and M.A. Marra, *Applications of next-generation sequencing technologies in functional genomics.* Genomics, 2008. **92**(5): p. 255-64.
14. Wold, B. and R.M. Myers, *Sequence census methods for functional genomics.* Nat Methods, 2008. **5**(1): p. 19-21.
15. Buenrostro, J.D., et al., *ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide.* Curr Protoc Mol Biol, 2015. **109**: p. 21 29 1-9.
16. Zhong, J., et al., *Mapping nucleosome positions using DNase-seq.* Genome Res, 2016. **26**(3): p. 351-64.
17. Naumova, N., et al., *Analysis of long-range chromatin interactions using Chromosome Conformation Capture.* Methods, 2012. **58**(3): p. 192-203.
18. van de Werken, H.J., et al., *4C technology: protocols and data analysis.* Methods Enzymol, 2012. **513**: p. 89-112.
19. van de Werken, H.J., et al., *Robust 4C-seq data analysis to screen for regulatory DNA interactions.* Nat Methods, 2012. **9**(10): p. 969-72.
20. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.
21. Jiang, W., et al., *RNA-guided editing of bacterial genomes using CRISPR-Cas systems.* Nat Biotechnol, 2013. **31**(3): p. 233-9.
22. Wang, H., et al., *One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering.* Cell, 2013. **153**(4): p. 910-8.
23. Wei, C., et al., *TALEN or Cas9 - rapid, efficient and specific choices for genome modifications.* J Genet Genomics, 2013. **40**(6): p. 281-9.

24. Sanborn, A.L., et al., *Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes.* Proc Natl Acad Sci U S A, 2015. **112**(47): p. E6456-65.
25. Wilson, N.K., et al., *Integrated genome-scale analysis of the transcriptional regulatory landscape in a blood stem/progenitor cell model.* Blood, 2016. **127**(13): p. e12-23.
26. Will, A.J., et al., *Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog).* Nat Genet, 2017. **49**(10): p. 1539-1545.
27. Cannavo, E., et al., *Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks.* Curr Biol, 2016. **26**(1): p. 38-51.
28. Hong, J.W., D.A. Hendrix, and M.S. Levine, *Shadow enhancers as a source of evolutionary novelty.* Science, 2008. **321**(5894): p. 1314.
29. Swami, M., *Transcription: Shadow enhancers confer robustness.* Nat Rev Genet, 2010. **11**(7): p. 454.
30. Guo, M.H., et al., *Comprehensive population-based genome sequencing provides insight into hematopoietic regulatory mechanisms.* Proc Natl Acad Sci U S A, 2017. **114**(3): p. E327-E336.
31. Ruf, S., et al., *Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor.* Nat Genet, 2011. **43**(4): p. 379-86.
32. Zhang, D.E., et al., *Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice.* Proc Natl Acad Sci U S A, 1997. **94**(2): p. 569-74.
33. Hark, A.T., et al., *CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus.* Nature, 2000. **405**(6785): p. 486-9.
34. Flavahan, W.A., et al., *Insulator dysfunction and oncogene activation in IDH mutant gliomas.* Nature, 2016. **529**(7584): p. 110-4.
35. Wouters, B.J., et al., *Distinct gene expression profiles of acute myeloid/T-lymphoid leukemia with silenced CEBPA and mutations in NOTCH1.* Blood, 2007. **110**(10): p. 3706-14.
36. Zabidi, M.A., et al., *Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation.* Nature, 2015. **518**(7540): p. 556-9.
37. Mansour, M.R., et al., *Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element.* Science, 2014. **346**(6215): p. 1373-7.
38. Erickson, P., et al., *Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt.* Blood, 1992. **80**(7): p. 1825-31.
39. Gamou, T., et al., *The partner gene of AML1 in t(16;21) myeloid malignancies is a novel member of the MTG8(ETO) family.* Blood, 1998. **91**(11): p. 4028-37.
40. Nucifora, G., et al., *Involvement of the AML1 gene in the t(3;21) in therapy-related leukemia and in chronic myeloid leukemia in blast crisis.* Blood, 1993. **81**(10): p. 2728-34.
41. Knezevich, S., et al., *Concurrent translocation of BCL2 and MYC with a single immunoglobulin locus in high-grade B-cell lymphomas.* Leukemia, 2005. **19**(4): p. 659-63.
42. Streubel, B., et al., *Molecular cytogenetic evidence of t(14;18)(IGH;BCL2) in a substantial proportion of primary cutaneous follicle center lymphomas.* Am J Surg Pathol, 2006. **30**(4): p. 529-36.
43. Groschel, S., et al., *A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia.* Cell, 2014. **157**(2): p. 369-381.
44. Pabst, T., et al., *AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia.* Nat Med, 2001. **7**(4): p. 444-51.

45. Ptasinska, A., et al., *Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding.* Leukemia, 2012. **26**(8): p. 1829-41.
46. Ptasinska, A., et al., *Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal.* Cell Rep, 2014. **8**(6): p. 1974-1988.
47. Sun, X.J., et al., *A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis.* Nature, 2013. **500**(7460): p. 93-7.
48. Wang, L., et al., *The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation.* Science, 2011. **333**(6043): p. 765-9.
49. Nimer, S.D. and M.A. Moore, *Effects of the leukemia-associated AML1-ETO protein on hematopoietic stem and progenitor cells.* Oncogene, 2004. **23**(24): p. 4249-54.
50. Davis, J.N., L. McGhee, and S. Meyers, *The ETO (MTG8) gene family.* Gene, 2003. **303**: p. 1-10.
51. Lindberg, S.R., et al., *Interactions between the leukaemia-associated ETO homologues of nuclear repressor proteins.* Eur J Haematol, 2003. **71**(6): p. 439-47.
52. Liu, Y., et al., *The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity.* Cancer Cell, 2006. **9**(4): p. 249-60.
53. Qiu, J., et al., *Decreased intranuclear mobility of acute myeloid leukemia 1-containing fusion proteins is accompanied by reduced mobility and compartmentalization of core binding factor beta.* Oncogene, 2006. **25**(28): p. 3982-93.
54. Rao, S.S., et al., *A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.* Cell, 2014. **159**(7): p. 1665-80.
55. Sanyal, A., et al., *The long-range interaction landscape of gene promoters.* Nature, 2012. **489**(7414): p. 109-13.
56. Gratz, S.J., et al., *CRISPR/Cas9-mediated genome engineering and the promise of designer flies on demand.* Fly (Austin), 2013. **7**(4): p. 249-55.
57. Li, T., et al., *TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain.* Nucleic Acids Res, 2011. **39**(1): p. 359-72.
58. Sanjana, N.E., et al., *A transcription activator-like effector toolbox for genome engineering.* Nat Protoc, 2012. **7**(1): p. 171-92.
59. Brocken, D.J.W., M. Tark-Dame, and R.T. Dame, *dCas9: A Versatile Tool for Epigenome Editing.* Curr Issues Mol Biol, 2017. **26**: p. 15-32.
60. Kwon, D.Y., et al., *Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC.* Nat Commun, 2017. **8**: p. 15315.
61. Franke, M., et al., *Formation of new chromatin domains determines pathogenicity of genomic duplications.* Nature, 2016. **538**(7624): p. 265-269.
62. Rao, S.S.P., et al., *Cohesin Loss Eliminates All Loop Domains.* Cell, 2017. **171**(2): p. 305-320 e24.
63. Taskesen, E., et al., *Prognostic impact, concurrent genetic mutations, and gene expression features of AML with CEBPA mutations in a cohort of 1182 cytogenetically normal AML patients: further evidence for CEBPA double mutant AML as a distinctive disease entity.* Blood, 2011. **117**(8): p. 2469-75.
64. Radomska, H.S., et al., *CCAAT/enhancer binding protein alpha is a regulatory switch sufficient for induction of granulocytic development from bipotential myeloid progenitors.* Mol Cell Biol, 1998. **18**(7): p. 4301-14.
65. Zhang, P., et al., *Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha.* Immunity, 2004. **21**(6): p. 853-63.
66. Guo, H., S. Cooper, and A.D. Friedman, *In Vivo Deletion of the Cebpa +37 kb Enhancer Markedly Reduces Cebpa mRNA in Myeloid Progenitors but Not in Non-Hematopoietic Tissues to Impair Granulopoiesis.* PLoS One, 2016. **11**(3): p. e0150809.

67. Ye, M., et al., *C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence.* Nat Cell Biol, 2013. **15**(4): p. 385-94.
68. Jaitin, D.A., et al., *Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types.* Science, 2014. **343**(6172): p. 776-9.
69. Jaitin, D.A., et al., *Each cell counts: Hematopoiesis and immunity research in the era of single cell genomics.* Semin Immunol, 2015. **27**(1): p. 67-71.
70. Hoppe, P.S., et al., *Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios.* Nature, 2016. **535**(7611): p. 299-302.
71. Paul, F., et al., *Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors.* Cell, 2015. **163**(7): p. 1663-77.
72. Yanez, A., et al., *IRF8 acts in lineage-committed rather than oligopotent progenitors to control neutrophil vs monocyte production.* Blood, 2015. **125**(9): p. 1452-9.
73. Tamura, T., *Regulation of mononuclear phagocyte development by IRF8.* Rinsho Ketsueki, 2017. **58**(7): p. 798-805.
74. Kurotaki, D., et al., *Transcription Factor IRF8 Governs Enhancer Landscape Dynamics in Mononuclear Phagocyte Progenitors.* Cell Rep, 2018. **22**(10): p. 2628-2641.
75. Faro, A., et al., *T-cell factor 4 (tcf7l2) is the main effector of Wnt signaling during zebrafish intestine organogenesis.* Zebrafish, 2009. **6**(1): p. 59-68.
76. Yu, Q., A. Sharma, and J.M. Sen, *TCF1 and beta-catenin regulate T cell development and function.* Immunol Res, 2010. **47**(1-3): p. 45-55.
77. Hong, C.S., et al., *Isolation of biologically active and morphologically intact exosomes from plasma of patients with cancer.* J Extracell Vesicles, 2016. **5**: p. 29289.
78. Hornick, N.I., et al., *AML suppresses hematopoiesis by releasing exosomes that contain microRNAs targeting c-MYB.* Sci Signal, 2016. **9**(444): p. ra88.
79. Hornick, N.I., et al., *Serum Exosome MicroRNA as a Minimally-Invasive Early Biomarker of AML.* Sci Rep, 2015. **5**: p. 11295.
80. Schofield, R., *The relationship between the spleen colony-forming cell and the haemopoietic stem cell.* Blood Cells, 1978. **4**(1-2): p. 7-25.
81. Yu, V.W. and D.T. Scadden, *Hematopoietic Stem Cell and Its Bone Marrow Niche.* Curr Top Dev Biol, 2016. **118**: p. 21-44.
82. Itkin, T., et al., *Distinct bone marrow blood vessels differentially regulate haematopoiesis.* Nature, 2016. **532**(7599): p. 323-8.
83. Yu, V.W.C., et al., *Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells.* Cell, 2016. **167**(5): p. 1310-1322 e17.
84. Scadden, D.T., *Blood and Bone.* N Engl J Med, 2016. **374**(19): p. 1891-3.
85. Hong, C.S., et al., *Isolation and characterization of CD34+ blast-derived exosomes in acute myeloid leukemia.* PLoS One, 2014. **9**(8): p. e103310.
86. Huan, J., et al., *RNA trafficking by acute myelogenous leukemia exosomes.* Cancer Res, 2013. **73**(2): p. 918-29.
87. Mass, E., et al., *A somatic mutation in erythro-myeloid progenitors causes neurodegenerative disease.* Nature, 2017. **549**(7672): p. 389-393.

A

THESIS_Roberto_Avellino.indd 198 07-05-18 10:18

ADDENDUM

THESIS_Roberto_Avellino.indd 199 07-05-18 10:18

THESIS_Roberto_Avellino.indd 200 07-05-18 10:18

ENGLISH SUMMARY

The **second chapter** of this thesis comprises a review that discusses novel concepts about the role of *CEBPA* in normal hematopoiesis and in disease. It focuses mainly on the many layers of transcriptional regulation of *CEBPA,* with particular interest on the function of the +42kb *CEBPA* enhancer in hematopoiesis. Hypothetical insights were discussed about the enhancer role in maintaining *CEBPA* levels at a steady-state and protect the bone marrow from undergoing HSC consumption, exhaustion and later bone marrow failure. Moreover, the +42kb enhancer can potentially be targeted by oncogenic mechanisms to minimize *CEBPA* levels down to a threshold that it's insufficient to drive the myeloid differentiation program and predispose myeloid committed progenitors to acute myeloid leukemia. In addition, adequate *CEBPA* expression levels are also required for the initiation of AML, indicating that *CEBPA* expression levels require tight regulatory mechanisms in bone marrow progenitors throughout the lifetime of an organism to protect from disease predisposition and initiation.

The **third chapter** describes thoroughly the non-coding region of the *CEBPA* locus in humans and mice to identify potential *CEBPA* enhancers in myelopoiesis. Combining different technical-approaches including active histone modifications, chromosome conformation capture and genome editing technology (CRISPR/Cas9 approach) revealed that a primary responsive enhancer in the *Cebpa* locus, located at +42kb and +37kb in humans and mice, respectively, activates *Cebpa* and the neutrophilic gene expression program. This enhancer engages with *CEBPA* in myeloid cells only, suggesting tissue specificity. Germ-line knockout of the enhancer in murine models compromises *Cebpa* expression and halts the neutrophilic differentiation program causing neutropenia. In addition, germ line enhancer knockout influences the HSC population, which leaves an open-end observation that requires further investigations. In conclusion, this study characterized an important *CEBPA* enhancer that has crucial implications in neutrophilic development and HSC integrity.

The **fourth** and the **fifth chapter** focus on the role of the non-coding regulatory region of the *CEBPA* locus in AML. A substantial number of AML patients exhibit low *CEBPA* expression levels compared to other AMLs with relatively higher CEBPA levels . Within these subsets are AML patient subsets with known hematological abnormalities such as the fusion oncoprotein AML1-ETO, ectopic high EVI1, FLT3-ITD, DNMT3, and a subset of AMLs with no recurrent abnormalities. For the latter group, the link to low *CEBPA* expression levels was hypothesized to occur via mutations in the non-coding genome of the *CEBPA* locus. However, screening 200 AML patients including, did not find any recurrent point mutations or chromosomal abnormalities, except for one patient that harbored a bi-allelic deletion of *CEBPA*. These findings lead to conclusions that the *CEBPA* locus in AML is not targeted by mutations but potentially deregulated at the chromatin level. The *CEBPA* locus was

investigated under the influence of AML1-ETO, the most common recurrent abnormality in AMLs with low *CEBPA* expression levels. AML1-ETO binds to the *CEBPA* locus via the +42kb enhancer, reverses its active chromatin state, disengage enhancer-gene interactions, down-regulate *CEBPA* expression and predispose bone marrow progenitors for the onset of acute myeloid leukemia.

In the **sixth chapter**, we tackled the unanswered question from **Chapter** three about the drop in HSC numbers in *Cebpa* enhancer-deleted and neutropenic mice. The main question addressed is two-way; (1) whether deleting the enhancer leads to the loss of the HSC sub-population that express *Cebpa* and shuts down the neutrophil differentiation program at the HSC stage, thus explained as a cell-autonomous effect; or (2) the neutropenia negatively influences the HSCs as an extrinsic effect in the absence of *Cebpa*. Using single-cell RNA-seq in combination with bulk RNA-seq in hematopoietic stem and progenitor cells, the most primitive HSC population does not express *Cebpa*, thus excluding that enhancer deletion influence HSCs in a cell autonomous way. The second hypothesis was tested using a transplantation model. Transplanted enhancer-deleted (neutropenic) bone marrow cells induced HSC loss and latent bone marrow failure in recipient mice, suggesting that neutropenic bone marrow cells compromise hematopoiesis by using, undefined, extrinsic mechanisms.

In conclusion, this thesis reveals new mechanisms of how a *Cebpa* enhancer acts as a primary switch to control neutrophilic differentiation and maintain bone marrow integrity. This enhancer is also a target for oncogenic deregulation of *CEBPA* expression in human AML, thus studying its function serves as a paradigm for epigenetic therapeutic targeting in the field of leukemia.

DUTCH SUMMARY

In dit proefschrift wordt de transcriptionele controle van *CEBPA* genexpressie tijdens hematopoëse bij muis en mens bestudeerd. Transcriptionele controle van genexpressie kent vele aspecten: dit proefschrift focust op de ontdekking van de rol van enhancers, die verantwoordelijk zijn voor de transcriptionele controle van *CEBPA* genexpressie in gezond en ziek beenmerg.

Hoofdstuk **twee** is een review, die nieuwe concepten over de rol van *CEBPA* in normale hematopoëse en bij ziekte bediscussieert. De focus van de review ligt op de nieuwe aspecten van transcriptieregulatie van *CEBPA* en met name op de functie van de +42kb *CEBPA* enhancer in hematopoëse. Hypotheses over de rol van de enhancer bij het onderhouden van de normale *CEBPA* expressieniveaus en over het beschermen van het beenmerg tegen consumptie van bloedstamcellen (HSCs), uitputting en uiteindelijk beenmergfalen worden hierin besproken. Verder wordt beschreven hoe de +42kb enhancer mogelijk kan worden gebruikt door oncogene mechanismen om de *CEBPA* expressie te minimaliseren, zodat de expressie onder de drempelwaarde ligt om de myeloïde differentiatie in gang te zetten. Hierdoor worden de myeloïde voorlopercellen vatbaar om acute myeloïde leukemie (AML) te ontwikkelen. Tevens is een minimaal *CEBPA* expressieniveau nodig voor de initiatie van AML, wat erop duidt dat de *CEBPA* expressie strikt gereguleerd moet worden in beenmerg voorlopercellen gedurende de gehele levensduur van een organisme om deze te beschermen tegen ziekte-aanleg en -initiatie.

In hoofdstuk **drie** wordt het niet-coderende gebied van de *CEBPA* locus in mens en muis in kaart gebracht om de mogelijke *CEBPA* enhancers in myelopoeise te identificeren. Door de combinatie van verschillende technieken, zoals analyse van actieve histon modificaties, chromosome capture en genoom-editing technologieën (CRISPR/Cas9), werd een primaire responsieve enhancer in het *Cebpa* locus geïdentificeerd, op +42kb bij de mens en +37kb bij muizen. Deze activeert *Cebpa* en het neutrofiele genexpressieprogramma. Deze enhancer werkt alleen in myeoloïde cellen, wat een weefselspecifieke rol suggereert. Een knock-out van de enhancer in de germ-line in muizenmodellen vermindert *Cebpa* expressie en blokkeert het neutrofiele differentiatie programma, waardoor er neutropenie ontstaat. Verder beïnvloedt de germ-line enhancer knock-out de HSC-populatie, De mechanismen hierachter zijn onduidelijk en worden verder bediscusieerd in Hoofdstuk 6. Samenvattend: deze studie beschrijft een belangrijke *CEBPA* enhancer met een essentiële rol voor neutrofiele ontwikkeling en HSC-integriteit.

Hoofdstuk **vier** en **vijf** focussen op de rol van de niet-coderende regulerende regio van de *CEBPA*-locus in AML. Een substantieel deel van de patiënten met AML heeft een relatief lage

CEBPA-expressie. In deze groep bevinden zich AML-patiënten met bekende hematologische afwijkingen, zoals het fusie-eiwit AML-ETO, abnormaal hoge EVI1, FLT3-ITD, DNMT3, maar ook een subgroep zonder bekende afwijkingen. Voor de laatste groep was de hypothese dat het lage *CEBPA* expressieniveau door mutaties in het niet-coderende genoom van het *CEBPA* locus zou kunnen worden veroorzaakt. Een screening van 200 AML-patiënten, leverde geen veelvoorkomende puntmutaties of chromosomale afwijkingen op, met de uitzondering van één patiënt waarbij een bi-allelische deletie van het *CEBPA*-locus werd gedetecteerd. Deze bevindingen leiden tot de conclusie dat het *CEBPA*- locus in AML geen doelwit is voor mutaties, maar mogelijk wordt gedereguleerd op het niveau van chromatine. Verder werd de invloed van AML1-ETO, de meest voorkomende afwijking bij AML met lage *CEBPA*-expressie, bestudeerd. AML1-ETO bindt aan het *CEBPA* locus via de +42k enhancer, maakt de actieve chromatine status ongedaan, inhibeert enhancer-gen interacties, zorgt voor lagere *CEBPA*-expressie en maakt de voorlopercellen in het beenmerg vatbaar voor de ontwikkeling van acute myeloïde leukemie.

In hoofdstuk **zes** wordt de nog open vraag uit hoofdstuk **drie** beantwoord, over de vermindering van HSC aantallen in *Cebpa* enhancer-gedeleteerde en neutropene muizen. De hoofdvraag leidt in twee richtingen: Ten eerste, is de deletie van de enhancer intrinsiek verantwoordelijk voor de vermindering van de HSC subpopulatie? In dit geval gaan we er van uit dat HSCs *Cebpa* tot expressie brengen en dat de neutrofiele differentiatie al wordt geblokkeerd in de HSCs. Een cel-autonoom effect zou dan het verlies van de aantallen HSCs verklaren. Of wordt, ten tweede, de HSC populatie negatief beïnvloed een extrinsiek effect als gevolg van de neutropenie veroorzaakt door de afwezigheid van *Cebpa*?

Uit RNA-seq van hematopoïetische stam-/voorloper-cellen uitgevoerd op individuele cellen in combinatie met RNA-seq van meerdere cellen, bleek dat de meest primitieve HSC-populatie geen *Cebpa* tot expressie brengt. Dit sluit uit dat de enhancer-deletie de HSC op een cel-autonome manier beïnvloedt. De tweede hypothese werd getest door middel van een transplantatie model. Enhancer gedeleteerde (neutropene) beenmergcellen werden getransplanteerd in muizen, wat leidde tot een afname van HSC-cellen en latent beenmergfalen bij deze muizen. Dit suggereert dat de neutropene beenmergcellen de hematopoëse verstoren via nog onbekende, extrinsieke mechanismen.

Samenvattend: dit proefschrift beschrijft nieuwe mechanismen waardoor de *Cebpa*-enhancer als een schakelaar de neutrofiele differentiatie en daarbij beenmergintegriteit aanstuurt. Deze enhancer is ook een target voor de oncogene deregulatie van *CEBPA*-expressie in menselijke AML, zodat het bestuderen van de functie van deze enhancer kan fungeren als een paradigma voor het zoeken naar epigenetische therapeutische strategieën bij leukemie.

LIST OF ABBREVIATIONS

3C	Chromatin Conformation Capture
4C-seq	Circularized Chromatin Conformation Capture
AML	Acute Myeloid Leukemia
BM	Bone marrow
PB	Peripheral Blood
BRD4	Bromodomain-4
CEBPA	CCAAT enhancer binding protein alpha
CEBPB	CCAAT enhancer binding protein beta
CEBPD	CCAAT enhancer binding protein delta
CEBPE	CCAAT enhancer binding protein epsilon
CEBPG	CCAAT enhancer binding protein gamma
CEBPZ	CCAAT enhancer binding protein zeta
CHD7	Chromodomain 7
CMP	Common myeloid progenitors
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTCF	CCCTC binding factor
DNA	Deoxyribonucleic acid
DNAse1 - HS	DNAse1 Hypersensitive
FACS	Fluorescence Activated Cell sorting
GMP	Granulocytic and Monocytic progenitors
GTF	General transcription factor
H3K27ac	Histone 3 lysine 27 acetylation
H3K27me3	Histone 3 lysine 27 tri-methylation
H3K9ac	Histone 3 Lysine 9 acetylation
HAT	Histone acetyl transferases
HDAC	Histone deacetyl transferases
HSC	Hematopoietic Stem cell
HSPC	Hematopoietic Stem and Progenitor Cell
KMT	Lysine Methyl transferase
LK	Lineage negative C-Kit positive
LMPPs	Lymphoid primed multi-potent progenitors
LSK	Lineage negative C-Kit positive Sca-1 Positive
LT-HSCs	Long term hematopoietic stem cells
LTF	Lineage specifc transcription factors
MEP	Megakaryocytic Erythroid progenitors
MLL	Mixed lineage leukemia
MPP	Multi-potent progenitor

PCR	polymerase chain reaction
PolII	RNA polymerase II
RNA	Ribonucleic acid
ST-HSCs	Short-term hematopoietic stem cells
TAD	Topological associated domains
TALENS	Transcription activator like effector nucleases
TBP	TATA binding protein
TF	Trasncription factor
TSS	Trasnscriptional start site

CURRICULUM VITAE

Roberto Avellino was born in Malta on the 25th of December 1980. After receiving his foundation Diploma in Pathology(1999-2003), followed by a Bachelors in Biomedical Sciences from the university of Malta (2003-2006), he worked in the clinical haematology laboratory at the general hospital of Malta, St.Luke's hospital. There he worked on routine and molecular diagnostics of acute and chronic leukemias and, in addition, he was involved in a project under the name of 'Molecular Pathology' in collaboration with Godfrey Grech at the University of Malta. The aim of this project was to implement new technologies locally that could improve the diagnosis of patients with leukemias. During this time, Roberto wanted to further his horizons and dig further in the molecular field of hematological malignancies. He applied for competitive funding from the European Union, and once he got the funds, he decided to work in a lab abroad. Working in a diagnostic haematology lab in Dublin (Ireland) for a period of 6 months, he studied recurrent gene mutations in an aggressive form of leukemia known as acute myeloid leukemia under the supervision of Mark Catherwood (2006). After leaving the lab in Ireland, Roberto immediately decided to enrol himself in a Master's programme at the University of Newcastle (United Kingdom). Working on mapping breakpoints using BAC clones in combination with fluorescent in situ hybridization chromosomal abnormalities involving the light chains of the immunoglobulin receptor in acute lymphoblastic leukemia, Roberto obtained his Masters degree in Medical and Biomedical Sciences with Distinction under the supervision of Christine Harrison (2009). From then onwards, Roberto joined the group of Ruud Delwel for a PhD program at the Erasmus Medical Centre in Rotterdam (The Netherlands), and he studied how high-order chromatin structures in the cell nucleus influence the transcription control of a gene known as *CEBPA* in the healthy bone marrow and in disease.

THESIS_Roberto_Avellino.indd 207 07-05-18 10:18

LIST OF PUBLICATIONS

1. M. Alberich-Jorda, B. Wouters, M. Balastik, C. Shapiro-Koss, H. Zhang, A. Di Ruscio, H. S. Radomska, A. K. Ebralidze, G. Amabile, M. Ye, J. Zhang, I. Lowers, **R. Avellino**, A. Melnick, M. E. Figueroa, P. J. Valk, R. Delwel, and D. G. Tenen, 'C/Ebpgamma Deregulation Results in Differentiation Arrest in Acute Myeloid Leukemia', J Clin Invest, 122 (2012), 4490-504.

2. **R. Avellino**, and R. Delwel, 'Expression and Regulation of C/Ebpalpha in Normal Myelopoiesis and in Malignant Transformation', Blood, 129 (2017), 2083-91.

3. **R. Avellino**, M. Havermans, C. Erpelinck, M. A. Sanders, R. Hoogenboezem, H. J. van de Werken, E. Rombouts, K. van Lom, P. M. van Strien, C. Gebhard, M. Rehli, J. Pimanda, D. Beck, S. Erkeland, T. Kuiken, H. de Looper, S. Groschel, I. Touw, E. Bindels, and R. Delwel, 'An Autonomous Cebpa Enhancer Specific for Myeloid-Lineage Priming and Neutrophilic Differentiation', Blood, 127 (2016), 2991-3003.

4. J. L. Glass, D. Hassane, B. J. Wouters, H. Kunimoto, **R. Avellino**, F. E. Garrett-Bakelman, O. A. Guryanova, R. Bowman, S. Redlich, A. M. Intlekofer, C. Meydan, T. Qin, M. Fall, A. Alonso, M. L. Guzman, P. J. M. Valk, C. B. Thompson, R. Levine, O. Elemento, R. Delwel, A.Melnick, and M. E. Figueroa, 'Epigenetic Identity in Aml Depends on Disruption of Nonpromoter Regulatory Elements and Is Affected by Antagonistic Effects of Mutations in Epigenetic Modifiers', Cancer Discov, 7 (2017), 868-83.

5. F. Grebien, M. Vedadi, M. Getlik, R. Giambruno, A. Grover, **R. Avellino**, A. Skucha, S. Vittori, E. Kuznetsova, D. Smil, D. Barsyte-Lovejoy, F. Li, G. Poda, M. Schapira, H. Wu, A. Dong, G. Senisterra, A. Stukalov, K. V. M. Huber, A. Schonegger, R. Marcellus, M. Bilban, C. Bock, P. J. Brown, J.Zuber, K. L. Bennett, R. Al-Awar, R. Delwel, C. Nerlov, C. H. Arrowsmith, and G. Superti-Furga, 'Pharmacological Targeting of the Wdr5-Mll Interaction in C/Ebpalpha N-Terminal Leukemia, Nature Chemical and Structural Biology, 2015 Oct;11(10):815.

PHD PORTFOLIO

Name : Roberto Avellino
PhD period : July 2010-May2017
Erasmus MC Department: Hematology

	Year
Courses and workshops	
Research Management for PhDs	2010
Photoshop and Illustrator (Basic)	2012
Basic Course on R	2013
Writing Grant proposals	2014
Basic introduction course on SPSS	2015
Ingenuity Pathway Analysis Workshop	2015
Photoshop and Illustrator (Advanced)	2018
Workshop 'Molecular aspects of Hematological malignancies'	2014-2018
Scientific Meetings Department of Hematology	
Work discussion (Weekly)	2010-2017
Erasmus Hematology lectures (Monthly)	2010-2017
AIO/post-doc (Monthly)	2010-2015
Journal club (bi-monthly)	2010-2017
National/International conferences	
Chromatin and Epigenetics, EMBO, Heidelberg, Germany	2013
Molecular aspects of Hematological Malignancies (oral)	2014
American Society of Hematology, San Francisco, US (poster)	2014
Molecular aspects of Hematological Malignancies (oral)	2015
American Society of Hematology, Orlando, US (oral)	2015
Tumor Cell biology meeting KWF meeting, The Netherlands	2015
Molecular Medicine day, Rotterdam, The Netherlands (poster)	2016
Modern Trends in Leukemia and Cancer, Wilsede, Germany (poster)	2016
Superivising and organisation activities	
Supervising Master Internship (Stefan Groeneweg)	2011 (6months)
Supervising Master Internship (Rob Swenker)	2011-2012
Supervising Master Internship (Lena Kourkouta)	2013-2014
Supervising Master Internship (Muriel Kuipers)	2014-2015
Organization and supervision PhD lunch with invited speakers	2011-2012

THESIS_Roberto_Avellino.indd 210 07-05-18 10:18

WORD OF THANKS

Starting and finishing a PhD, feels like a long 360 degrees ride taken on a roller coaster. This long journey now comes to an end; a journey that strengthened and prepared me both scientifically and as a person to reach out and be exposed for what is next in my personal life and in my scientific career. This journey couldn't be accomplished without the contribution and help of several people from inside and outside the lab. I will take some time to thank all of you, including family, friends and colleagues.

The person I would like to thank the most is my supervisor Ruud Delwel. Dear **Ruud**, for all these years, through thick and thin, you were patient enough to give me the time I needed for myself to grow and develop scientifically. Starting a whole program from scratch and developing new projects together to make this whole journey happen was not the easiest task at all. It taught me how hard it is to survive in science and how much time it requires from one's personal life. However, you also taught me not to give up once it becomes a passion. Your excellence in networking and negotiating scientific business makes your research of a high impact level. This is a huge asset for yourself and the people working with you. Because of this, I was exposed to a numerous leading scientific researchers in the hematology field. This exposure also helped me to choose the next step in my career. All these years working together, did not only build my scientific capabilities, but helped me develop as a person. I was honored to be present when you won the Dutch prize (KWO), the highest award a scientist can get in the cancer field in The Netherlands. Outside science, you are also a dedicated family man that, although having a big family, you are always there for them providing care and support. It was a real pleasure to meet your wife Wilma, your sons and your daughter. Thank you for everything!

I am honored that you, Prof.dr. Ivo Touw, accepted to be my promoter. Dear **Ivo**, I am highly fond of your vast scientific knowledge, fond of the time you take to help other people in the whole department to develop their scientific projects and your constant participation in discussions during floor meetings. I am honored that I had the opportunity to be there when you gave the 'Game of Clones' lecture for the Ham-Wasserman Award at the American Society of Hematology Congress 2015. Such a prize would only be achieved by scientists that reach your scientific level. We had many fruitful discussions about our fields of research and also about science in general. It was a great pleasure to get to know you and I truly appreciate the time you took to give me advice about my projects and my career development.

To the members in the small scientific committee, Prof.dr. **Marc** Raaijmakers, Prof.dr. **Jan**. J Schuringa, Prof.dr. **Peter** Verrijzer, thank you for thoroughly reading my thesis and for your suggestions. To the other members of the committee, Prof. **Bo** Porse, dr.**Peter** Valk and Prof. dr. **Wouter** de Laat, I am very grateful to have you in the committee of my thesis defense.

Dear **Claudia**, you were the first person that approached me when I came to Ruud's lab. I am truly grateful that I had the opportunity to work with you throughout all these years. You are highly valuable in Ruud's group. Being meticulous, professional and highly experienced, makes you an outstanding scientist that every group wishes to have as a member. You were vital for the whole *CEBPA* program. Without you this could have never been fully accomplished. Apart from being a great scientist you are also a very kind person, and because of this, you were my first choice to be beside me during my thesis defence as my paranymph. Thanks for everything!

Dear **Eric**, as you clearly remember, the beginning of this journey was tough for me to settle down in the lab and find the line of research that fit me most. Despite the difficulties, I am still grateful for the patience that you showed when I started. You helped me overcome difficult times and on several occassions you were also a point of reference when I asked for your opinion about experiments and new ideas. I understand that your position was not the easiest one, which sometimes also caused some havoc. However, one way or another we managed to go through it all. Together with the expertise of **Marije**, we worked together to accomplish the mouse studies for the Blood paper, and I am truly grateful for that. I would like to thank **Marije** for the time she dedicated to this project and contributed effectively to make it for publication.

I wish all the best for the 'new' recruited members in Ruud's group for the EVI1 program including **Tim**, **Leonie**, **Sophie** and **Andrea**. Tim, your recent publication in New England Journal of Medicine is of huge impact for your career. I wish you have more similar achievements and I hope that you reach your desired milestones. Leonie, Sophie and Andrea; I am sure you will make a good team together and be able to unravel your question with a valuable publication. **Nelleke**, as a new member in our group, I wish you all the best for your future. It's a flight of stairs that you have to take to reach the end of it, just take one step at a time!

I would like to thank former members in Ruud's group including **Bas**, **Stefan** and **Veronika**. Dear **Bas**, I took over the work you started some years ago about *CEBPA* in AML patients. You are a dedicated scientist, a humble and very pleasant person. It's always a pleasure to work with you and I wish you success in your scientific career. **Stefan**, you were a person to learn from scientifically and you were vital for the EVI1 project. The publication in Cell put you and Ruud at a high rank in the scientific community. I admire what you have already achieved in your scientific career. I wish you all the best with your future! **Veronika**, your achievements are nonetheless outstanding. Becoming an associate professor at such a young age makes me proud that I had the opportunity to work with you and also to have you as a close friend, especially in very difficult times.

I started my PhD together with 7 other PhD candidates, **Farshid**, **Kasia**, **Davine**, **Patricia** , **Roel**, **Noemi** and **Lucia**. One way or another we all made it through this whole journey and I feel honored that you were part of it. I am sure we will meet again later in our careers and be able to collaborate together. I wish all of you the best of luck to your future. To the current PhDs and post-docs, **Ping, Keane, Adil, Jess, Maurice, Emanuele, Cansu, Burak, Eline, Patricia O., Nils** and **Helene**, I wish you reach your career goals successfully.

Dear **Renee**, I highly admire you as a person as well as a scientist. Your dedication and commitment to science never failed to impress me. Even while taking care of three adorable kids, you manage to guide your research in the right direction. We kept our strong friendship even though we live in different countries. Your scientific advice is always valuable to me. I wish you all the best and hope we will collaborate soon.

To the bioinformaticians, **Mathijs**, **Remco**, **Roger** and **Elodie**. You are highly valuable for the department and your contribution to my work was utterly important. I express gratitude for all the times I just ran into your room asking for help and without hesitation, you were always ready to answer my questions. I wish you success in your careers and personal lives. Dear **Harmen**, it's an honor to get to know you and would like to thank you for your great help in analysing our 4C-seq data.

Dear **Stefan** and **Hans**. I am grateful for all discussions we had and the time you dedicated to discuss the CRISPR mouse studies. Stefan, your mentorship to Farshid was impressive and I believe you are now highly established at the Immunology department. Dear **Rebekka**, your kind help with the murine histology preparations and examinations was crucial for the neutropenia project. Your achievements are highly admirable, and you are considered as a role model for each and everyone of us. Dear **Emma**, as young PI, I believe you have a promising future ahead. I would like to thank you for all the fruitful discussions we had in the lab. I also wish to thank the other PIs for discussions and advice during the floor meetings including **Bob, Peter Valk, Eric, Jan, Pieter Sonneveld, Anita, Tom, Mojca** and **Frank**.

Dear **Lisa**, it was a great pleasure for having you as my mentor during my Masters degree project at Newcastle University, UK. Even back then, you already showed competent skills as an independent scientist and in managing members in your group. Your support before my first-time talk at ASH 2015 was vital! I wish you all the best for your future career and in your personal life. From the University of Munster, **Daniel,** I feel grateful for discussing projects and contemplating how to progress further with new possibilities for 4C-seq analysis. You are a great scientist and I am sure that you will succeed in your career. From the University of Regensburg, **Claudia**, you are a collaborator and a friend that I will always treasure!

I would like to thank the administrative stuff at our department **Lenke**, **Ans** and **Tessa** and to **Egied** for his great help to adjust and organize the layout of the whole thesis.

Thank you **Stefan**, **Rob**, **Lena** and **Muriel**. Working with masters students taught me a lot about science in general but also about myself. You were la bearing that showed me where I stand in the field and helped me to see gaps which are rather obscured when tackled by oneself. Thank you all, you helped me to take my program a step further. Dear **Elwin** and **Michael**, your excellence in cell sorting experience was vital to our projects. I appreciate your work and contribution. To all the other members in the haematology department, I wish to thank everyone for all the contribution that you provided, one way or another. I would like also to thank former lab members **Arturo, Tania, Rasti, Erik, Suming, Simone, Shirley, Piotr, Szaby, Julia, Joyce, Julienne**.

The 13th floor outside EMC : Dear **Kasia, Jana, Keane, Adrian, Onno, Paulette, Noemi, Monica, Marshall, Si Chen, Julienne, Adil, Ping, Tania, Alex, Mira, Emmanule, Cansu**. I don't know from where to start. We experienced very good times together, which buffered all the tension that accumulated at work. We managed to see famous DJs and bands playing in local festivals. Radiohead and Nils Frahm hit the list (I am called a groupie for attending 8 Nils Frahm concerts in a period of 4 years, and it's never enough). Although at times my reserved personality could be mistaken as detachment, I have no words to describe how grateful I am for your constant care and support. **Kasia**, we built more than a friendship based on trust and honesty, and that is the reason why you are my paranymph. I hope that we all manage to keep connected, even though moving to different countries!

My interest in hematology dates back to 2001. When I was a bachelor's student I was introduced to hematology by **Dr. Alicia Grochowska**, Miss **Agnes Saunders** and Miss **Antionette Mifsud**. You accepted me to be part of your team. Without you I would have never reached my career goals. I consider you as seniors and mentors that introduced me to this highly explorative field of hematology. **Sherif**, you were the first person to show me a stained bone marrow slide of a 5-year-old kid with pre-B acute lymphoblastic leukemia. You also exposed me to the field of Molecular Hematology more than 15 years ago. From that time onwards, I got intrigued by what appeared on that slide under the microscope as big, dark, ugly cells called leukaemic blast cells. I wanted to learn more about what was going on inside those cells, and since then, I involved myself enough and drifted away from our tiny island to widen horizons in my career. It was a pleasure to work with haematology consultants **Dr.D.Camilleri** and **Dr.A.Gatt**, as well as **Patricia, Michael, Denis, Elton, Dieter, Neville, Mario, Marvic, Jacklyn, Silvana, Aldo, Darren, Margaret, Patrick, Josielle, Silvana, Ian, Lisa, Justin, Mario, Michelle** and **John**.

To the University of Malta, I would like to thank **Prof.A.Cuschieri** and **Prof.A.Xuereb** for mentoring me during my Bachelor's degree. I am devotedly thankful for **Prof.G.Grech** for introducing me to the field of molecular biology. And also would like to thank other colleagues and friends **Sean, Chris, Robert, Melissa, Joseph** and **Stef**.

Outside the lab, I want to show gratitude to my parents **Joseph** and **Louise**, my aunt **Carmen**, my sister **Katya**, her husband **Adrian**, and my niece and nephew **Iulia** and **Andreas**. Your love and support throughout these years were unconditional. Despite the distance, you kept on supporting me. Dear **Alexia**, we have known each other for almost 14 years and I am grateful for your constant support to accomplish my PhD throughout these years. You are a smart, caring, and loving person. Although we might soon be living in different countries, you will still be my soul-mate. Dear **Ryan**, we have known each other back from secondary school, about 26 years ago, and our friendship grew stronger by time. I always saw you as an example to all of us. Smart, diligent, fair, humble and correct. It was a great pleasure to be your best-man on your wedding day with your amazing wife **Sinead**. Despite we rarely see each other, you are still my brother and I am so happy that we shared so much experiences together! Dear **Andrea** (iz-Zuli), the sharp, smart gentleman, you are family to me. Your constant visits to The Netherlands made things much simpler for me. We re-lived the 90s once again here in The Netherlands with many music events and festivals. Can never forget kilometres of walking to Kralingse Bos and back. Dear **Elton** and **Lorraine**, you have been a point of reference from the first day till the end of this journey. As you always said, 'go and get what's yours'. Your words of wisdom stayed with me all along the journey and were my asset to get me going when the sea got rough! Dear **Matthew**, we are always told that we look alike, although my pale skin color compared to your dark skin. Though we are cousins, we were raised up together as brothers and shared many great things together. Despite the distance that separated us these last number of years, you are still my younger brother that I dearly miss. Dear **Mikiel** and **Diane**, you were an important part of this journey. Overnight bbqs on your terrace eating abundantly and tasting homemade wine, priceless! Your journey took a different lead, from dating to getting married to having your first adorable child **Jude**. **Michael**, our friendship dates back to 18 years ago when we started University, old times that can never be forgotten. Dear **James** and **Denis**, our endless discussions over a pint of beer or more, is what I look forward to from time to time. Your intelligence and mind-power always fascinated me, and it's always a pleasure having you around. Last but not least, I would to give my regards to all my other friends **Rosanne**, **Ray, Chris, Jean** and **Terri, Ian** and **Jody, Ian** and **Daniela, Nils and Maike, David, Giulian** and **Steve.** Even though we don't frequently meet, your friendship stays there forever!

THESIS_Roberto_Avellino.indd 216 07-05-18 10:18